The Costly Path from Percolation to Full Connectivity *

Olivier Dousse
School of Computer and Communication Sciences
EPFL - Lausanne
Switzerland
olivier.dousse@epfl.ch

Massimo Franceschetti
Dept. of Electrical and Computer Engineering
University of California
San Diego, CA, USA
massimo@ece.ucsd.edu

Patrick Thiran
School of Computer and Communication Sciences
EPFL - Lausanne
Switzerland
patrick.thiran@epfl.ch

Abstract

Requiring all nodes of a wireless multihop network to be connected is expensive and results in a poor scalability of properties such as transport capacity. We show however that it is no longer the case if we only slightly loosen the connectivity requirement, by just imposing that most nodes be connected to each other (so that the network "percolates"). This feature is found in models neglecting interferences, taking interferences as noise or taking a more information theoretic approach.

1 Introduction

Connectivity of a multi-hop wireless network is usually understood as the *full connectivity* of all its nodes. Lower bounds on its transport capacity [1] are obtained by routing data in networks that are already assumed to be fully connected. In this paper, we replace the costly requirement of a fully connected network by that of a network where not all, but most of the nodes, are connected to each other. More precisely, we want a giant component to appear in the network, that contains a vast majority of the nodes, but we can leave a tiny number out of it (in other words, we want the network to percolate). This apparently benign change gives a much more optimistic perspective on the scalability of wireless multihop networks. We consider three different models: the simplest Boolean model, with a circular connectivity range for each node (Section 2), a model with interferences, where nodes connect if their signal to noise ratio is above some given threshold (Section 3), and a more "information theoretic" model, where nodes connect if they can exchange data at a rate higher than some given rate (Section 4). This shows therefore that the high cost of full connectivity can be spared without prejudice. On the contrary, trading full connectivity for a giant component makes it possible to construct a scheduling and routing scheme that matches the upper bound on the transport capacity of the network and offers multiple paths between most nodes (Section 5).

*The work presented in this paper was supported (in part) by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

2 The Boolean model

The first and simplest model for tackling the connectivity issue is the Boolean model. The main assumption is that nodes have a *connectivity range* r, within which they can wirelessly connect to their neighbors. We assume furthermore that the range is the same for all nodes. Therefore, two nodes are directly connected if the distance between them is less than r.

We take now a finite area, and model the node distribution using a 2-dimensional Poisson point process of intensity λ (λ is thus the average number of nodes per square meter). *Full connectivity* occurs when one can find a path joining any two nodes.

As a first observation, it is clear that the probability that the network is fully connected is always less than one, whenever the diameter of the network area is larger than r. Therefore, full connectivity can only be an asymptotic property, in the sense that this probability can only *tend* to one. Moreover, if one considers the (unrealistic) case where the network area is infinite, then the probability that the network is fully connected is always exactly zero.

However, in the case of an ad hoc network, one can say that a network still *well connected*, if disconnected nodes may exist but always represent a small fraction of the total number of nodes. We say that a node is *disconnected* if it is not connected to the *majority* of the other nodes. In fact, in the context of ad hoc networks, we would like most nodes to belong to the same huge connected component (which forms the network itself).

Percolation theory addresses the case where the network area is infinite, and the fundamental result is that if the node density λ and the range r are such that $\pi\lambda r^2 > N^*$, for a special constant $N^* \simeq 4.5$, then the network is indeed formed by a huge connected component (the network), plus a multitude of finite components (disconnected nodes). Moreover, the fraction of connected nodes is a deterministic function θ of the average node degree $\pi\lambda r^2$.

Therefore, this infinite network model is a good approximation for large networks. However, networks are never infinite, and one needs more specific results for the finite case. Penrose and Pisztora [2] showed that for a large but finite area, the fraction of connected nodes is always close to the deterministic function $\theta(\pi\lambda r^2)$. We call this *partial connectivity.*

Theorem 1 *Let $B(m)$ denote the square $[0,m]^2$ and set $\mathcal{P}_{\lambda,m} := \mathcal{P}_\lambda \cap B(m)$, a Poisson point process of intensity λ on $B(m)$. Suppose that $\lambda r^2 > N^*$. Let $0 < \varepsilon < \frac{1}{2}$, and let $E(m)$ be the event that (i) there is a unique cluster $C_b(B(m))$ on $\mathcal{P}_{\lambda,m}$ containing more than $\varepsilon\lambda\theta(\pi\lambda r^2)m^2$ points of $\mathcal{P}_{\lambda,m}$, and (ii)*

$$(1-\varepsilon)\lambda m^2\theta(\pi\lambda r^2) \leq \mathrm{card}(C_b(B(m)) \cap \mathcal{P}_{\lambda,m}) \leq (1+\varepsilon)\lambda m^2\theta(\pi\lambda r^2).$$

