

Reliable Density Estimates for Coverage and Connectivity in Thin Strips of Finite Length

Paul Balister[†] Béla Bollobás[†] Amites Sarkar[†] Santosh Kumar[‡]

[†] Dept. of Mathematical Sciences
University of Memphis
Memphis, TN 38152, USA

{pbalistr,bollobas}@memphis.edu, amites.sarkar@gmail.com

[‡] Dept. of Computer Science
University of Memphis
Memphis, TN 38152, USA

santosh.kumar@memphis.edu

ABSTRACT

Deriving the critical density (which is equivalent to deriving the critical radius or power) to achieve coverage and/or connectivity for random deployments is a fundamental problem in the area of wireless networks. The probabilistic conditions normally derived, however, have limited appeal among practitioners because they are often asymptotic, i.e., they only make high probability guarantees in the limit of large system sizes. Such conditions are not very useful in practice since deployment regions are always finite. Another major limitation of most existing work on coverage and connectivity is their focus on thick deployment regions (such as a square or a disk). There is no existing work (including traditional percolation theory) that derives critical densities for thin strips (or annuli).

In this paper, we address both of these shortcomings by introducing new techniques for deriving reliable density estimates for finite regions (including thin strips). We apply our techniques to solve the open problem of deriving reliable density estimates for achieving *barrier coverage* and connectivity in thin strips, where sensors are deployed as a barrier to detect moving objects and phenomena. We use simulations to show that our estimates are accurate even for small deployment regions. Our techniques bridge the gap between theory and practice in the area of coverage and connectivity, since the results can now be readily used in real-life deployments.

Categories and Subject Descriptors

C.2.1 [Computer-Communication Networks]: Network Architecture and Design—*network topology, wireless communication*; G.3 [Mathematics of Computing]: Probability and Statistics—*Stochastic Processes*

General Terms

Theory

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MobiCom '07, September 9–14, 2007, Montréal, Québec, Canada.
Copyright 2007 ACM 978-1-59593-681-3/07/0009 ...\$5.00.

Keywords

Wireless sensor networks, barrier coverage, coverage, connectivity, deployment, network topology, random geometric graphs, critical conditions.

1. INTRODUCTION

Deploying sensors deterministically is an expensive undertaking in terms of time, effort, and money. As the cost of wireless sensors continues to fall, the expense of deterministic deployment may become prohibitive (when compared with the cost of sensors), making random deployment a compelling alternative. When deploying sensors in inaccessible terrain (e.g., forests, mountains, enemy regions), deterministic deployment may not even be an option. Even when sensors have been deployed deterministically, significant errors in deployment, movement of sensors after deployment due to wind, unanticipated failure of sensors, or probabilistic activation of sensors, may all make the distribution of sensor locations close to a Poisson process as shown in [15]. For these and other reasons, the Poisson process is used in the majority of analytical work on coverage and connectivity in both wireless ad hoc and sensor networks [2, 7, 8, 9, 12, 13, 14, 15, 17, 18, 19, 21, 23, 24, 26]. Random deployment is also the model of choice in simulations when studying the performance of topology control and other algorithms [6, 16, 22, 25].

Deriving the critical density (which is equivalent to deriving the critical power or radius) needed to achieve coverage and/or connectivity is a fundamental problem. It can be used to predict the behavior of a network with regard to coverage and/or connectivity. At a density *lower than critical*, with high probability, the network *does not* provide coverage (and/or connectivity), and, at a density *higher than critical*, with high probability, the network *does* provide coverage; hence, the term critical. Such conditions, however, are asymptotic in nature. Since deployment regions are always finite in real life, such conditions are not too useful in practice.

Another major limitation of most existing work is that they limit themselves to thick deployment regions such as disks and squares. These results are not applicable to thin strips. The fact that percolation does not occur for thin regions is often cited as a primary reason for avoiding thin strips [8]. The results that do consider rectangles place a lower bound on the width to length ratio [19]. When sensors are deployed in thin strips, such as when deploying along international borders to detect intrusion, or around forests to

detect fire, no existing work can be used to derive the density of sensors needed for achieving coverage and/or connectivity.

If the goal of deploying sensors is to detect moving objects and phenomena (which is often the case in thin strip deployments), then the model of *barrier coverage* [14] may be a more appropriate model than *full coverage*, which is the focus of all existing work except [14]. Barrier coverage ensures that no moving object or phenomenon can cross the barrier of sensors without being detected, whereas full-coverage ensures that every point in the deployment region is covered. Further, when sensors are deployed for barrier coverage, achieving *s-t connectivity*, which ensures that a connected path exists between the two far ends of a thin strip, may be more appropriate than achieving *full connectivity*, which requires that every sensor be connected to every other sensor. The fact that some sensors may not be connected to the base station does not compromise the barrier coverage guarantee; all events can still be detected and communicated to the base station(s) even if the base station(s) is located at a far end.

The problems of deriving densities needed for achieving barrier coverage and/or *s-t connectivity* in thin strips (even asymptotic conditions) are open [14]. In addition, no existing work can be used to provide reliable density estimates for achieving full coverage and full connectivity in thin strips even though these models have been extensively studied for thick regions.

In this paper, *we derive reliable estimates for the density needed to achieve coverage and connectivity in thin strips for all four models of coverage and connectivity*. We develop a novel definition of *break* (a disruption in connectivity) that is critical in solving the problems of barrier coverage, *s-t connectivity*, and full connectivity in thin strips, all three of which are harder than full coverage. We use of the concept of covering boundary intersection points [11] in deriving a density estimate for full coverage. Our derivation is much simpler, yet far more accurate than any existing work.

As far as is possible, we avoid taking limits (which is pervasive in existing work) in all our derivations to come up with reliable estimates. We demonstrate the accuracy of our estimates for finite regions through extensive simulations. We intentionally use small regions in our simulations to show that, unlike existing work, the system size does not need to be large for our estimates to be accurate. Consequently, our estimates can be readily used in real-life deployments (including small regions such as perimeters of secure facilities) for determining the density of sensors needed, thus bridging the long-standing gap between theory and practice in the area of coverage and connectivity.

Organization: Section 2 describes the model and problem formulations. Section 3 presents our key results, main contributions, and some more related work. Section 4 presents our novel definition of *breaks* which makes it possible to derive reliable density estimates for coverage and connectivity in thin strips. We also establish in this section that break occurrence approximately follows a Poisson distribution. Section 5 describes heuristics that can be used to guess the density estimates. Section 6 presents the derivation of the break intensity expected in a thin strip. Once the break intensity is known, the density needed to achieve barrier coverage (or *s-t connectivity*) with a desired probability can be readily derived using the properties of the Poisson process. Simulation results, which show the reliability of our estimates,

appear in Section 7. Sections 8 and 9 present derivations of reliable density estimates for full connectivity and full coverage. Section 10 concludes the paper and mentions some future work. Due to space limitations, we do not include full proofs of our results here. For more detailed proofs see [1].

2. MODEL AND PROBLEM DEFINITIONS

Consider an infinite strip $S_h = \mathbb{R} \times [0, h]$ of width (or height) h . We place sensors inside this strip randomly according to a Poisson process \mathcal{P} of intensity $\lambda > 0$. Thus for any region $A \subseteq \mathbb{R}^2$, the number of sensors in A is given by a Poisson variable of mean $\lambda|A \cap S_h|$, where $|A \cap S_h|$ is the area of the region $A \cap S_h$. Moreover, the number of sensors in A is independent of the number of sensors in any disjoint region. We shall assume that each sensor has the ability to detect intruders within a certain distance r_s , and transmit information to other sensors within a (possibly unrelated) distance r_t . We note that for a given sampling frequency of sensors, and a given startup latency, a sensing range exists such that all events of interest within this sensing region are detected by this sensor with very high probability.

Construct an infinite random geometric graph $G_{h,r,\lambda}$ with vertex set given by the set of sensors \mathcal{P} , by joining every sensor to every other sensor that is strictly within (Euclidean) distance $r \in \{2r_s, r_t\}$. For $a < b$, define $G_{h,r,\lambda}(a, b)$ to be the subgraph of $G_{h,r,\lambda}$ consisting of sensors with x -coordinate between $a - \frac{r}{2}$ and $b + \frac{r}{2}$, and with two extra sensors s and t , where s is joined to all sensors of \mathcal{P} within distance $\frac{r}{2}$ of the line $x = a$ and t is joined to all sensors of \mathcal{P} within distance $\frac{r}{2}$ of the line $x = b$ (see Figure 1). Such a graph is called a *Coverage Graph* in [14]. We use $G_{h,r,\lambda}(a, b)$ to address two problems.

The first is *barrier coverage* [14]. We wish to detect moving objects or phenomena (such as intruders) crossing the strip, so we wish to know if there is a path contained within the region $[a, b] \times [0, h]$ from the top to the bottom of the strip which does not pass within a distance of $\frac{r}{2} = r_s$ of any sensor. We call such a path a *separating path*. It is clear that if such a path exists, then this path disconnects the graph $G_{h,r,\lambda}(a, b)$, since no point to the left of the path is within distance r of a point to the right of the path. The converse however is not true. There may be many small components of $G_{h,r,\lambda}(a, b)$ which do not result in such a path. However, if no separating path exists in $[a, b] \times [0, h]$, then we have *s-t connectivity* in $G_{h,r,\lambda}(a, b)$, i.e., there is a path from s to t in the graph $G_{h,r,\lambda}(a, b)$ (see Figure 1).

