

# A Deterministic Approach to Throughput Scaling in Wireless Networks\*

Sanjeev R. Kulkarni and Pramod Viswanath<sup>†</sup>

1 Nov, 2002

## Abstract

We address the problem of how throughput in a wireless network scales as the number of users grows. Following the model of Gupta and Kumar, we consider  $n$  identical nodes placed in a fixed area. Pairs of transmitters and receivers wish to communicate but are subject to interference from other nodes. Throughput is measured in bit-meters per second.

We provide a very elementary deterministic approach that gives achievability results in terms of three key properties of the node locations. As a special case, we obtain  $O(\sqrt{n})$  throughput for a general class of configurations. Results for random node locations can also be derived as special cases of the general result by verifying the growth rate of three parameters. For example, as a simple corollary of our result we obtain a stronger (almost sure) version of the  $\sqrt{n}/\sqrt{\log n}$  throughput for random node locations obtained by Gupta and Kumar. Results for some other interesting non-i.i.d. node distributions are also provided.

**Index Terms:** wireless networks, ad hoc networks, multi-hop, throughput, capacity, deterministic, individual sequence, random, scaling

---

\*This work was supported in part by the National Science Foundation under contract number ECS-9873451 and by the Army Research Office under grant number DAAD19-00-1-0466.

<sup>†</sup>Sanjeev R. Kulkarni is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544. His email address is [kulkarni@princeton.edu](mailto:kulkarni@princeton.edu). Pramod Viswanath is with the Department of ECE, University of Illinois, Urbana-Champaign, IL 61801. His email address is [pramodv@uiuc.edu](mailto:pramodv@uiuc.edu).

# 1 Introduction

The past decade has seen an emergence of wireless communication in a rapid and pervasive manner. Several innovations at the physical layer (smart use of multiple antennas at the transmitter and receiver and interference suppression techniques such as multiuser detection being some important examples), combined with advances in low power VLSI that allow synthesis of these algorithmic advances at the physical layer in hardware have driven this revolution. However, from an overall architecture point of view of wireless networks, the primary structure has been that of *base stations* connected with each other by a wired infrastructure. The base station performs duplex communication with mobiles in its neighborhood and acts as an access point for the mobile device to the wired network. A systematic study of the fundamental tradeoffs involved in the joint design of the wireless and wired architecture is a very important area of research; one that will significantly impact the way future wireless systems will be designed. In particular, the study of all wireless systems (or *wireless ad hoc networks*) is one aspect of the agenda, the protocols for which are being studied [7, 16, 17, 1].

One important facet of this research will be to shed light on how a completely wireless system scales with space and the number of users. Some significant steps towards such a systematic study have been initiated recently [3, 2, 4]. In [3], the authors studied a distributed all wireless communication model focusing on the case when the nodes are not mobile (or equivalently, the desired time scale of communication is assumed to be much faster than the mobility of the nodes). They showed that regardless of the initial location of the nodes and traffic pattern in a fixed area, the throughput in bit-meters per second cannot grow faster than  $\sqrt{n}$ , and for a very special arrangement of nodes and traffic pattern a throughput  $\sqrt{n}$  can be achieved. They also considered a random traffic pattern and random node distribution. In this case, they showed that the total throughput (in bits per second) can grow no faster than  $\sqrt{n}/\sqrt{\log n}$ , as a function of the number of nodes (denoted by  $n$ ), and they gave an achievability result that achieved this growth rate. These results assumed a simple communication model where the nodes only act as relays and the reliable reception of each transmission depends only on the signal to noise/interference ratio (SNR) at the receiver. If the transmitter-receiver pair of nodes are very close to each other, then reliable communication will cause little interference to the other nodes and the scenario is essentially that of a set of non-interfering point-to-point communication systems. However, in general, the transmitter-receiver pairs are not arbitrarily close to each other and an important physical insight that followed from the work of [3] was the need for local communication and use of multiple hops to reach the destination. The fact that nearest neighbors get closer as  $n$  grows means that the number of hops to reach the destination increases, which imposes a fundamental limit on the way throughput of the entire network scales as a function of  $n$ . Other related studies

are [5, 11].