Then there exist constants $c_1 > 0$ and $m_0 > 0$, such that

$$\mathbb{P}[E(m)] \geq 1 - \exp(-c_1 m), \quad m \geq m_0.$$

We can observe in the above theorem that if we let the area of the network m^2 tend to infinity (and thus also the number of nodes), then the fraction of connected nodes tends to the constant $\theta(\pi\lambda r^2)$. This result matches the first percolation result for infinite networks. However, there is a slight difference between the two: in the first one, we consider only an *infinite* network area, whereas in the second one, we consider a sequence of *finite* networks, and derive an *asymptotic* property when the number of nodes tend to infinity.

The same approach can be applied for full connectivity. As we only consider (larger and larger but still) finite networks, full connectivity can happen with positive probability. This approach is very frequent in the literature, and the following result has been proven first by Penrose in 1997 [3].

Theorem 2 *Let M_n denote the length of the longest edge of the minimal spanning tree connecting all nodes in the network. Then*

$$\lim_{n \to \infty} \mathbb{P}[n\pi M_n - \log n \le \alpha] = \exp(e^{-\alpha}), \quad \forall \alpha \in \mathbb{R}.$$

It is now interesting to observe how these two results, namely the one about partial connectivity (Theorem 1) and the one about full connectivity (Theorem 2), are related. We will start with an intuitive – but not rigorous – reasoning, that allows to link the probability that the network is fully connected with the function $\theta(\pi\lambda r^2)$. A rigorous and detailed version of this reasoning can be found in [4].

We assume without loss of generality that nodes are distributed according to a Poisson point process of unit density ($\lambda = 1$), and that the network area increases and is equal to n. The total number of nodes in the network is therefore approximately equal to n. We want to compute the critical range $r(n)$ to keep the network fully connected.

We start with the simple observation that the network is fully connected if and only if no node is disconnected. So we obtain immediately a lower bound on the probability that the network is connected:

$$\mathbb{P}[\text{the network is connected}] \ge 1 - nP[\text{node } i \text{ is disconnected}].$$

Furthermore, we know from Theorem 1 that the probability that a given node is isolated tends to $1 - \theta(\pi\lambda r^2)$ when the network area tends to infinity. Therefore

$$\mathbb{P}[\text{the network is connected}] \ge 1 - n(1 - \theta(\pi\lambda r^2)).$$

In fact, this lower bound is asymptotically tight, as when the number of nodes becomes large, the events that different nodes are connected become almost independent. Therefore we have

$$\mathbb{P}[\text{the network is connected}] \simeq \theta^n(\pi\lambda r^2). \tag{1}$$

In order to have the above probability tend to one when n tends to infinity, we must have $\theta(\pi\lambda r^2) \to 1$. This is only possible if $\lambda r^2 = \lambda r^2(n)$ grows with n.

To compute how fast $r(n)$ should grow, we derive an approximation of the function $\theta(\pi\lambda r^2)$ when $\pi\lambda r^2$ is large. According to Propositions 6.4-6.6 in [5], when the average node degree is large, the ratio between the probability that a node is disconnected, and the probability that a node is isolated (i.e. has degree zero) tends to one. Therefore, asymptotically, all disconnected nodes are isolated nodes. The probability that a node is isolated is easy to compute:

$$\mathbb{P}[\text{a given node is isolated}] = \exp(-\pi\lambda r^2).$$

Therefore, when $\pi\lambda r^2$ is large, $\theta(\pi\lambda r^2) \simeq 1 - \exp(-\pi\lambda r^2)$. Figure 1 shows a simulation based evaluation of $\theta(\pi\lambda r^2)$ for large values of $\pi\lambda r^2$.