The second problem is to determine whether the sensors that provide barrier coverage can relay information along the strip. Suppose that every sensor within $\frac{r_t}{2}$ of the line $x = a$, say, can relay information to a base station. Then it is enough that $G_{h,r,\lambda}(a, b)$ is *s-t connected* where $r = \min\{\frac{r_s}{2}, r_t\}$. Indeed, the sensors along any *s-t path* in the graph $G_{h,r,\lambda}(a, b)$ give barrier coverage, and one can relay information along this path. On the other hand, if $G_{h,r,\lambda}(a, b)$ is not *s-t connected*, then either $r = \frac{r_s}{2}$ and there is no barrier coverage, or $r = r_t$, and there is no communication path from s to t using any of the sensors. Note that even if $G_{h,r,\lambda}(a, b)$ is *s-t connected*, there may be many small components of the network that cannot communicate far along the strip (see Figure 1 and Figure 2), however, for the purposes of detecting movements, these small components are not important, since removing these sensors would not allow any moving object (or phenomena) to cross undetected.

If, on the other hand, each point in $R = [a, b] \times [0, h]$ is covered by at least one sensor, then the region R is said to be *fully covered*, and if the vertices of $G_{h,r,\lambda}$ that lie in R form a single connected component, then the network in R is said to be *fully connected*.

Although it is sufficient to ensure barrier coverage and s - t connectivity when the goal of sensor deployment is to detect movements, other applications (such as fine-grained tracking) may need full coverage and/or full connectivity, which are stronger guarantees than barrier coverage and s - t connectivity. We, therefore, derive reliable density estimates for all four models of coverage and connectivity.

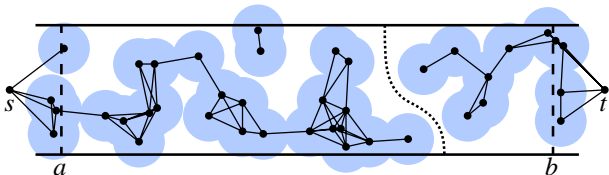


Figure 1: Model with sensor range $\frac{r}{2}$ (sensing regions indicated by shading) or transmission range r (indicated by graph edges). Dotted line indicates a separating path — either a possible path of an undetected intruder, or a path disconnecting s from t in the network.

3. SUMMARY OF OUR CONTRIBUTIONS

3.1 Key Results

In this section, we briefly summarize our key results.

First, we derive a closed form approximation for the break intensity expected in a strip deployment region (see Theorem 9, as well as (12) and (11)). Break intensity, which we define formally in Section 4, informally denotes how frequently we expect to encounter a loss in s - t connectivity as we traverse the length of the strip. We show that the breaks occur with an approximately Poisson distribution. Thus denoting the break intensity by $I_{h,r,\lambda}$,

$$\begin{aligned} \mathbb{P}(G_{h,r,\lambda}(0, \ell) \text{ is not } s\text{-}t \text{ connected}) \\ \approx 1 - \exp(-\ell I_{h,r,\lambda}) \approx \ell I_{h,r,\lambda}, \end{aligned}$$

provided $\ell I_{h,r,\lambda} \ll 1$. As we show in Section 2, the problems of barrier coverage and s - t connectivity are equivalent to determining whether $G_{h,r,\lambda}(0, \ell)$ is s - t connected. Hence, this relation holds for both barrier coverage and s - t connectivity.

Second, we derive a closed form approximation for the probability that the network is not fully connected (see (13) and (14)) in a thin strip. Finally, we derive a closed form approximation for full coverage in thin strips (see (15) and (17)).

We show in simulations that all of our estimates are almost identical to those observed in experiments even for small strips.

Comparing the estimates for the three models (since barrier coverage is equivalent to s - t connectivity), we find that for a long thin strip ($\ell \gg h, r$ and $\lambda r^2 \gg 1$), the probability of failure of barrier coverage is about the same as for full connectivity when $h < \frac{\pi}{2}r$, but much less than for full connectivity when $h > \frac{\pi}{2}r$. Moreover, the probability of failure of full coverage is always much larger than for barrier coverage, or full connectivity (compare Figure 13 to Figures 9 and 11 for similar values of λ).

3.2 Our Contributions and Related Work

Our main contributions in this paper are as follows.

First, ours is the first work to derive reliable density estimates for achieving barrier coverage and s - t connectivity in thin strips. This had been an open problem. We also settle the conjecture made in [14] in the negative, i.e., the conditions for weak barrier coverage derived in [14] are not enough to guarantee barrier coverage. See the lowermost dotted curve in Figure 9, which represents the probability of weak barrier coverage. The top curves in this figure represent the actual probability of barrier coverage and our estimates (which are almost indistinguishable).

Second, ours is also the first work to derive reliable density estimates for full connectivity in thin strips. As we discuss in more detail in Section 8, previous techniques [18] cannot be used in thin strips. This is because in thick regions, the probability that a small component exists is much larger than the probability that two or more large components exist, and hence the latter event is ignored. For thin strips, both events are significant and hence both must be estimated. Again, it is our novel definition of breaks that enables us to estimate the probability of the existence of two or more larger components.

Finally, ours is also the first work to derive reliable estimate for achieving full coverage in a thin strip region. As we show in Section 9, the effect of the boundary dominates in not only thin strip regions but also in thick deployment regions such as disks or squares. Although this behavior has been observed earlier for the square region in [23], no work on full coverage currently exists for thin strips. Further, our estimates are much more accurate since they are not asymptotic, unlike most existing work.

In summary, our work makes it possible, for the first time, to derive reliable density estimates for achieving coverage and connectivity in thin strips.

4. BREAKS

In this section, we define *breaks*, which is central to solving the coverage and connectivity problems in thin strips. We model s - t connectivity in $G_{h,r,\lambda}(a, b)$ by first considering the infinite graph $G_{h,r,\lambda}$. Since there almost surely are separating paths *somewhere* along the infinite strip, our answers are in the form of estimating the *frequency* $I_{h,r,\lambda}$ of these “breaks”. In other words, if $b - a$ is large (as compared with r or h), then the expected number of breaks in $G_{h,r,\lambda}(a, b)$ is about $(b - a)I_{h,r,\lambda}$ (with possibly a small $O(1)$ error due to end effects near $x = a$ and $x = b$). In practical applications, we are interested in the probability of a break occurring, or in the probability distribution of the number of breaks. We define the notion of a break below so as to make breaks “almost independent”. This will imply that the number of breaks in $G_{h,r,\lambda}(a, b)$ is given approximately by a Poisson variable of mean $(b - a)I_{h,r,\lambda}$, and in particular, the probability that no break occurs is about $\exp(-(b - a)I_{h,r,\lambda})$.

The definition of a break is more tricky than it might appear at first. For example, in Figure 2, should the two separating paths be considered as defining the same break, or two different breaks? Since there may be several small connected “islands” in the break, one may be able to construct very many separating paths; indeed, infinitely many topologically distinct paths if we do not mind paths crossing themselves. These two paths are also not really independent

since the existence of one makes the existence of the other much more likely. If we were to count the paths in Figure 2 as distinct breaks, then the probability distribution of the number of breaks in $G_{h,r,\lambda}(a,b)$ would be very far from Poisson, and the probability of a break would be much less than the expected number of breaks even when the expected number of breaks is much less than 1. For this reason, we wish to consider the situation in Figure 2 as a *single* break.

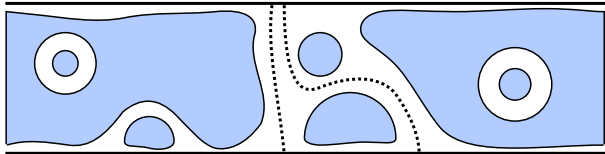


Figure 2: Ambiguity in counting breaks. Do the two paths indicate two separate breaks, or should we consider this as just one “compound” break? Also, there are several small components of $G_{h,r,\lambda}$ that do not cause any breaks.

There are several suitable definitions of a break, which although different, give the same asymptotic frequencies when h and r are large. The following definition is easy to compute in simulations and is fairly convenient theoretically.

Define a *good component* to be a (graph) component C of the graph $G_{h,r,\lambda}$ which contains a sensor strictly within distance $\frac{\sqrt{3}}{2}r$ of the top of S_h , and also contains a sensor strictly within distance $\frac{\sqrt{3}}{2}r$ of the bottom of S_h .

Let t'_C (resp. b'_C) be the positions of the sensors in C closest to the top (resp. bottom) boundary of S_h , and let t_C (resp. b_C) be the corresponding closest point on the boundary. Thus C is good if and only if the line segments $t_C t'_C$ and $b_C b'_C$ are both less than $\frac{\sqrt{3}}{2}r$ in length. We can order the good components from left to right according to the order of the points t_C along the top boundary of S_h . The definition of a good component ensures that this agrees with the ordering given by the points b_C along the bottom boundary of S_h since no good component can “jump” over or under another — this is the significance of the $\frac{\sqrt{3}}{2}r$ margin. To see this, define for each good component C , a path γ_C^0 across S_h from t_C to b_C consisting of line segments. The first line segment joins t_C to t'_C , then there are line segments corresponding to a (graph theoretic) path through C to b'_C , and finally a line segment to b_C (see Figure 3). If the ordering of the points t_C was not consistent with the ordering of the points b_C , then there would be paths $\gamma_{C_1}^0$ and $\gamma_{C_2}^0$ corresponding to two good components, C_1 and C_2 , that cross. Either an edge of one component would cross an edge of the other, or an edge of one component would cross one of the segments $t_C t'_C$ or $b_C b'_C$ of the other. It is a simple geometric exercise to show that in both cases, one of the points of C_1 would be within distance r of one of the points of C_2 , contradicting the assumption that C_1 and C_2 are distinct components of the graph $G_{h,r,\lambda}$.