In this paper, we revisit the throughput scaling problem of [3]. Our key contribution is the identification of deterministic properties in the location of the nodes that combined with a simple deterministic algorithm provides a throughput that is very easily computed. For example, we obtain the  $\sqrt{n}$  order growth of bit-meters per second for a general class of transmitter-receiver locations.

Moreover, our results capture fundamental properties of the set of node locations that affect throughput growth and can be easily calculated for different distributions of random node locations. Thus, the deterministic approach can be used to make almost sure statements for various random node location and traffic patterns. As one special case, we easily derive, and strengthen, the main results of [3]. That is, for i.i.d. uniform node locations and random transmitter-receiver pair case our algorithm obtains a throughput of  $\sqrt{n}/\sqrt{\log n}$  almost surely. We also consider some interesting non-i.i.d. node distributions.

## 2 Deterministic Lower Bounds

We begin with a description of the physical model that is very similar to the one introduced in [3]. The area of communication is assumed to be restricted to a *square* of area 1 meter-square. Assuming  $n$  nodes in the network, we denote the physical location of the node  $i$  by  $S(i) \stackrel{\text{def}}{=} (x_i, y_i)$  with  $x_i$  and  $y_i$  representing the two physical location coordinates. We assume that the nodes are not moving (at least, relative to the time scale of communication). In [3], the authors show that for every configuration of node locations and traffic patterns, the total throughput cannot grow faster than  $\sqrt{n}$  bit-meter per second and also demonstrate a specific instance of physical node locations and traffic pattern where the growth is indeed of that order. We provide a simple transmission scheduling scheme that combined a deterministic condition on the physical location of the nodes allows for an easy computation of the throughput. Using this, we can identify a general class of physical locations of the nodes that for every traffic pattern has a capacity growth of  $\sqrt{n}$  bit-meter per second. Finally, we can use this deterministic approach to analyze the situation of random locations and random transmitter-receiver pairs.

For expositional ease, we restrict ourselves to a simple model of communication between the nodes, one of the models in [3] called the *Protocol* model. Suppose node  $i$  transmits to node  $j$ . We assume that the transmission is successful if the distance between the two nodes is smaller by a constant factor than the distance between  $j$  and any

other node that is simultaneously transmitting. In our notation, we have a successful transmission if for every other node  $k$  that is transmitting simultaneously with  $i$ ,

$$|S(k) - S(j)| \geq (1 + \Delta) |S(i) - S(j)| .$$

Here  $\Delta$  is a fixed positive constant that represents the guard zone in the protocol model. Since the performance measure is the rate of growth of throughput as a function of the number of nodes, the role of the separation parameter  $\Delta$  is insignificant. We assume that time is slotted (and each slot is 1 second long, for simplicity) and a successful transmission in any slot conveys a fixed amount of data (denoted by a “packet”, henceforth).

## 2.1 Deterministic Structure and Conditions

With  $n$  nodes, the node configuration is defined by the set of physical node locations  $\{S(1), \dots, S(n)\}$ , together with a specification of transmitter-receiver pairs. We assume that half of the nodes are designated as transmitters and the other half as receivers, with each receiver identified uniquely with one transmitter node. We emphasize that all nodes are identical and act as both transmitters and receivers. The specific designation of a node as a transmitter means that new traffic originates from that node. Since we are interested in how the throughput grows with  $n$ , we actually consider a sequence of configurations, indexed by the number of nodes  $n$ .