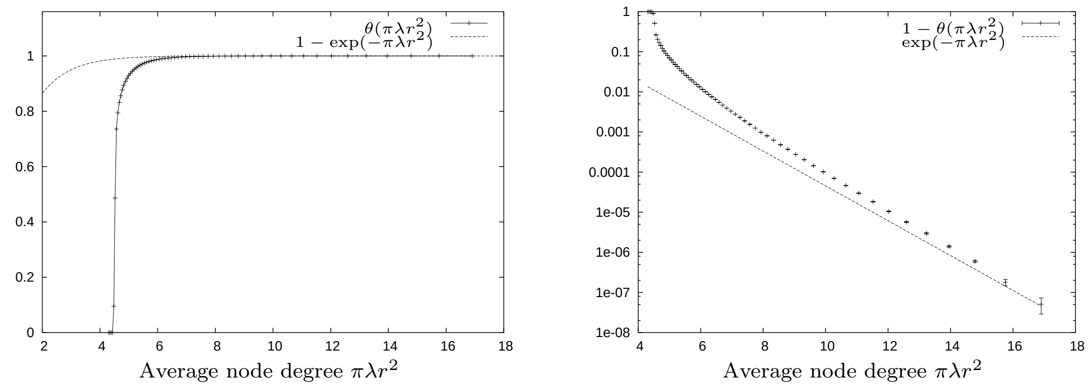

Figure 1: The fraction of disconnected nodes in the Boolean model.

Using this approximation in (1), we obtain

$$\mathbb{P}[\text{the network is connected}] \simeq [1 - \exp(-\pi\lambda r^2(n))]^n.$$

The latter expression tends to one provided $n \exp(-\pi\lambda r^2(n))$ tends to zero. Taking the logarithm of this expression leads to the conclusion that the network is asymptotically connected if and only if

$$\pi\lambda r^2(n) - \log n \to \infty.$$

This means that the average node degree must grow approximately like $\log n$, when the number of nodes in the network increases. However, it appears in this intuitive derivation of the critical range for full connectivity that the *most isolated node* is determining the result. In fact, the network becomes connected when the last node joins the network. This means that full connectivity is not really a global property of the network; it just answers the question "how isolated is the most isolated node". As we just saw above, the distance to the first neighbor of the most isolated node increases with the number of nodes. But this is a pure statistical effect: we are taking a set of randomly distributed distances, and pick the largest one. As we increase the sample set, the largest element becomes longer and longer. This explains why the range of the nodes has to increase, even though the node density remains constant.

3 Taking interferences into account

In this section, we introduce a slightly more sophisticated model for connectivity. We keep the Poison distribution of the nodes, but assume that two nodes are neighbors if they can communicate wirelessly, despite the interferences from all other nodes. More precisely, Node i and Node j are neighbors if

$$\frac{Pl(||x_i - x_j||)}{N_0 + \gamma \sum_{k \neq i,j} Pl(||x_k - x_j||)} > \beta \text{ and } \frac{Pl(||x_j - x_i||)}{N_0 + \gamma \sum_{k \neq i,j} Pl(||x_k - x_i||)} > \beta, \qquad (2)$$

where P is the emitting power of all nodes, $l(d)$ is the attenuation of the signal over distance d, N_0 is the ambient noise, γ is a weighting factor for interferences (inverse processing gain of the CDMA system) and β is the SNIR threshold for successful decoding.

In this model, as we take interferences into account, it is not enough to increase the range (i.e. the emitting power) of the nodes to keep the network connected. Increasing the emitting power would also increase interferences, and at the limit, when P tends to infinity, one can see in (2) that the SNIR converges to a maximum.

However, it has been proven in [6] that percolation can still occur in this model, for appropriate parameters λ, P, N_0, γ and β. Regarding partial connectivity, the situation is thus similar to the previous section. Let us make this precise. We assume in the sequel that λ, N_0 and γ are fixed, and that P and β are our parameters (varying β correspond to adapting the data rate, for example as in 802.11).

Let us consider the case where the network spreads over the whole plane first, and denote by $\theta(P, \beta)$ the probability that a given node belongs to an infinite connected component (i.e. is connected to the network). The following proposition can be inferred from the results in [6]

Proposition 1 *For any value of P, there exists a critical value $\beta^*(P)$ such that if $\beta < \beta^*(P)$, then there exists an infinite connected component a.s., and $\theta(P, \beta) > 0$.*

Although it was not formally proven yet, we can conjecture that a similar theorem as Theorem 1 exists for this model, and that for a finite network, the fraction of connected nodes tend to $\theta(P, \beta)$ when the network area grows.