Define a *break* to be the gap between two consecutive good components. In other words, a break is a partition of the set of good components into two classes, those on the left of the break, and those on the right, which is compatible with the left-right ordering of the good components. The following simple topological result indicates that our definition of a break is reasonable.

LEMMA 1. *Any separating path γ partitions the good components into those that lie to the left of γ and those that lie to the right of γ , so in particular defines a break. Conversely, for any break there exists such a separating path γ . Indeed, if the break occurs between good components C_1 and C_2 then we can choose γ to lie between the paths $\gamma_{C_1}^0$ and $\gamma_{C_2}^0$ defined above.*

PROOF. It is clear that any separating path disconnects $G_{h,r,\lambda}$, and so defines a break. For the converse, suppose C is a good component. The region S consisting of γ_C^0 and all the sensing regions of sensors in C forms a connected subset of the plane. No sensor outside of C is within distance $\frac{r}{2}$ of γ_C^0 or of the sensing range of any sensor in C . Thus no sensor outside of C is within $\frac{r}{2}$ of S . Since S intersects both the top and bottom boundaries of S_h , the external boundary path of S crosses S_h both on the left and on the right of S . Taking a minimal subpath of this boundary that crosses S_h on the right of S gives a suitable separating path corresponding to the break just to the right of C . \square

Note that translational invariance and long range independence of the model $G_{h,r,\lambda}$ imply that horizontal translation is an ergodic transformation on the probability space of this model. Thus breaks (or any other event that can be defined in a translational invariant manner) occur almost surely with a well defined frequency along the strip. Write $I_{h,r,\lambda}$ for the frequency (or intensity) of breaks in $G_{h,r,\lambda}$.

We shall now make more precise our claim above about the independence of breaks. For this we first need to make some assumptions about λ , r , and h . Firstly, we note that by scaling all distances by a factor of $\sqrt{\lambda}$ the graph $G_{h,r,\lambda}$ (resp. $G_{h,r,\lambda}(a,b)$) can be identified with $G_{h\sqrt{\lambda},r\sqrt{\lambda},1}$ (resp. $G_{h\sqrt{\lambda},r\sqrt{\lambda},1}(a\sqrt{\lambda},b\sqrt{\lambda})$). In Sections 4 through 7, we shall therefore make the simplifying assumption that $\lambda = 1$ and drop λ from the notation, for example, writing $G_{h,r}$ and $I_{h,r}$ instead of $G_{h,r,1}$ and $I_{h,r,1}$. We note here for future reference that

$$I_{h,r,\lambda} = \sqrt{\lambda} I_{h\sqrt{\lambda},r\sqrt{\lambda},1}. \quad (1)$$

Next we note that if r is too small then $G_{h,r}$ has very few large components. In this case, good components are rare, but it is easy to find separating paths. Thus for small r , our definition of a break is not very informative. There are few breaks, but the breaks are “wide”, so most of the strip lies in some break. Indeed, if r is small then even if we deployed the sensors in the whole plane, the components would typically be small. However, if r is above the critical threshold for the Gilbert model [10] (numerically about 1.1984, [20, 3]), sensors deployed in the whole plane would give rise to an infinite connected component. In this case, provided h is large, most of S_h would be in (or surrounded by) the sensing regions of good components. For our theoretical results, we generally need r to be at least 6, but numerically r only needs to be slightly above 1.1984 for many of our results to be valid. The last assumption is that the product hr is large, since this will ensure that breaks are rare and “thin” (see below).

For any good component C define x_C^+ to be the smallest x -coordinate of any point on any separating path to the right of γ_C^0 . This is equivalent to the requirement that the separating path is to the right of C , except in the (generally rare) case when C is not within distance $\frac{r}{2}$ of either the top or the bottom of the strip S_h . Similarly define x_C^- to

be the largest x -coordinate of any point on any separating path to the left of γ_C^0 . Note that although γ_C^0 is not uniquely determined, x_C^+ and x_C^- are independent of the choice of γ_C^0 (see Figure 3).

Now define the *width* of a good component C to be $x_C^+ - x_C^-$. If this is positive, then it is also equal to the minimum horizontal distance between separating paths γ^+ and γ^- where γ^+ corresponds to the break immediately to the right of C , and γ^- corresponds to the break immediately to the left of C . Note that the width of a good component may be negative. Define the *width* of a break between good components C_1 and C_2 to be $x_{C_2}^- - x_{C_1}^+$. This is also the maximum horizontal extent of any separating path γ between C_1 and C_2 that lies between $\gamma_{C_1}^0$ and $\gamma_{C_2}^0$. Note that when the strip is finite, the separating widths of the good components and the widths of the breaks add up to the total length of the strip.

LEMMA 2. Assume $r \geq 6$. Then the average width of a break is at most $\max\{5h, 1/h + 2h\}$ and the proportion of good components with width less than $w \geq 0$ is at most $(w + 7)e^{-h}$.

It should be noted that due to the ergodicity of the model $G_{h,r,\lambda}$, terms such as ‘‘average width’’ and ‘‘proportion’’ are well defined.

PROOF. (Sketch) The full details are too long to give here. For further details see [1]. The idea is to tile S_h with small $a \times b$ rectangles, where h/b is an integer, $a \approx b$, and the diameter $(a^2 + b^2)^{1/2}$ of these rectangles is $\frac{r}{2}$. The collection of rectangles intersecting a separating path γ must form a connected region that is devoid of sensors. One can put a lower bound on the size of this region in terms of h and the horizontal extent of γ . One can then bound the number of such possible connected regions of any given size, and bound the probability that such a region is empty. This places a bound on the probability that such a γ exists starting in some fixed rectangle, and hence gives an upper bound on the number of wide breaks. We obtain a lower bound on the total number of breaks by considering empty $r \times h$ rectangles, with all h/b of the small $a \times b$ rectangles immediately to the left containing a sensor (thus guaranteeing a good component exists immediately to the left of the empty rectangle). Using these two bounds gives a bound on the expected width of a break of $5h$ provided $h \geq \frac{r}{2}$. If $h < \frac{r}{2}$ then every sensing region crosses S_h , and one can bound the width of a break by $X + 2h$, where X is an exponential random variable of mean $1/h$ giving the extra horizontal distance between sensors above the minimum needed to form the break. This gives a bound of $1/h + 2h$ on the expected width of a break. For the width of a good component, we fix a ‘‘leftmost’’ path γ_C^0 and use a similar tiling strategy to estimate the probability that a separating path exists to the right of γ_C^0 that gets within horizontal distance w of γ_C^0 . For each rectangle at the top of S_h , we bound the probability that such a separating path originates there by bounding, as above, the size of the collection of rectangles that must be devoid of sensors, and the number of possible such collections. This probability decays exponentially as the horizontal distance d of the starting rectangle from γ_C^0 increases (since the width of the separating path must be at least $d - w$ and this forces an increasing number of empty rectangles as $d \rightarrow \infty$). Thus we obtain a finite bound on the probability that any such

separating path exists, and hence on the probability that the good component has width $< w$. Integrating this bound then gives a bound on the expected width of a good component. In all cases, we carefully optimize a and b for different values of h so as to obtain our precise bounds. \square

As a result of Lemma 2, the breaks are generally small and well separated. Using Theorem 1.21 in [4] and some work one can deduce that the probability distribution of the number of breaks in $G_{h,r}(a, b)$ is approximately Poisson. Thus, we get the following result. (For detailed proof see [1].)

LEMMA 3. Assume $r \geq 6$ and $x > 0$. Then the probability that $G_{h,r}(0, x/I_{h,r})$ contains exactly k breaks tends to $e^{-x} x^k / k!$ as $h \rightarrow \infty$.

Note that although breaks are defined only for the infinite strip, under the conditions of Lemma 3 breaks are unlikely to intersect the lines $x = a$ or $x = b$. Hence any ambiguity in ‘‘the number of breaks in $G_{h,r}(a, b)$ ’’ caused by breaks overlapping the ends of the strip is unimportant.

For practical purposes one would aim to make the existence of any break very unlikely. In this case one can use the following more precise result.

LEMMA 4. If $r \geq 6$ and $h \geq 1$ then

$$\mathbb{P}(G_{h,r}(0, \ell) \text{ is not } s\text{-}t \text{ connected}) \leq (\ell + 5h)I_{h,r}.$$

PROOF. If a separating path exists in $G_{h,r}(0, \ell)$, then $[0, \ell]$ must intersect one of the intervals $[x_{C_i}^+, x_{C_{i+1}}^-]$ corresponding to the breaks in $G_{h,r}$. Equivalently, either $x_{C_i}^+ \in [0, \ell]$, or $0 \in [x_{C_i}^+, x_{C_{i+1}}^-]$ for some i . The expected number of i 's with the first property is $\ell I_{h,r}$ since the asymptotic density of points $x_{C_i}^+$ equals the density of breaks $I_{h,r}$. The expected number of i 's with the second property is at most $5h I_{h,r}$ since the expected width of a break is at most $5h$ (for $h \geq 1$) and their density is $I_{h,r}$. Thus the probability of a separating path in $G_{h,r}(0, \ell)$ is at most $(\ell + 5h)I_{h,r}$. \square

By Lemma 3, the bound given by Lemma 4 is very close to the truth when $h \ll \ell$ and $\ell I_{h,r} \ll 1$ (see Figure 9 and Figure 10). Indeed, the main error appears to be due to the finite size of a break as compared to the length of the strip causing breaks to overlap the ends of the strip, as suggested by the proof of Lemma 4.