If the transmitter-receiver pairs are arbitrarily close to each other yet far from other transmitter-receiver pairs, the communication problem reverts to a set of non-interfering point-to-point communications problem with a well understood solution. The results in [2] can be rephrased to conclude that if every transmitter-receiver distance is of the order  $O\left(\frac{1}{\sqrt{n}}\right)$ , and far from other transmitter receiver pairs, then we are in the point-to-point non-interfering scenario. On the other hand, if all the transmitter-receiver pairs are close to one another, then we can effectively scale the area of the domain under consideration. Thus, here we are interested in the scenario when the transmitter and corresponding receivers are not arbitrarily close to each other. In particular, we will require the sum of the transmitter-receiver pair distances to be  $O(n)$ . We will show that this condition is satisfied almost surely when the transmitters and receivers are placed randomly according to various distributions.

These issues also relate to the use of bit-meters per second to measure throughput as opposed to simply bits per second. If bits per second is used then upper bounds tighter than  $O(n)$  are not possible for *every* configuration of nodes. We could construct configurations with transmitters arbitrarily close to corresponding receivers yet far from other nodes, and thereby achieve  $O(n)$  bits/sec by simply avoiding interference. However,

as shown in [3] this problem is avoided by using bit-meters per second, so that upper bounds on throughput can be obtained for *every* configuration. For the lower bounds as presented here, the throughput results would also hold using bits per second, and in fact, in this case we would not need the condition that the sum of transmitter-receiver distances be  $O(n)$ . However, to have the matching upper bounds from [3], we use bit-meters/sec and impose the condition on the sum of transmitter-receiver distances.

We are now ready to describe the additional structure we impose on the configurations, summarize our deterministic condition, and describe the corresponding simple scheduling algorithm. Consider a fixed sequence of configurations. For every  $n$ , we divide the unit area into  $\frac{1}{s_n^2}$  squares of area  $s_n^2$  each by drawing equally spaced lines horizontal and vertical lines with spacing  $s_n$ . We refer to these smaller squares as *squarelets* and we require each squarelet to have at least one node.

An important quantity in our scheduling algorithm and network throughput is captured by the *crowding factor* of nodes in the squarelets. The crowding factor, denoted by  $c_n$ , is defined to be the maximum number of nodes in any of the  $\frac{1}{s_n^2}$  squarelets.  $c_n$  depends on the number of nodes  $n$ , the physical node locations, and our choice of squarelet size  $s_n$ . For any sequence of node locations, as the squarelet area size increases, the crowding factor generally increases and always satisfies the relation  $c_n s_n^2 \geq n$ .

We summarize the parameters and conditions on the structure below:

1. **Tx-Rx Distances:** The sum of transmitter-receiver distances grows at least as  $O(n)$ .
2. **Nonempty Squarelets:**  $s_n$  is such that for sufficiently large  $n$ , no squarelet is empty.
3. **Crowding Factor:**  $c_n$  is the maximum number of nodes in any squarelet.

## 2.2 Equivalence Classes of Squarelets

To maximize throughput, we would like to schedule many simultaneous transmissions. Towards this end, in our scheduling algorithm, transmission will occur only between neighboring squarelets. That is, when a given node is scheduled to transmit, it will send packets only to nodes in the four squarelets adjacent to itself. Since this is a local transmission, other squarelets sufficiently far away (depending on the parameter  $\Delta$  as discussed below) could simultaneously transmit without interference. This allows us to

define “equivalence classes” of squarelets such that one node in each squarelet of the equivalence class can transmit simultaneously without interfering with one another.

For a systematic method, for each squarelet, we define its equivalence class to be those squarelets that are a vertical and horizontal distance of exactly some multiple of an integer number of squarelets away. The shaded squarelets in Figure 1 represent one such equivalence class.

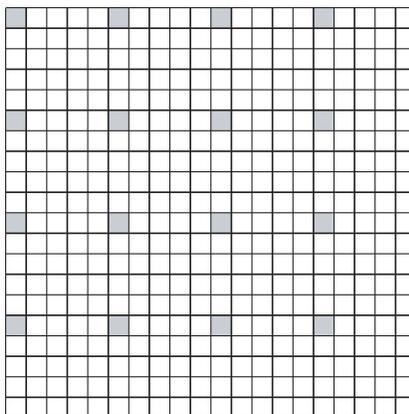


Figure 1: An example of an equivalence class of squarelets with  $K = 5$ . The shaded squarelets all belong to the same equivalence class.