For full connectivity, we must study more precisely the shape of the function $\theta(P, \beta)$. Simulation based evaluations of $\theta(P, \beta)$ are presented on Figure 2. As a first observation, we notice that if β is fixed and P arbitrarily large, the fraction of connected nodes $\theta(P, \beta)$ is bounded above by a constant $\overline{\theta}(\beta)$, that is strictly smaller than one. In fact, letting P tend to infinity is equivalent to let N_0 tend to zero in (2). Therefore, at the limit, we obtain a purely interference-limited network. It can be shown that such a network always contain a non-zero fraction of disconnected nodes. We skip this proof because of space limitation.

However, to have full connectivity, we must have $\theta(P, \beta) \to 1$ when $n \to \infty$. The only way to achieve that is to have $\beta \to 0$, as shown on the right-hand side of Figure 2. Therefore, to achieve full connectivity, the rate on each link has to decrease with the number of nodes.

A drawback of this model is that we assume that all nodes emit with power P at any instant. This assumption leads of course to strong interferences, and therefore a poor connectivity. One can improve the model by introducing a random TDMA scheme, where each node picks a time slot at random, out of t time slots. Then, at each instant, only a fraction $1/t$ of the nodes are emitting, and the average interference level is t times lower. For fixed P and β, one can now increase the number of time slots t to improve the network connectivity. This setting has also been studied in [6], and it has been shown that the network can be made at least super-critical (i.e. partially connected) by setting t sufficiently large.

But whether we decrease β or increase t, we always end up to reduce of the throughput per link. It turns out in this model that increasing the emitting power does not allow to reconnect the network. The only way to achieve full connectivity is to reduce the throughput.

4 Information theoretic connectivity

In this section, we consider a slightly more sophisticated (and more realistic) definition of connectivity: two nodes are connected, if it is possible to transmit data from one node to the other at rate at least R, all other nodes of the network serving as relays. Nodes are

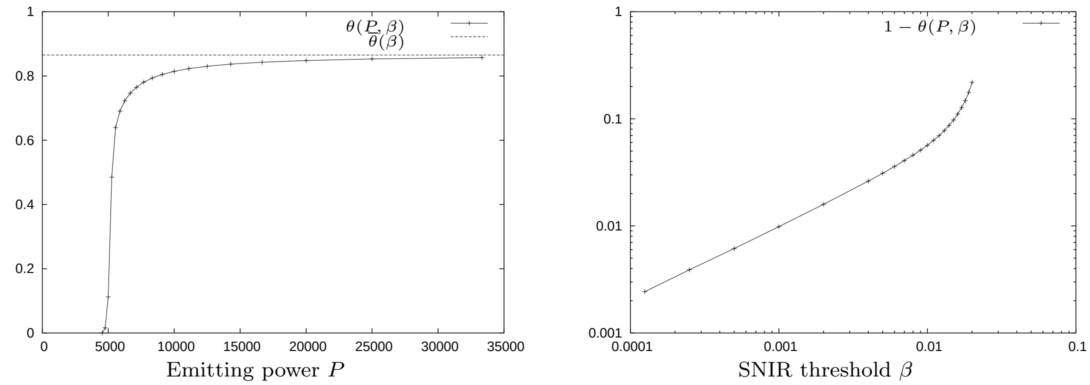

Figure 2: On the left: the fraction of connected nodes, as a function of the emitting power P, for a fixed $\beta = 0.02$. The function converges to $\hat{\theta}(\beta) < 1$. On the right, the fraction of disconnected nodes, as a function of the SNIR threshold β, with fixed $P \simeq 7.7 \cdot 10^3$.

still distributed according to a Poisson point process of intensity λ over an area of size A. This model has been first introduced by Liu and Srikant [7]. It has in fact many more parameters than the Boolean model, like the maximum emitting power of the nodes, the attenuation exponent or the ambient noise. Here, we assume them all fixed, but the node density λ, the network area A and the rate R.

We raise the same two questions as above: is the network is fully connected, and does it contain a giant connected component with most nodes. It turns out that the first question lead to asymptotic results similar to those obtained for the Boolean model:

Theorem 3 *[7] For an attenuation function of the form $l(d) = d^{-\alpha}$, with $\alpha > 1$, and if $R(A) \geq \frac{c_2 \lambda^\alpha}{(\log \lambda A)^{\alpha-1}}$ where $c_2 = \frac{48P4^{\alpha-1}(1+\varepsilon)(2\alpha-1)}{(1-\varepsilon)^\alpha N_0(\alpha-1)}$ for some $\varepsilon > 0$, the network is disconnected w.h.p. when $A \to \infty$.*

We observe that when the network size A increases, the rate must decrease to keep the network connected. Equivalently, if we fix the rate, then the emitting power of the nodes should increase. Again, this result is a pure statistical effect, as the probability to find an arbitrarily isolated node in the network tends to one when the network area grows to infinity. In their paper, Liu and Srikant directly use this argument on the most isolated node to find the upper bound on the achievable rate R.