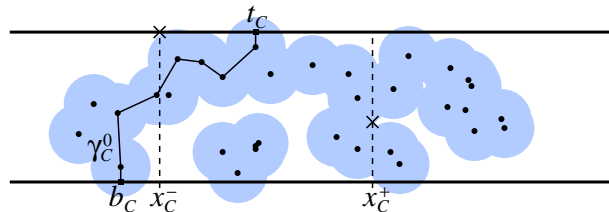


Figure 3: Definition of γ_C^0 (solid path) and x_C^\pm . Crosses indicate the rightmost point to the left of γ_C^0 , or leftmost point to the right of γ_C^0 that a separating path can pass through.

5. HEURISTICS

In this section, we introduce some non-rigorous heuristics about the number and type of components in $G_{h,r}$. The purpose is to provide intuitive methods for estimating the probability of coverage and/or connectivity that do not require extensive derivations, but at the same time provide a fairly accurate first guess. We generally assume r is large, so that in most areas the graph $G_{h,r}$ is highly connected. The main idea is the concept of “excluded area”, that is that rare configurations can be described by the absence of sensors within some nice region A . Outside A , the density of points will be assumed to be close to the normal expected density, which will result in a high degree of connectivity.

Given a region A , the probability that it contains no sensors is $e^{-|A|}$, where $|A|$ is the area of A (assuming $\lambda = 1$). We consider the minimal regions A that can force a component, or a break, under the assumptions that there is a reasonable density of sensors outside of A . For example, a break is very likely to be caused by an excluded $r \times h$ rectangle across the strip (see Figure 4). On either side of this rectangle, the components are very likely to be good (for large r), but the region disconnects $G_{h,r}$. This region is the smallest such region to give a break, so we expect most breaks to be approximately rectangular, and the frequency of breaks $I_{h,r}$ to be about e^{-hr} . (More precisely he^{-hr} , since such a rectangle can be placed after any sensor, and there are about h sensors per unit distance along the strip.) Small components can form near the boundary of S_h with an excluded area of $\frac{1}{2}\pi r^2$, or in the interior of S_h with excluded area πr^2 . Hence these are likely to occur with frequencies about $e^{-\pi r^2/2}$ and $e^{-\pi r^2}$ respectively. Note that if $h < \frac{\pi}{2}r$, then breaks should be more common than these small components, and so most components are likely to be good. On the other hand, if $h > \frac{\pi}{2}r$, then most components are likely to be small (bad) components that do not form breaks.

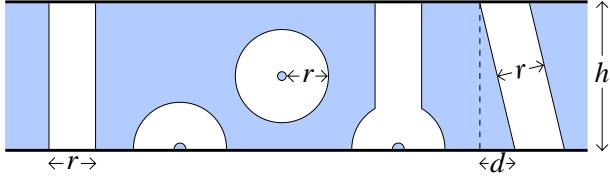


Figure 4: Minimal excluded areas for breaks, small components near the boundary and in the interior, compound breaks, and diagonal breaks.

Compound breaks (breaks containing homotopically inequivalent separating paths) need an excluded area of at least $hr + (\frac{\pi}{3} - \frac{\sqrt{3}}{4})r^2 > hr$ (see fourth example in Figure 4), so at first sight these seem far rarer than simple breaks. Generally this is true, however we need to take care of combinatorial issues — how many ways such breaks can occur. It is possible for compound breaks to be more common than simple breaks if h is extremely large, in particular if $1 \ll r^2 \ll \log h$. This is because although we lose a factor of $e^{-O(r^2)}$ in the frequency of compound breaks due to the extra $O(r^2)$ excluded area, we gain a combinatorial factor of h due to the choice of the vertical position of the small component inside the break. (The excluded area is still $hr + O(r^2)$ even if the small component is in the center of the break.)

Although we expect most breaks to be vertical excluded rectangles, diagonal breaks are also possible. If the top of the break is displaced by a distance d , the excluded area becomes $r\sqrt{h^2 + d^2} \approx hr + \frac{rd^2}{2h}$. If $rd^2 \sim h$ this does not impose a large penalty, so we may expect breaks to often deviate by about $\sqrt{h/r}$ from vertical.

These heuristics tend to be good for large r , however, corrections are needed for the effects of finite r . These corrections tend to make breaks more likely as we shall see in the next section.

6. ESTIMATING BREAK INTENSITY

Having shown in Section 4 that breaks occur with distribution close to a Poisson process, we now derive the break intensity. Once we have the break intensity, it is a fairly simple matter to either estimate or bound the probability that no breaks exist in a strip of given length (which is equivalent to the strip being barrier covered) using either Lemma 3 or Lemma 4.

We are mainly interested in the case when h is larger than r . To estimate $I_{h,r}$ for large h , we shall reduce it to the case of small h . Thus we shall need to study the small h case in some detail. We count the number of points that are the rightmost point of some good component, since this is equivalent to counting good components, and hence breaks. The ideas used naturally reflect the heuristics of the last section, since if a sensor v is to be the rightmost point of a good component, then there must be an excluded area to the right of v .

For $h < \frac{\sqrt{3}}{2}r$, all components are good, so the probability that a fixed sensor v is the rightmost sensor of a good component is given by the probability that there is no sensor to the right of v that is joined in $G_{h,r}$ to v or to any sensor to the left of v . To calculate this probability, fix v , and place sensors to the left of v according to a Poisson process. Then conditioned on the process to the left of v , the probability that v is the rightmost sensor of a component is $e^{-|A|}$ where A is the region to the right of v that is within distance r of v or any sensor to the left of v . The probability that we are interested in is just $\mathbb{E}(e^{-|A|})$ where the expectation is over the position of v and the state of the Poisson process to the left of v . The intensity of such sensors, and hence of breaks, is then given by

$$I_{h,r} = h \mathbb{E}(e^{-|A|}), \quad (2)$$

since the intensity of sensors per unit length of the strip is just h . It remains to calculate $\mathbb{E}(e^{-|A|})$. The excluded area is the union of a number of disks. We first approximate these areas by parabolic regions, replacing the disk $(x - x_0)^2 + (y - y_0)^2 < r^2$ about a sensor (x_0, y_0) by the parabolic region $x - x_0 < r - (y - y_0)^2/2r$. We then estimate the excluded area as $hr - |B|$, where B is the shaded region in Figure 5 which lies to the right of all these approximating parabolas and to the left of the vertical line which is at distance r to the right of v . The advantage of this approximation is that the two parameters r and h can be reduced to a single parameter $z = hr^{-1/3}$ by rescaling the strip by a factor $r^{1/3}$ along the x -axis and by a factor $r^{-1/3}$ along the y -axis. This rescaling leaves the density of the Poisson process and all areas unchanged, but puts the parabolas in the form $x = z_0 - (y - y_0)^2/2$. If for convenience we swap the x and y -coordinates, B can be defined by placing a Poisson process with intensity 1 in the half infinite strip $[0, z] \times [0, \infty)$ plus one more point chosen

uniformly at random on $[0, z] \times \{0\}$ and then taking the area below all the parabolas $y = y_0 + (x - x_0)^2/2$, where (x_0, y_0) ranges over all of these points. Define $\varepsilon(z)$ by

$$\varepsilon(z) = z \mathbb{E}(e^{|B|}). \quad (3)$$

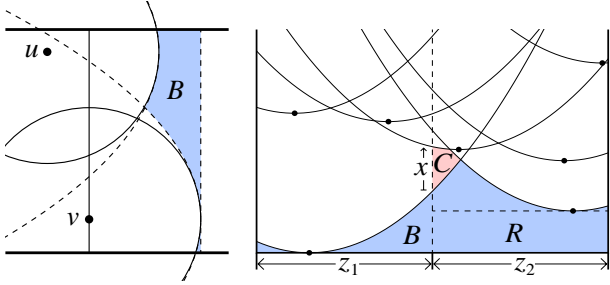


Figure 5: Approximating excluded disks by parabolas. The point v is the rightmost point of the good component, and u is another point to the left of v . The excluded area A consists of the interior of all the circles that lies to the right of v , so $|A| \approx hr - |B|$. For clarity, the picture is drawn with $h > r$. Diagram on the right is used in the proof of Lemma 5.