The distance needed to allow squarelets in the same equivalence to simultaneously transmit without interference can be computed in terms of  $\Delta$  as follows. Let  $K$  be an integer such that squarelets in an equivalence class are integer multiples of  $K$  apart.

Suppose node  $i$  is transmitting to node  $j$ . Since transmissions will be between neighboring squarelets, the maximum distance between  $i$  and  $j$  is  $\sqrt{5}s_n$ . The closest another simultaneously transmitting node  $k$  (in a different squarelet of the same equivalence class) could be to node  $j$  is  $(K - 2)s_n$ . (For example, see the arrangement at the bottom left of Figure 2, and the node in the top right.) To avoid interference, we need  $(K - 2)s_n \geq (1 + \Delta)\sqrt{5}s_n$ , so that  $K \geq 2 + (1 + \Delta)\sqrt{5}$ . Thus, we can take  $K = \lceil 2 + (1 + \Delta)\sqrt{5} \rceil$ . Note that there are only a finite number of equivalence classes (namely,  $K^2$  of them).

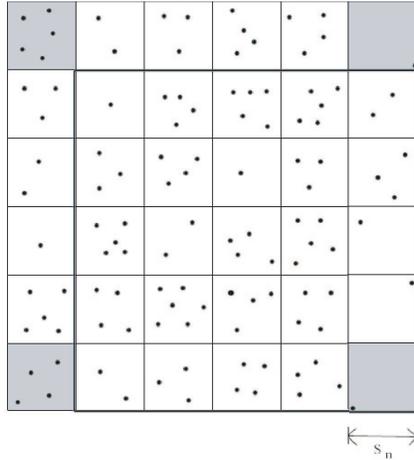


Figure 2: A zoomed in version showing several squarelets and nodes within the squarelets. The arrangement resulting in maximum transmission distance of a single hop is shown in the bottom right. This together with the node in the top right shows the arrangement governing the bound on  $K$ .

### 2.3 Scheduling Algorithm and Throughput

As a prelude to describing the scheduling algorithm, we review a closely related problem studied intensely in the parallel and distributed computing research community. Consider a square of  $l^2$  processing units (PUs) with  $l$  in each row and column as shown in Fig. 3. Each processing unit can communicate with its immediate vertical and horizontal neighbors in a single slot, i.e., every PU can receive and transmit to its neighbors (no more than 4) simultaneously. Here we are continuing the assumption of a slotted time and communication occurring over slots and communication occurs between neighboring PUs and contains a *packet*. Suppose each PU is the source and destination of exactly  $k$  packets. The problem of routing the total  $kl^2$  packets to their destinations is a well studied problem in the literature under the nomenclature  $k \times k$  *permutation routing*. The following result characterizes the performance of  $k \times k$  permutation routing algorithms with minimal queue length requirements at the PUs. We have denoted  $o(n)$  to denote a function that satisfies the property  $o(n)/n \rightarrow 0$  as  $n \rightarrow \infty$ .

**Lemma 2.1** [10, 8]  *$k \times k$  permutation routing in a  $l \times l$  mesh can be performed deterministically in  $kl/2 + o(kl)$  steps with maximum queue size at each PU equal to  $k$ . Further, every routing algorithm takes at least  $kl/2$  steps.*