However, if we only require partial connectivity, with an arbitrarily low fraction of disconnected nodes, the asymptotic behavior of R when the area tends to infinity dramatically changes: R can be kept constant when the network size tend to infinity. The two following theorems have been proven in [8], in the case where the network area tends to infinity.

Theorem 4 *For any $0 < \check{\theta} < 1$, there exists a rate $R > 0$ independent of n, such that a fraction at least $\check{\theta}$ of the nodes can exchange data at rate R w.h.p.*

Theorem 5 *For any rate $R > 0$, the fraction of nodes that can send data to some destination at that rate is at most $\hat{\theta}$ w.h.p., where*

$$\hat{\theta} = \mathbb{P}[I \geq \frac{N_0}{P}\left(e^{2R} - 1\right)],$$

where I is the shot-noise *defined by $I = \sum_{x \in N} l^2(||x||)$ and N is a Poisson point process of unit density over $\mathbb{R}^2$.*

In other words, the two above theorems say that when the network size tends to infinity, a constant fraction of the nodes have the property that for any pair of such nodes, a link with throughput R can be established between them. This remains true when the two nodes are arbitrarily distant. The converse says that for any rate R, there will be a non-zero fraction of nodes that cannot communicate with any other node with rate at least R.

Again, with this model of connectivity, the results have the same flavor: requiring full connectivity makes the network performance drop when the number of nodes is large. On the contrary, if we only require partial connectivity (even with a very small fraction of disconnected nodes), the network performance scales perfectly.

5 The cost of full connectivity

5.1 The capacity under uniform traffic matrix

In the above section, we considered the case where only one link is active at a time. Although this assumption is realistic for lightly loaded sensor networks, it is also interesting to consider the other extreme case, where *all* nodes want to transmit data at the same time.

We consider here the model where each node picks a destination at random, and transmit data to it (uniform traffic matrix). As before, the node density is kept constant, whereas the network area increases. n denotes the average number of nodes.

This situation has been first studied in [1], where a constructive scheme is found, that achieves a throughput of order $1/\sqrt{n \log n}$. However, if the nodes are placed in a more regular (non-random) fashion, a throughput of order $1/\sqrt{n}$ is achievable. The key for this gap is that we require that *each node* benefits from the same throughput. Of course, this rules out the possibility of having a few disconnected nodes (which would have zero throughput). In this section, we follow the alternative approach from [9], and show how the requirement of full connectivity is responsible for this $\sqrt{\log n}$ factor.

We consider here a square area $B(m) = [0, m] \times [0, m]$, and we assume without loss of generality that $\lambda = 1$. Hence $n = m^2$. We divide the area $B(m)$ into squarelets of size $c \times c$. There are thus approximately n/c^2 squarelets, each containing in average c^2 nodes. We set the nodes' power so that each node in a squarelet can reach any node in the 4 contiguous squarelets. Then, for each source-destination pair, we use the following deterministic routing scheme (see Figure 3): we draw a straight line from source to destination, and pick one node per squarelet cut by this line. These nodes will be the relays for this flow. It can be shown that with a simple TDMA scheme, each squarelet benefits from a constant throughput, which must be shared by all nodes inside of the squarelet. Therefore, the actual throughput per flow will depend on the number of routes that cross a typical squarelet.

However, this deterministic routing scheme only works if there is at least one node in each squarelet (this happens with probability $p = 1 - \exp(-c^2)$ in each squarelet). Clearly, if c is a constant, there is always a positive probability that each squarelet is empty. Asymptotically, the probability that there are at least k empty squarelets in the network goes to one, for any finite k. Even worse, the probability that there are occupied squarelets surrounded by empty squarelets also tends to one. Therefore, our network is disconnected w.h.p.