LEMMA 5. *There exists absolute constants α and β such that $\log \varepsilon(z) = \alpha z + \beta + o(1)$ as $z \rightarrow \infty$.*

(Note, throughout this article, \log denotes the natural logarithm.) A detailed proof of Lemma 5 is too long to include here, but is given in [1]. The basic idea is to show that $\varepsilon(z_1 + z_2)/\varepsilon(z_1)\varepsilon(z_2)$ tends rapidly to some constant ($= e^{-\beta}$) as $\min\{z_1, z_2\} \rightarrow \infty$. To do this, one decomposes the area B for $z = z_1 + z_2$ into two pieces, the part above $[0, z_1]$ and the part above $[z_1, z_1 + z_2]$ (see Figure 5). If y_1 (resp. y_2) is the height of the lowest vertex on the left (resp. right), then the set of vertices can be written as a union of $\mathcal{P}_1 + (0, y_1)$ and $\mathcal{P}_2 + (z_1, y_2)$, where \mathcal{P}_1 and \mathcal{P}_2 are corresponding configurations used in the definition of $\varepsilon(z_1)$ and $\varepsilon(z_2)$. The area $|B|$ can be calculated as $|B_1| + |B_2| + |R| - |C|$, where B_1 and B_2 are the corresponding areas for $z = z_1$ and $z = z_2$, R is either $[0, z_1] \times [0, y_1]$ or $[z_1, z_1 + z_2] \times [0, y_2]$ depending on whether y_2 or y_1 is zero, and C is the set of points that are below all the parabolas corresponding to vertices on their own side, but above a parabola corresponding to a vertex on the opposite side. Fixing the configurations \mathcal{P}_1 and \mathcal{P}_2 , one obtains (after some calculation)

$$\varepsilon(z_1 + z_2) = z_1 z_2 \mathbb{E}_{\mathcal{P}_1} \mathbb{E}_{\mathcal{P}_2} e^{|B_1|} e^{|B_2|} \int_{-\infty}^{\infty} e^{-|C|} dx,$$

where $\mathbb{E}_{\mathcal{P}_1}$ and $\mathbb{E}_{\mathcal{P}_2}$ are expectations over the corresponding configurations and x is the distance of the top left point of B_2 over the top right point of B_1 , so that $|C| = 0$ when $x = 0$. The integral of $e^{-|C|}$ is $O(d_1 + d_2 + 1)$ where d_i is the distance to $x = z_1$ of the furthest vertex v_i of \mathcal{P}_i that defines the parabolic boundary of B_i at some point $x \in [z_1 - 1, z_1 + 1]$. One can show that when one takes expectations over \mathcal{P}_1 and \mathcal{P}_2 , the configurations with large d_i make an insignificant contribution to the expectations and when the v_i are not the lowest vertices, the integral does not depend on z_1 and z_2 . From this one can show that for large z_1 and z_2 , $\varepsilon(z_1 + z_2)$ is

very close to $cz_1 z_2 \mathbb{E}_{\mathcal{P}_1} \mathbb{E}_{\mathcal{P}_2} e^{|B_1|} e^{|B_2|} = c\varepsilon(z_1)\varepsilon(z_2)$ for some constant $c > 0$ independent of z_1 and z_2 .

Computer simulations were performed to estimate $\varepsilon(z)$ for various values of z (see Figure 6). Using these results, the constants in Lemma 5 were estimated as

$$\begin{aligned} \alpha &= 1.12794 \pm 0.00001 \\ \beta &= -1.05116 \pm 0.00005 \end{aligned} \quad (4)$$

(errors are ± 1 standard deviation). The $o(1)$ error term in Lemma 5 appears to tend to zero extremely rapidly as $z \rightarrow \infty$. Indeed, the approximation $\varepsilon(z) \approx e^{\alpha z + \beta}$ is within 2% of the correct value when $z > 0.85$, and for $z > 3$ the error is insignificant (see Figure 6).

For small z , one can expand $\varepsilon(z)$ as a power series in z . One can show that the only non-zero terms are of the form cz^{1+3k} . The first few terms are

$$\varepsilon(z) = z + \frac{1}{12}z^4 + 0z^7 + \frac{1}{64800}z^{10} - \frac{1}{2721600}z^{13} + \dots$$

We obtain these coefficients via symbolic integration using Mathematica.

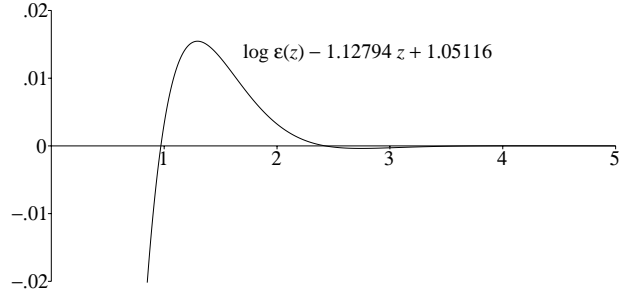


Figure 6: Plot of $\log \varepsilon(z) - 1.12794z + 1.05116$ against z .

THEOREM 6. *If $h \leq \frac{\sqrt{3}}{2}r$ then*

$$I_{h,r} = r^{1/3} \varepsilon(hr^{-1/3}) \exp(-hr + O(hr^{-5/3})). \quad (5)$$

PROOF. Substitute $z = hr^{-1/3}$ into (3) and use (2) with $|A| \approx hr - |B|$. To estimate the error in this approximation, we note that the parabolas $y = y_0 + (x - x_0)^2/2$ used in the definition of $\varepsilon(z)$ need to be replaced by ellipses, which can be approximated as $y = y_0 + (x - x_0)^2/2 + O(r^{-4/3}(x - x_0)^4)$. Following the proof of Lemma 5 we see that the error in the exponent is $O(zr^{-4/3}) = O(hr^{-5/3})$ provided that the slopes of the boundaries of the excluded regions do not pass through vertical. This occurs provided the excluded disks in the original formulation do not have horizontal boundaries inside the strip. But the disks are of radius r and the height of the strip is less than r , so this does not happen. \square

To extend Theorem 6 to larger h , two additional lemmas are required. The first shows that for $h = o(r^2)$, most breaks are “rectangular”. Let C be a good component, and let γ_b be the rightmost boundary of the excluded area to the right of C , i.e., the boundary of the set of points in S_h that are within distance r of some vertex of C .

LEMMA 7. *If $r \geq 6$ then the proportion of good components for which γ_b deviates more than an angle θ from vertical at any point is at most $e^{-r^2(\theta - \sin \theta) + ch}$ for some absolute constant $c > 0$.*

The proof (details given in [1]) involves a tiling argument similar to that used in Lemma 2, bounding the number of choices for the rectangles meeting the rightmost possible path γ_C^0 of C , and the size of the excluded area resulting from each choice.

LEMMA 8. *If $r \geq 6$ and $h = h_1 + h_2$, $h_1, h_2 \geq r$, then*

$$C_1 r^{-1} I_{h_1, r} I_{h_2, r} \leq I_{h, r} \leq C_2 h I_{h_1, r} I_{h_2, r}$$

where C_1 and C_2 are positive absolute constants.

PROOF. (Sketch) Cut the strip S_h horizontally into two strips S_{h_1} and S_{h_2} (with S_{h_1} on the top). If a separating path exists for S_h then it also separates S_{h_1} and S_{h_2} . For at least half of all breaks in S_h , the breaks caused in S_{h_1} and S_{h_2} must be within $10h$ of each other (otherwise the average width of a break in S_h would be more than $5h$ contradicting Lemma 2). The frequency of breaks in S_{h_1} which can be extended to a break in S_h of width at most $10h$ is bounded by $15h I_{h_1, r} I_{h_2, r}$ (adding another $5h$ to take account of the widths of the breaks in S_{h_1} and S_{h_2}). Thus the total frequency of breaks in S_h is at most $30h I_{h_1, r} I_{h_2, r}$. Similarly, for the lower bound, a coincidence of a break in S_{h_1} and in S_{h_2} can give rise to a break in S_h . Typically however, the relative positions of these breaks must be correct to within $O(1/r)$, but this still gives a lower bound on $I_{h, r}$ of $C_1 r^{-1} I_{h_1, r} I_{h_2, r}$. For more details see [1]. \square

THEOREM 9. *For $r \geq 6$ and all $h > 0$,*

$$I_{h, r} = r^{1/3} \varepsilon (hr^{-1/3}) \exp(-hr + O(hr^{-5/3})).$$

PROOF. (Sketch) Using Lemma 7, the argument used in Theorem 6 can be extended to $h = o(r^2)$. For larger h , one uses Lemma 8 to reduce inductively to the case $h \sim r^{5/3+\varepsilon}$, the error introduced by the $C_1 r^{-1}$ and $C_2 h$ terms being swamped by the $O(hr^{-5/3})$ error term in the exponent. \square

Note that even without Theorem 6, Lemma 8 implies that for large, but fixed r ,

$$I_{h, r} = e^{-\alpha_r h + O(\log h)} \quad (6)$$

as $h \rightarrow \infty$, where α_r is some constant dependent on r . Using Theorem 9 we can identify α_r as

$$\alpha_r = r - \alpha r^{-1/3} + O(r^{-5/3}) \quad (7)$$

where $\alpha \approx 1.12794$ is as in Lemma 5.

7. SIMULATION RESULTS FOR BARRIER COVERAGE

In this section, we provide results of our simulations and compare them with our estimates. Note that we are mostly interested in cases when the probability of achieving barrier coverage is closer to 1 (lower portion of the curves in Figures 7 and 9, both of which have logarithmic scale for the probability). This event occurs for sufficiently large values of r and h so that the conditions of our theorems are satisfied. The main result is that our estimates are almost indistinguishable from that observed in simulations even for short strips. All our simulations are sufficiently extensive to provide extremely accurate results, for example, the simulations estimating the probability of a break in a short strip were run over ten million times for statistical validity. As

discussed in Section 4, we assume that all distances have been scaled by a factor of $\sqrt{\lambda}$ so that we may assume $\lambda = 1$.