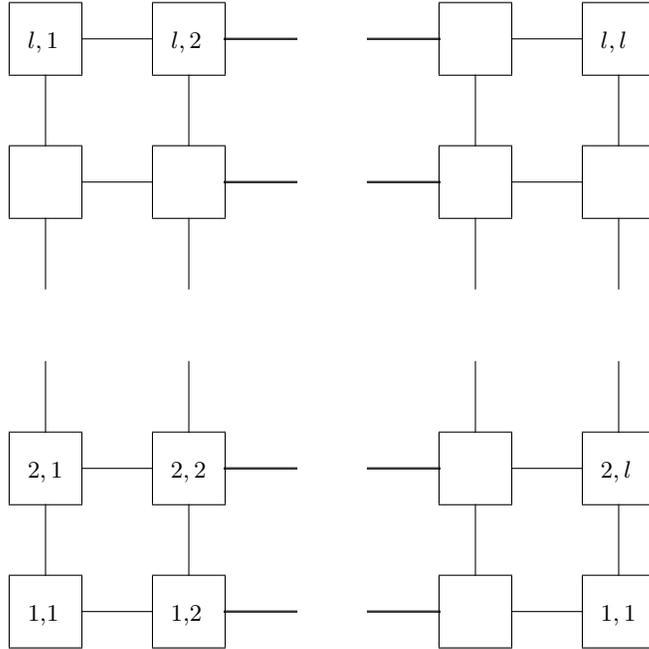


Figure 3: A mesh network of PUs each connected to its vertical and horizontal neighbors.

We can now use the  $k \times k$  permutation routing algorithms of packets across a mesh of PUs to our ad hoc wireless network. Let us fix  $n$ , the number of users. The first correspondence is between PUs and squarelets by letting  $l = 1/s_n$ . Next, we let each user have  $m$  packets and since a squarelet can have no more than  $c_n$  users, the total number of packets in any squarelet is no more than  $mc_n$ . By letting  $k = mc_n$  and observing that the users in each squarelet transmit and receive no more than  $mc_n$  packets, we have formed a correspondence in the traffic pattern between our wireless network and the mesh network of PUs, i.e., we associate the  $k$  packets associated with a PU with the packets of the users contained in the corresponding squarelet. By allowing a buffer of size  $mc_n$  for each user, we are ready to transfer the routing algorithm of Lemma 2.1 to one on the wireless network.

Communication in the mesh of PUs allows vertical and horizontal neighbors to communicate in a single step. Thus each PU can transmit and receive up to 4 packets in the same slot. In our wireless network, we restrict communication to be between users in neighboring squarelets (vertical and horizontal neighbors), but only users in the equivalence classes of squarelets can transmit over the same slot. Further, a user can only transmit one packet in one slot. We make the routing protocol on the mesh of PUs *sequential* to meet these two constraints. First we make sequential the transmissions of the PUs which are not in the same equivalence class. Since there are  $K^2$  equivalence classes,

this increases the total number of steps by this factor. Second, we make sequential the transmissions of a given PU. Since there are no more than 4 simultaneous transmissions of any PU (corresponding to the maximum number of neighbors), this increases the total number of steps by no more than a factor of 4.

Thus we can convert the routing protocol on the mesh of PUs to the wireless network and conclude that the  $m$  packets of each of the  $n/2$  users reach their destination in number of slots equal to

$$4K^2 \frac{kl}{2} = \frac{2K^2 mc_n}{s_n}.$$

Further, due to our condition on Tx-Rx distances, the total distance traveled by the  $mn/2$  packets is at least order  $O(mn)$ . This allows us to make the following statement.

**Proposition 2.1** *The throughput in bit-meters per second (and bits per second) for a network with squarelet size  $s_n$  and crowding factor  $c_n$  is at least  $O(\frac{ns_n}{c_n})$ .*

In particular, when the node location sequence is such that  $s_n$  can be chosen to be  $\frac{1}{\sqrt{n}}$  and the crowding factor is of order 1, then we recover the largest possible growth of throughput,  $\sqrt{n}$  bit-meters per second.

### 3 I.i.d. Node Locations

In this section, we show how to use our deterministic results to easily recover the achievability result of [3] with random node locations. Suppose the nodes are equally divided into transmitters and corresponding receivers and their locations are chosen independently and randomly from a uniform distribution on the square. Our main result is the following.