To keep the network connected, we must let c grow as a function of n. By taking $c = c(n) \sim \sqrt{\log n}$ we obtain a connected network, with order $\Theta(\log n)$ nodes per

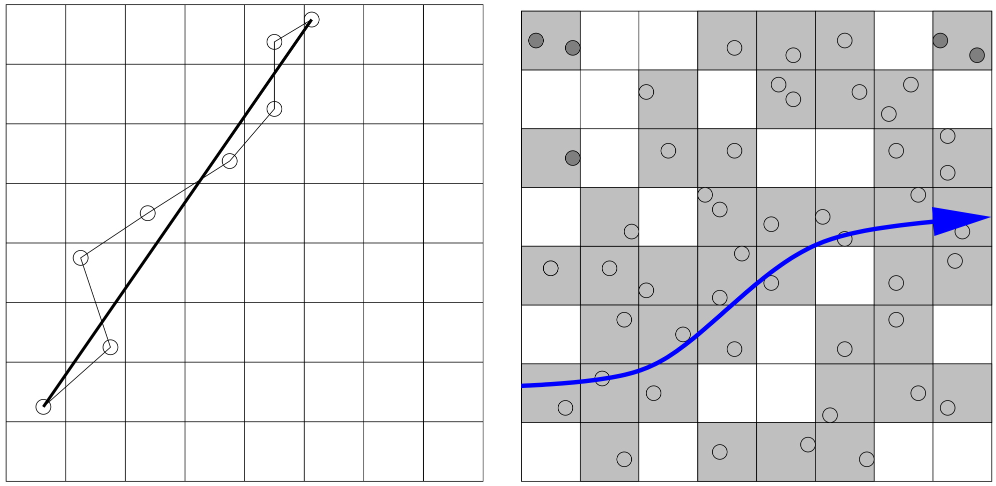

Figure 3: On the left: the deterministic routing scheme. In each squarelet cut by the bold line, we pick on relay. On the right, the site percolation model induced by the occupancy of squarelets. Filled disks represent the isolated nodes in our scheme: the deterministic routing scheme cannot be successfully applied to these nodes

squarelet, and $\Theta(\sqrt{n/\log n})$ routes crossing each squarelet. The resulting throughput is thus $\Theta(1/\sqrt{n \log n})$.

An alternative is to allow for some empty squarelets, and to keep c constant. This lead to a *site percolation* model on a square grid, where nodes are occupied with probability $p = 1 - \exp -c^2$ and empty with probability $1 - p$. It has been shown in [10] that if c is large enough, one can find a routing scheme on this incomplete grid, so that each node has a throughput of order $1/\sqrt{n}$.

However, with a constant squarelet size c, there are still disconnected nodes w.h.p. To cope with them, we use a separate time slot, for draining the traffic from disconnected nodes to the connected part. This is done by using longer hops. As this traffic is only local (from remote nodes to the closest connected nodes), it can be shown that the throughput per node during this special time slot decreases slowlier than $1/\sqrt{n}$. Therefore, as the overall throughput is determined by the minimum of the throughput during the two time slots, introducing this draining phase does not affect the asymptotic behavior of the result. The difference between this strategy and the previous one is that here we treat the problems of transport capacity and connectivity *separately* (in two distinct time slots), whereas by letting c tend to infinity, we solve the capacity *and* connectivity issues simultaneously, with high cost on the throughput.

Note that in this model, we can grantee service to all the nodes (and therefore keep the network fully connected) only because the throughput is decreasing with n. There is therefore no contradiction with Theorem 3.

5.2 The number of paths

As suggested by the above section, increasing c (the average hop length) with n is counter productive. On the other hand, if c is too small, the fraction of empty squarelets is large, and no route can be found from one side of the network to the other (sub-critical case). There should therefore be an optimal value for c, that maximizes the throughput.

In this section, we address this trade-off in a more general context by studying the

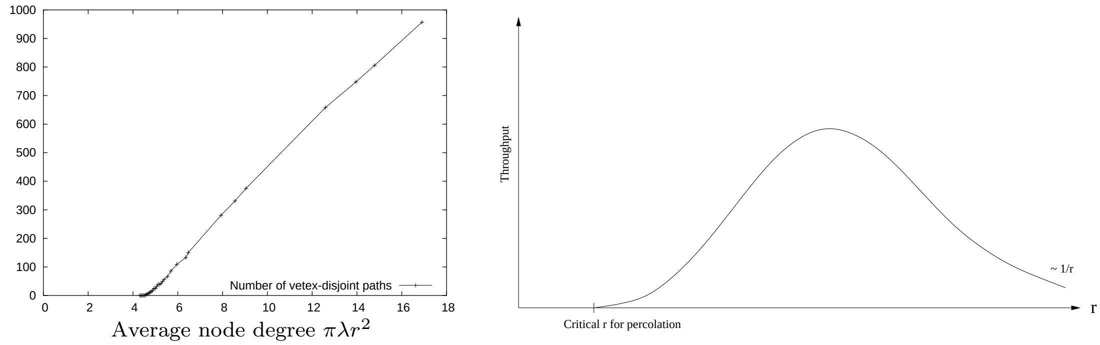

Figure 4: On the left, the number of vertex disjoint path in a Boolean model. On the right, a sketch of the horizontal throughput as a function of the average node degree.