Our theoretical results strongly suggest that (6) can be strengthened to

$$I_{h, r} = e^{-\alpha_r h - \beta_r + o(1)}, \quad (8)$$

or more generally

$$I_{h, r, \lambda} = \sqrt{\lambda} \exp(-\alpha_r \sqrt{\lambda} h - \beta_r \sqrt{\lambda} + o(1)) \quad (9)$$

where for fixed r the $o(1)$ term tends to zero as $h \rightarrow \infty$ and

$$\begin{aligned} \alpha_r &= r - \alpha r^{-1/3} + O(r^{-5/3}), \\ \beta_r &= -\frac{1}{3} \log r - \beta + O(r^{-4/3}) \end{aligned} \quad (10)$$

as $r \rightarrow \infty$. This does not quite follow from the above, since we may still have a small $O(\log h)$ error term in the exponent. We believe however that this is just an artifact of the proof. Indeed, the $\log h$ term in (6) comes from the $O(h)$ bound on the expected width of a break (Lemma 2) used in Lemma 8. In practice, one should use the amount by which the breaks in S_{h_1} and S_{h_2} can be moved relative to one another and still form a break in S_h . This should be $O(1)$, and in particular should not depend noticeably on h .

Figure 7 plots the logarithm of the frequency of breaks observed in simulation against h for various values of r together with the best fit estimate for the approximating linear function $-\alpha_r h - \beta_r$. The estimated values of α_r and β_r are listed in Table 1. We observe that these linear approximations provide an extremely good fit to the simulation data except when r is close to the critical radius for the Gilbert model $r_c \approx 1.1984$. As mentioned in Section 4, the network becomes highly disconnected in this case, and so the slope α_r tends to zero, as r decreases to r_c .

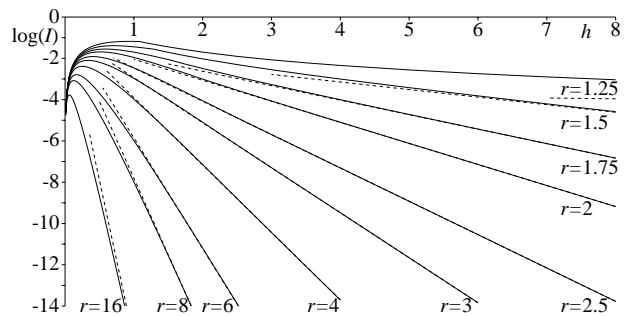


Figure 7: Plots of $\log(I_{h, r})$ against h for various values of r . Dotted lines indicate the linear approximation $-(\alpha_r h + \beta_r)$.

The simulated values of α_r do indeed appear consistent with the theoretical estimate in (7). Using simulations, one can estimate the error terms for α_r and β_r giving

$$\begin{aligned} \alpha_r &\approx \alpha_r^{\text{est}} := r - 1.12794 r^{-1/3} - 0.20 r^{-5/3} \\ \beta_r &\approx \beta_r^{\text{est}} := -\frac{1}{3} \log r + 1.05116 + 0.27 r^{-4/3} \end{aligned} \quad (11)$$

Note that the constants 1.12794 and 1.05116 are the constants from Lemma 5, and only the last coefficients (0.20 and 0.27) were estimated from the simulation estimates of α_r and β_r . From Table 1 one sees that the approximations in (11) are extremely good for $r \geq 3$, but get progressively less accurate for smaller values of r .

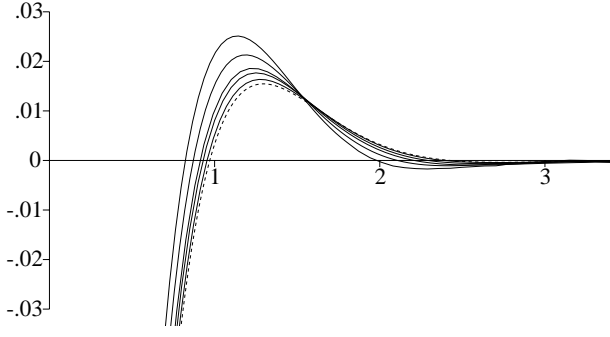


Figure 8: Plots of $\log(I_{h,r}) + \alpha_r h + \beta_r$ against $z = hr^{-1/3}$ for $r = 3, 4, 6, 8, 16$ (outer to inner solid lines respectively). Dotted line indicates the estimate $\log \varepsilon(z) - \alpha z - \beta$.

By comparison with (5), the $o(1)$ term in (8) should be approximately equal to $\log \varepsilon(z) - \alpha z - \beta$ where $z = hr^{-1/3}$. Figure 8 shows the values of this error term obtained from simulations for $r \geq 3$ (for $r < 3$ the error is much larger). Once again, the theoretical result is very close to the results from simulations.

| r | α_r | β_r | α_r^{est} | β_r^{est} |
|------|------------|------------|-------------------------|------------------------|
| 1.25 | 0.0428[1] | 3.6200[15] | 0.0650 | 1.1773 |
| 1.5 | 0.3668[1] | 1.6790[15] | 0.4129 | 1.0732 |
| 1.75 | 0.7018[1] | 1.2198[10] | 0.7353 | 0.9927 |
| 2 | 1.0195[2] | 1.0240[10] | 1.0418 | 0.9273 |
| 2.5 | 1.6152[2] | 0.8490[10] | 1.6255 | 0.8253 |
| 3 | 2.1809[2] | 0.7548[10] | 2.1859 | 0.7474 |
| 4 | 3.2678[2] | 0.6342[10] | 3.2696 | 0.6316 |
| 6 | 5.3686[2] | 0.4801[10] | 5.3692 | 0.4787 |
| 8 | 7.4296[1] | 0.3751[2] | 7.4298 | 0.3749 |
| 16 | 15.5504[1] | 0.1337[2] | 15.5504 | 0.1337 |

Table 1: Estimates of α_r and β_r from simulations, together with the approximations given by (11). Numbers in square brackets indicate approximate 1 standard deviation errors in the last decimal place.

Finally, Figure 9 shows the probability that $G_{h,r,\lambda}(0, \ell)$ is not s - t connected in the special case $h = 2, r = 1, \ell = 10$, with varying λ , together with the estimate $1 - \exp(-\ell I_{h,r,\lambda})$ based on the assumption that breaks occur with an exact Poisson distribution with mean $\ell I_{h,r,\lambda}$. Figure 10 gives a more detailed comparison in this case.

In summary, despite the stated condition that $h \rightarrow \infty$ in (8), if $r \geq 3$ and $z = hr^{-1/3} \geq 2$, the $o(1)$ term in this equation is infinitesimal (as seen from Figure 8), and the break intensity can be well approximated by the following simple expression

$$I_{h,r} \approx e^{-\alpha_r h - \beta_r}, \quad (12)$$

where α_r and β_r are given by (11). Further, notice that although we required $r \rightarrow \infty$ in (10), one can see from Table 1 that the estimates of α_r and β_r given by (11) are quite accurate for $r \geq 3$. To achieve barrier coverage with a reasonably high probability, the values of h and r will generally satisfy these conditions, and therefore as demonstrated in Figure 9 our estimates are quite reliable even for short strips.

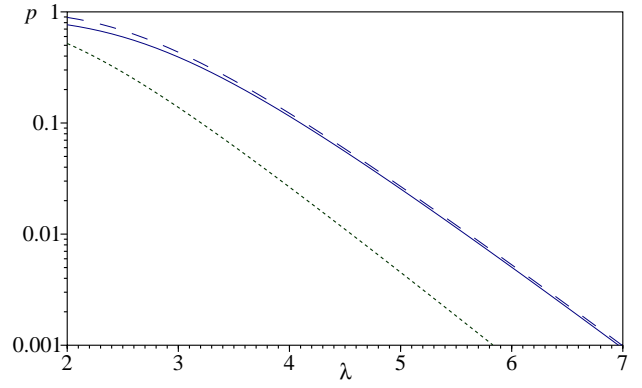


Figure 9: Comparing the probability of the existence of a separating path in $G_{h,r,\lambda}(0, \ell)$ (dashed line) with the estimate $1 - \exp(-\ell I_{h,r,\lambda})$ (solid line) for $\ell = 10, h = 2, r = 1$ and varying λ . The lower dotted curve is the estimate from the asymptotic conditions for weak barrier coverage from [13].

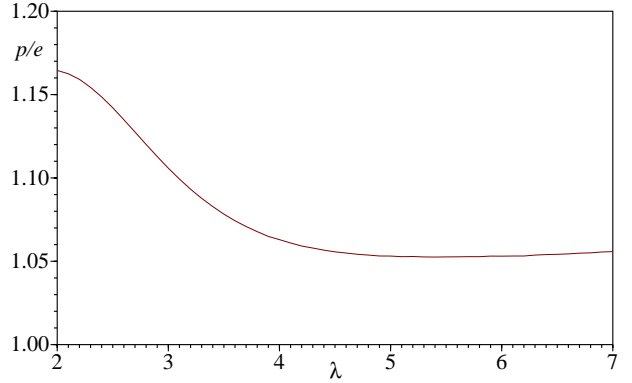


Figure 10: Ratio of the probability p of the existence of a separating path in $G_{h,r,\lambda}(0, \ell)$ to the estimate $e = 1 - \exp(-\ell I_{h,r,\lambda})$ for $\ell = 10, h = 2, r = 1$ and varying λ .