**Proposition 3.1** *A throughput  $O(\sqrt{n}/\sqrt{\log n})$  is achievable almost surely.*

This result will follow immediately from our deterministic result of the previous section (Proposition 2.1) and the following claim.

**Claim 3.1** 1. *For squarelet size  $s_n = \frac{\sqrt{3 \log n}}{\sqrt{n}}$ , no squarelet is empty almost surely.*

2. For  $s_n$  as above, the crowding factor satisfies  $c_n \leq 3e \log n$  almost surely.
3. The sum of the distances between transmitter-receiver pairs is  $O(n)$  meters almost surely.

Once we show this claim, we have from Proposition 2.1 that, for almost every realization of transmitter and receiver node locations the throughput growth is at least of the order  $\frac{\sqrt{n}}{\sqrt{\log n}}$  bit-meter per second. Since the average distance between a transmitter-receiver pair is of the order 1, this result recovers the bit per second result of [3] in the random node location case. Much of the technical difficulty of [3] in proving that no node is overloaded is easily avoided here by exploiting our deterministic structure and a simple scheduling algorithm. We now proceed to show our claim above.

**Nonempty Squarelets** In an  $n$  node system, the probability that any fixed squarelet is empty is equal to  $(1 - s_n^2)^n$ . From a simple union bound and using the fact that there are  $\frac{1}{s_n^2}$  squarelets, we conclude that the probability that at least one node is empty is upper bounded by  $p_n \stackrel{\text{def}}{=} \frac{(1 - s_n^2)^n}{s_n^2}$ . With the value of  $s_n$  equal to  $\frac{\sqrt{3 \log n}}{\sqrt{n}}$ , we see that

$$p_n < \frac{\exp(-ns_n^2)}{s_n^2} \leq n^{-2},$$

using  $1 - x \leq \exp(-x)$  and hence that  $\sum_{n=1}^{\infty} p_n < \infty$ . Hence, using the Borel-Cantelli lemma, we conclude that almost surely no squarelet is empty for sufficiently large  $n$ .

**Crowding Factor** We fix  $n$  and look at one particular squarelet. The number of nodes in this squarelet (denoted by, say,  $Z_n$ ) is a binomial random variable with parameters  $(s_n^2, n)$ . By a simple Chernoff bound we have, for all  $a > 0, \theta > 0$ ,

$$\mathbb{P}[Z_n > a \log n] \leq \frac{\mathbb{E}[\exp(\theta Z_n)]}{\exp(\theta a \log n)}. \quad (1)$$

Now,

$$\begin{aligned} \mathbb{E}[\exp(\theta Z_n)] &= \left(1 + (e^\theta - 1) s_n^2\right)^n, \\ &\leq n^{3(e^\theta - 1)}. \end{aligned}$$

Choosing  $\theta = 1$  and  $a = 3e$  in (1) we arrive at

$$\mathbb{P}[Z_n > 3e \log n] \leq \frac{1}{n^3}.$$

Using a simple union bound we see that

$$\mathbb{P}[c_n > 3e \log n] \leq \frac{1}{n^3 s_n^2} \leq \frac{1}{n^2}.$$

Using the Borel-Cantelli lemma, we conclude that, almost surely, the crowding factor is no more than  $3e \log n$ .

**Tx-Rx Distances** Finally, we need to estimate the sum of the transmitter-receiver pair distances. Since all distances are less than  $\sqrt{2}$ , the distance squared is less than the product of  $\sqrt{2}$  and the distance. Further, the sum of scaled distances is lower bounded by the sum of squared scaled distances. The squared distance between any transmitter and receiver has the same distribution as that of  $(x_1 - y_1)^2 + (x_2 - y_2)^2$  with  $x_1, x_2, y_1, y_2$  being i.i.d. uniform on  $[0, 1]$  random variables. Further, the squared distances between the transmitter-receiver pairs (there are  $\frac{n}{2}$  such pairs) are all independent and identically distributed. Hence we have, for  $\theta > 0, a > 0$ ,