number of paths in the percolation cluster, under the protocol model, as defined in [1] and the model presented in Section 3.

We start with the Boolean model. We consider a large square network, with a fixed node density λ, and each node having the same connecting range r. We try here to find an optimal value of r for maximizing the throughput from the left side of the network to the right.

We know already that if the range r is too small, the network is sub-critical, and there exists no path crossing the network from left to right. Therefore, r has to be such that $\pi\lambda r^2 > N^*$. Furthermore, when r increases, the connectivity graph becomes richer, and the number of left-right crossings increases. This provides a better left-right throughput.

However, using a large range causes many interferences. To evaluate their impact on throughput, we use the protocol model [1]: each transmission occupies a footprint of area $\pi(\Delta r)^2$. Therefore, if the size of the network area is n, the maximum number of simultaneous transmissions cannot be larger than $n/\pi(\Delta r)^2$. On the other hand, each transmission can transport data at a distance at most r. Thus, the total transport capacity is $n/\pi\Delta^2 r$. As the length of a left-right crossing is at least $m = \sqrt{n}$, the horizontal throughput across the network cannot be larger than $\sqrt{n}/\pi\Delta^2 r$.

Figure 4 shows the shape of the throughput as a function of r. For small average node degrees, no path are found, because the network is sub-critical. Above the percolation threshold, the number of vertex-disjoint paths (and thus the throughput) slowly increases. This corresponds to the noise-limited regime. For high node degree, many paths are found, but they interfere with each other, and the overall throughput decreases like $1/r$, for the reason mentioned above. This corresponds to the interference-limited regime. Therefore, there should exist an optimal value for the range r.

We finish this section with a slightly different simulation model. Here we consider the model of Section 3, with the random TDMA scheme with t time slots. In this model, the SNIR is at least β on each link. Thus, the throughput on each path is guaranteed, and it is enough to count the number of vertex-disjoint path to estimate the throughput. However, if we use TDMA, we have to divide the result by the number of time slots t.

Figure 5 shows the throughput obtained for several number of time slots. We took the parameters of the model such that the network is sub-critical when no TDMA is used ($t = 1$). This is the interference-limited case. Then, when we increase t, interferences decrease, and more links are available. The number of paths increases. But when t becomes very large, the interferences are so low that the noise becomes dominant, and the connectivity graph saturates. Thus, the throughput decreases like $1/t$.

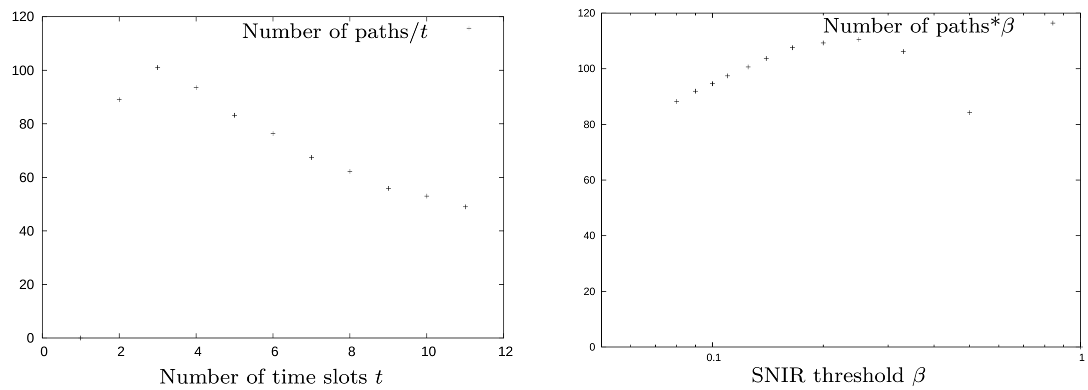

Figure 5: The horizontal throughput in the signal-to-interference model. On the left, the normalized number of left-right crossings with TDMA. In this simulation, the optimal number of time slots is $t = 3$. On the right, the normalized number of left-right crossings with a lower SNIR threshold β.