8. FULL CONNECTIVITY

We now consider *full connectivity* of the network. Here we deploy the sensors according to a Poisson process in the rectangle $R = [0, \ell] \times [0, h]$ and join two sensors if they are within distance r of each other. We require that the resulting graph G is connected. In [18] this problem was considered for deployment in a square. To bound the probability that the graph is disconnected, one usually bounds (a) the probability that a small component exists, and (b) the probability that (at least) two large components exist. Here “small” typically means diameter $O(\sqrt{\log A})$ where $A = \ell h$ is the area of the deployment region. In the case of a square (or any “thick” convex region), the probability (b) is much smaller than (a). Moreover, provided we are in a regime where small components are unlikely, (a) is dominated by the occurrence of isolated vertices, i.e., components with just one sensor. For the case of thin strip deployment, (a) is once again dominated by the occurrence of isolated vertices, however the contribution from (b) may also be important. Indeed, large components can occur as the result of a break. Heuristically, we expect small components (or isolated sensors) near

the boundary with frequency about $e^{-\lambda\pi r^2/2}$ and breaks to occur with frequency about $e^{-\lambda hr}$. Thus if $h < \frac{\pi}{2}r$ then breaks will dominate, while if $h > \frac{\pi}{2}r$ then isolated sensors will dominate.

The expected number of isolated sensors in a deployment region R can be calculated exactly since a sensor u is isolated if and only if the region $D_r(u) \setminus \{u\}$ is empty, where $D_r(u)$ is the disk of radius r about u . Even conditioned on the existence of u , this occurs with probability $\exp(-\lambda|R \cap D_r(u)|)$. Thus the expected number of isolated sensors is

$$E_1 = \int_R \lambda \exp(-\lambda|R \cap D_r(u)|) du.$$

Similarly, the expected number of components of size k is

$$E_k = \frac{1}{k!} \int \dots \int \lambda^k \exp(-\lambda|R \cap (\cup_i D_r(u_i))|) du_1 \dots du_k,$$

where the integral is over all configurations of k sensors u_1, \dots, u_k that form a connected set and the $\frac{1}{k!}$ factor compensates for the fact that permutations of u_1, \dots, u_k give rise to the same component. Calculating E_1 or even E_2 may be feasible, however these integrals very quickly become unmanageable. Define E'_k as for E_k , but restrict the integration over u_1, \dots, u_k to those configurations that form a bad component. In [18] it is shown that $\sum_{k=2}^{\infty} E'_k = O(E_1/\lambda r^2)$. Indeed, the proof in [18] first estimates the number of small components, assuming that these components do not get close to opposite sides of the region. (Since [18] assumes a large square region, this is automatic.) Then [18] estimates the number of large components. However, the proof relies on the fact that a large component must be surrounded by a large empty region, which is not valid for thin strip regions when these components are close to both sides of the strip. Hence in both cases, the proofs in [18] are valid *provided* we restrict to counting only bad components. But good components are counted by the intensity of breaks $I_{h,r,\lambda}$. Thus

$$\begin{aligned} \mathbb{P}(G \text{ disconnected}) &\leq \mathbb{P}(G_{h,r,\lambda}(0, \ell) \text{ has break}) + \sum_{k=1}^{\infty} E'_k \\ &\leq \ell I_{h,r,\lambda} + E_1 + O(h I_{h,r,\lambda} + E_1/\lambda r^2). \end{aligned} \quad (13)$$

Moreover, using Theorem 1.21 from [4], one can show that the isolated vertices and breaks are approximately Poisson distributed. Thus the inequality in (13) is close to equality when the right hand side is small compared with 1.

Assuming $h > 2r$, and ignoring effects at the end of the strip (or, by imposing connectivity at the end of the strips as in $G_{h,r,\lambda}(a, b)$), one can calculate E_1 explicitly as

$$E_1 = \lambda \ell (h + 2(\alpha - 1)r) e^{-\lambda\pi r^2}$$

where

$$\alpha = \int_0^{\pi/2} e^{\lambda r^2(\theta - \cos \theta \sin \theta)} \sin \theta d\theta.$$

Approximating this integral and using (13) we obtain

$$\begin{aligned} \mathbb{P}(G \text{ disconnected}) &\approx \ell I_{h,r,\lambda} + \lambda \ell h e^{-\lambda\pi r^2} + \frac{\ell}{r} e^{-\lambda\pi r^2/2} \end{aligned} \quad (14)$$

when the right hand side is small compared with 1.

Figure 11 shows the probability that the graph is disconnected, together with an estimate formed by adding contributions from $I_{h,r,\lambda}$ (breaks), $E'_1 = E_1$, E'_2 , and E'_3 (small

bad components). In Figure 11, $h = 2$ is close, but larger than $\frac{\pi}{2}r \approx 1.5$, so for large λ , isolated sensors dominate this probability. However, the correction terms to the estimate $\ell e^{-\lambda hr}$ for breaks result in the breaks dominating when λ is small.

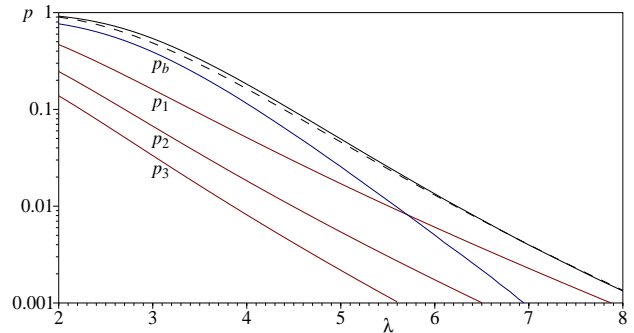


Figure 11: Probability that model in $[0, 10] \times [0, 2]$ is not connected (dashed line) with estimates based on separating paths ($p_b = 1 - \exp(-\ell I_{h,r,\lambda})$), bad components of sizes 1, 2, and 3, ($p_i = 1 - \exp(-E'_i)$) and the sum of these ($1 - \exp(-\ell I_{h,r,\lambda} - \sum_{i=1}^3 E'_i)$, top solid line).

9. FULL COVERAGE

In this section, we consider *full coverage*, i.e., we wish the entire rectangle to be covered by some sensing region. Suppose we place sensors in the rectangle $R = [0, \ell] \times [0, h]$ with density λ . We wish to know if every point in R is covered by at least one sensor. Let the sensing region of the sensor u be $D(u) = D_{r/2}(u)$. Let ∂R and $\partial D(u)$ be the boundaries of these regions (so $\partial D(u)$ is a circle of radius $\frac{r}{2}$). Following [11] (see also [13]), we note that R is covered by the sensor regions if and only if all the following conditions hold.

- ∂R intersects some circle $\partial D(u)$.
- If z is an intersection of the boundary ∂R and some circle $\partial D(u)$, then $z \in D(v)$ for some $v \neq u$.
- If $z \in R$ is an intersection of two circles $\partial D(u)$ and $\partial D(v)$, then $z \in D(w)$ for some sensor $w \neq u, v$.

Since each of the intersections above consist of at most 8 points, this gives a polynomial time algorithm for determining whether or not a specific set of sensors covers the region R . Compared to using a virtual grid as in [15], this algorithm is both more efficient and more accurate.

We now estimate the contributions from (a), (b), and (c). For (a), failure only occurs if there is no sensor within $\frac{r}{2}$ of the boundary of the rectangle, so occurs with probability $\exp(-rP/2 + r^2)$, where $P = 2\ell + 2h$ is the perimeter of the rectangle. In any practical application, this probability is extremely small. For (b) we estimate the number of uncovered intersections with the boundary. Ignoring the effects of corners for the moment, the number of points within $\frac{r}{2}$ of the boundary is about $\lambda r P/2$. There are twice this number of intersection points, each of which is uncovered with probability $\exp(-\lambda\pi r^2/8)$. Thus we obtain on average $\lambda r P \exp(-\lambda\pi r^2/8)$ uncovered intersection points. However,

if there is one uncovered intersection, then there must be an uncovered interval in ∂R , and hence at least two uncovered intersections. Thus we can bound the probability of an uncovered intersection point by half this number. Hence the probability that (b) fails is at most

$$P_1 = \frac{\lambda r}{2}(2\ell + 2h + (\alpha - 4)r)e^{-\lambda\pi r^2/8}, \quad (15)$$

where, for $h, \ell \geq r$,

$$\alpha = 2 \int_0^{\pi/2} e^{\frac{\lambda r^2}{8}(\theta - \cos\theta \sin\theta)} (2 - \sin\theta) \sin\theta d\theta \quad (16)$$

is a correction term for the corners.

For (c) we split the estimate into two cases. The first is when all uncovered intersections are at least $\frac{r}{2}$ from the boundary. In this case, each sensor region intersects at most $\lambda\pi r^2$ others on average, leading to $2\lambda\pi r^2$ intersection points. Multiplying by the average number of sensors and dividing by 2 to correct for double counting, we have on average $\lambda^2\pi r^2 A$ intersections, where $A = \ell h$ is the area of the rectangle. There are at least three uncovered intersections, if there are any, so the probability of an uncovered region is at most $P_2 = \frac{1}{3}\lambda^2\pi r^2 \exp(-\lambda\pi r^2/4)$. Note that $P_2 \leq \frac{4\pi A}{3P_1^2} P_1^2 \leq \frac{\pi}{12} P_1^2$, so if P_1 is small, P_2 is much smaller.

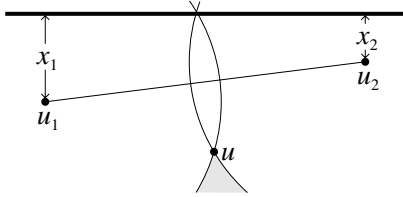


Figure 12: Estimating the probability that an internal intersection is uncovered.