Similarly to Section 3, an alternative strategy would be to reduce the rate on the links (reduce β). As the required SNIR would be lower, one would obtain more links, but with lower rate. This strategy leads to a similar trade-off a with the TDMA scheme. If the threshold is too high, the network does not percolate, and no path is found. When the threshold tends to zero, the actual range of the nodes is bounded from above: according to (2), we have

$$\beta \leq \frac{Pl(||x_i - x_j||)}{N_0 + \gamma \sum_{k \neq i,j} Pl(||x_k - x_j||)} \leq \frac{Pl(||x_i - x_j||)}{N_0},$$

which implies

$$l(||x_i - x_j||) \geq \frac{\beta N_0}{P}.$$

If we assume a power-law attenuation of the form $l(d) = d^{-\alpha}$, we obtain

$$||x_i - x_j|| \leq \left(\frac{\beta N_0}{P}\right)^{-\frac{1}{\alpha}} := r. \tag{3}$$

The connectivity graph in this model is thus a subgraph of the connectivity graph in the Boolean model with the range r defined above. Therefore, we can bound the number of paths by counting the number of path in the Boolean model.

We consider a slice of width r that cuts the network from top to bottom. Clearly, in the Boolean model, each vertex-disjoint path must have one node in this slice. The number of left-right crossings is thus limited by the number of nodes in the slice, which increases linearly with r. Because of (3, we know that r increases like $\beta^{-1/\alpha}$. So finally, we conclude that the number of paths increases like $\beta^{-1/\alpha}$.

On the other hand, according to Shannon's formula, the rate of the links is linear in β when β is small. Therefore, the actual horizontal throughput across the networks is of order $\beta \cdot \beta^{-1/\alpha} = \beta^{1-1/\alpha}$. For any $\alpha > 0$, this function tends to zero when β tends to zero. We conclude that lowering the threshold β eventually decreases the horizontal throughput. This reasoning is confirmed by the simulation result in the right-hand side of Figure 5.

6 Conclusion

We have seen that in the Boolean model as well as in the information theoretic connectivity model, full connectivity do not scale when the network size increases. At the contrary, if we allow for a (possibly very small) fraction of disconnected nodes, then the range (respectively the rate) does not need to be adjusted when the number of nodes tend to infinity.

When several flows have to share the available bandwidth and interferences are critical, full connectivity turns out to be very costly in terms of throughput. In fact, keeping the most isolated nodes connected consumes a lot of resource, and affects greatly the overall performance of the network. This situation leads to a trade-off between capacity and connectivity. Under several models, keeping the connectivity graph quite sparse leads to optimal throughput.

References

[1] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Trans. Inform. Theory*, 46(2):388–404, March 2000.

[2] M. Penrose and A. Pisztora. Large deviations for discrete and continuous percolation. *Adv. Appl. Prob.*, 28:29–52, 1996.

[3] M. Penrose. The longest edge of the random minimal spanning tree. *Ann. Appl. Probability*, 7:340–361, 1997.

[4] P. Gupta and P. R. Kumar. Critical power for asymptotic connectivity in wireless networks. *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*, 1998. Edited by W.M. McEneany, G. Yin, and Q. Zhang, (Eds.) Birkhuser.

[5] R. Meester and R. Roy. *Continuum percolation.* Cambridge University Press, 1996.

[6] O. Dousse, F. Baccelli, and P. Thiran. Impact of interferences on connectivity of ad hoc networks. *IEEE/ACM Trans. Networking*, to appear.

[7] X. Liu and R. Srikant. An information theoretic view of connectivity in random sensor networks. preprint, http://www.cs.ucdavis.edu/∼liu/research.html, 2004.

[8] O. Dousse, M. Franceschetti, and P. Thiran. Information theoretic bounds on the throughput scaling of wireless relay networks. In *submitted to IEEE Infocom 2005.*

[9] S. R. Kulkarni and P. Viswanath. A deterministic approach to throughput scaling in wireless networks. In *Proc. IEEE International Symposium on Information Theory*, November 2002.

[10] M. Franceschetti, O. Dousse, D. Tse, and P. Thiran. Closing the gap in the capacity of random wireless networks. In *Proc. of Information Theory Symposium (ISIT)*, Chicago, Illinois, July 2004.