Finally we need to consider the case when the boundary of R is covered, but there are uncovered intersections within $\frac{r}{2}$ of the boundary of R . We count the uncovered intersection points that are the closest such point to the boundary of R . Assume for simplicity that we are far from a corner of R and suppose the uncovered intersection u is the intersection of $\partial D(u_1)$ and $\partial D(u_2)$, and u_i is at distance x_i from the boundary. Since the intersection is the closest point to the boundary that is uncovered, u_1 and u_2 must be on opposite sides of u , and u must be at least $\frac{1}{2}(x_1 + x_2)$ from the boundary itself (see Figure 12). Since the area $|D(u) \cap R|$ is a convex function of the distance $x \in [0, \frac{r}{2}]$ of u to the boundary ∂R , we have $|D(u) \cap R| \geq \frac{\pi r^2}{8}(r + 2x) \geq \frac{\pi r^2}{8}(r + x_1 + x_2)$ when $x \leq \frac{r}{2}$. Also, u_1 and u_2 must be within horizontal distance r of one another for any such intersection to exist. Thus we can bound the probability of such an uncovered point by

$$\begin{aligned} rP \int_0^\infty \int_0^\infty e^{-\frac{\lambda\pi r}{8}(r+x_1+x_2)} \lambda^2 dx_1 dx_2 \\ = \lambda^2 r P e^{-\frac{\lambda\pi r^2}{8}} \left(\frac{8}{\lambda\pi r}\right)^2 \leq \frac{128}{\pi^2 \lambda r^2} P_1. \end{aligned}$$

It can be checked that if u is near a corner, the estimate above is still an overestimate, so once again, in all practical cases P_1 dominates the probability that a point is uncovered. Indeed, the probability that the strip is not fully covered is at most

$$P_1(1 + 128/\pi^2 \lambda r^2 + P_1\pi/12) + \exp(-rP/2 + r^2), \quad (17)$$

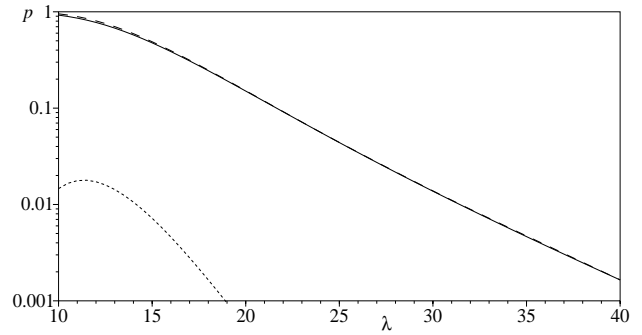


Figure 13: Probability that model in $[0, 10] \times [0, 2]$ with $r = 1$ is not fully covered (dashed line), with estimate $1 - \exp(-P_1)$ based on number of uncovered intersections on the boundary (solid line). (Note that these two curves are almost indistinguishable.) The probability that the boundary is covered but the interior is not is also shown (dotted line).

which is very close to P_1 in all practical cases. One can show (see [1]) that uncovered intervals on the boundary ∂R are not strongly correlated, so if P_1 is small there are unlikely to be more than one such interval. Hence the expected number of such intervals P_1 is close to the probability that one exists. Figure 11 shows the probability that a rectangle is uncovered, together with the estimate based on P_1 and the (much smaller) probability that ∂R is covered, but R is not.

10. CONCLUSIONS AND FUTURE WORK

In this paper, for the first time, we derive reliable density estimates for achieving barrier coverage and s - t connectivity in thin strips. Ours is also the first work to derive reliable density estimates for achieving full coverage and full connectivity. Central to our derivations is a novel definition of *breaks* (gaps in connectivity). We show, using simulations, that our estimates are very accurate predictions of the actual behavior even for small deployment regions. Such accurate estimates make our results readily usable in practice, bridging the gap between theory and practice.

Since ours is the first work for thin strip regions, there are several interesting problems that remain unaddressed. First, instead of 1-coverage (and 1-connectivity), k -coverage (and k -connectivity) may be needed in real deployments for fault-tolerance and load balancing. Second, the deployment regions may be non-rectangular. Our preliminary results suggest that our estimates are sufficient for non-rectangular regions, as well. But, it needs further investigation. Third, disks are not always the best model for sensing and communication. Other models have been proposed in the literature [5, 27]. Finally, other deployment models such as deterministic deployment on grids that are subject to errors in placement may be more appropriate in certain deployment scenarios. We propose to investigate these and other variations in future.

11. REFERENCES

- [1] P. Balister, B. Bollobás, A. Sarkar, and S. Kumar. Reliable density estimates for coverage and connectivity in thin strips of finite length. Technical report, University of Memphis, Available at:

- <https://umdrive.memphis.edu/pbalistr/public/ThinStripComplete.pdf>, 2007.
- [2] P. Balister, B. Bollobás, A. Sarkar, and M. Walters. Connectivity of random k -nearest-neighbour graphs. *Advances in Applied Probability*, 37:1–24, 2005.
 - [3] P. Balister, B. Bollobás, and M. Walters. Continuum percolation with steps in the square or the disk. *Random Structures & Algorithms*, 26:392–403, 2005.
 - [4] B. Bollobás. *Random Graphs*. Cambridge University Press, 2001.
 - [5] Q. Cao, T. Yan, J. A. Stankovic, and T. F. Abdelzaher. Analysis of target detection performance for wireless sensor networks. In *DCOSS*, 2005.
 - [6] M. Cardei, M. Thai, and W. Wu. Energy-efficient target coverage in wireless sensor networks. In *IEEE INFOCOM*, 2005.
 - [7] O. Dousse, F. Baccelli, and P. Thiran. Impact of interferences on connectivity in ad hoc networks. *IEEE/ACM Transactions on Networking*, 13(2):425–436, 2005.
 - [8] O. Dousse, P. Thiran, and M. Hasler. Connectivity in ad-hoc and hybrid networks. In *IEEE Infocom*, 2002.
 - [9] R. B. Ellis, X. Jia, and C. Yan. On random points in the unit disk. *Random Structures and Algorithms*, 29(1):14–35, 2006.
 - [10] E. Gilbert. Random plane networks. *Journal of the Society for Industrial Applied Mathematics*, 9:533–543, 1961.
 - [11] E. Gilbert. The probability of covering a sphere with n circular caps. *Biometrika*, 56:323–330, 1965.
 - [12] P. Gupta and P. R. Kumar. Critical power for asymptotic connectivity in wireless networks. In *IEEE 37th Conference on Decision and Control*, pages 1106–1110, Tampa, FL, 1998.
 - [13] P. Hall. On the Coverage of k -Dimensional Space by k -Dimensional Spheres. *Annals of Probability*, 13(3):991–1002, 1985.
 - [14] S. Kumar, T. H. Lai, and A. Arora. Barrier coverage with wireless sensors. In *ACM MobiCom*, pages 284–298, Cologne, Germany, 2005.
 - [15] S. Kumar, T. H. Lai, and J. Balogh. On k -coverage in a mostly sleeping sensor network. In *ACM MobiCom*, pages 144–158, Philadelphia, PA, 2004.
 - [16] S. Kumar, T. H. Lai, M. E. Posner, and P. Sinha. Optimal Sleep Wakeup Algorithms for Barriers of Wireless Sensors. In *Fourth International Conference on Broadband Communications, Networks, and Systems (IEEE BROADNETS)*, Raleigh, NC, 2007.
 - [17] X. Y. Li, P. J. Wan, Y. Wang, and C. Yi. Fault tolerant deployment and topology control in wireless networks. In *ACM MobiHoc*, pages 117–28, Annapolis, MD, 2003.
 - [18] M. D. Penrose. *Random Geometric Graphs*. Oxford University Press, 2003.
 - [19] E. Peserico and L. Rudolph. Robust network connectivity: when it’s the big picture that matters. *SIGMETRICS Perform. Evaluation Review*, 34(1):299–310, 2006.
 - [20] S. Quintanilla, S. Torquato, and R. Ziff. Efficient measurement of the percolation threshold for fully penetrable discs. *J. Phys. A: Math. Gen.*, 33(42):L399–L407, 2000.
 - [21] P. Santi. The critical transmitting range for connectivity in mobile ad hoc networks. *IEEE Transactions in Mobile Computing*, 4(3):310–317, 2005.
 - [22] S. Slijepcevic and M. Potkonjak. Power efficient organization of wireless sensor networks. In *ICC*, 2001.
 - [23] P. Wan and C. Yi. Coverage by randomly deployed wireless sensor networks. *IEEE Transactions on Information Theory*, 52:2658–2669, 2006.
 - [24] F. Xue and P. R. Kumar. The number of neighbors needed for connectivity of wireless networks. In *Wireless Networks*, volume 10, pages 169–181, March 2004.
 - [25] H. Zhang and J. Hou. Maintaining sensing coverage and connectivity in large sensor networks. In *NSF International Workshop on Theoretical and Algorithmic Aspects of Sensor, Ad Hoc Wirelsss, and Peer-to-Peer Networks*, 2004.
 - [26] H. Zhang and J. Hou. On deriving the upper bound of α -lifetime for large sensor networks. In *ACM MobiHoc*, pages 121–132, Tokyo, Japan, 2004.
 - [27] G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic. Impact of radio irregularity on wireless sensor networks. In *ACM MobiSys*, Boston, MA, 2004.