$$\mathbb{P}\left[\text{sum of distances between transmitter-receiver pairs} < \frac{n}{a}\right] \leq \frac{\left(\mathbb{E}\left[\exp\left(-\theta(x-y)^2\right)\right]\right)^n}{\exp\left(-\frac{\theta\sqrt{2}n}{a}\right)}, \quad (2)$$

where  $x$  and  $y$  are i.i.d. uniform on  $[0, 1]$  random variables. Now

$$\begin{aligned} \mathbb{E}\left[\exp\left(-\theta(x-y)^2\right)\right] &= 2 \int_0^1 \int_0^y \exp\left(-\theta t^2\right) dt dy, \\ &= \int_0^1 (1-t) \exp\left(-\theta t^2\right) dt, \\ &< \frac{\sqrt{\pi}}{\sqrt{\theta}}. \end{aligned} \quad (3)$$

Substituting (3) in (2) and choosing  $a = \sqrt{2}\theta$  and  $\theta = 2\pi e^2$  we arrive at

$$\mathbb{P}\left[\text{sum of distances between transmitter-receiver pairs} < \frac{n}{2\pi e^2}\right] < \frac{1}{2^{n/2}}. \quad (4)$$

Appealing to the Borel-Cantelli lemma and using (4) we conclude that almost surely the sum of the distances between transmitter-receiver pairs grows at least linearly with  $n$ . This concludes the proof of the claim.  $\square$

## 4 Some Non-i.i.d. Examples

To further show the power of the deterministic approach, in this section we consider some examples where the node locations are random but with some dependence between

nodes. With our approach, again we need only quantify the growth rates of  $s_n$  and  $c_n$ , and verify that no squarelet is empty and that the sum of transmitter receiver distances grows as  $O(n)$ . In the examples, below, the lack of independence only slightly complicates the analysis.

## 4.1 Users on Roads

Consider the  $n$  node locations drawn randomly as follows. In the unit square, independently draw  $\sqrt{n}/2$  horizontal lines uniformly, and independently draw  $\sqrt{n}/2$  vertical lines uniformly. On each of the  $\sqrt{n}$  lines, draw  $\sqrt{n}$  nodes independently and uniformly on the line. We can think of the lines as “roads” and the nodes as users randomly placed along the roads.

From the intuition in the previous sections, we might expect the throughput to suffer somewhat since the users will be more concentrated in some sense. Hence, we might expect to need a larger  $s_n$  to guarantee that no squarelet is empty, but this might result in a larger crowding factor  $c_n$ . However, it turns out that we don’t lose much in the throughput as the following result shows.

**Proposition 4.1** *For nodes distributed along lines as described above, a throughput  $O(\sqrt{n}/(\log n))$  is achievable almost surely.*

As before, this result will follow immediately from Proposition 2.1 once we show the following claim.

- Claim 4.1**
1. *For squarelet size  $s_n = \frac{6 \log n}{\sqrt{n}}$ , no squarelet is empty almost surely.*
  2. *For  $s_n$  as above, the crowding factor satisfies  $c_n \leq 7e^6(\log n)^2$  almost surely.*
  3. *The sum of the distances between transmitter-receiver pairs is  $O(n)$  meters almost surely.*

**Nonempty Squarelets** First, consider a fixed squarelet. The probability that no node from a particular line falls in this squarelet is given by

$$\mathbb{P}[\text{no node from one line falls in fixed squarelet}] = (1 - s_n) + s_n(1 - s_n)^{\sqrt{n}}$$

since with probability  $1 - s_n$  the line does not even intersect the squarelet, and when it does (with probability  $s_n$ ), each of the  $\sqrt{n}$  nodes has an independent probability of  $1 - s_n$  of not falling in the squarelet.

Therefore, with  $\sqrt{n}$  independent lines, the probability that a fixed squarelet is empty is

$$\mathbb{P}[\text{a fixed squarelet is empty}] = \left[ (1 - s_n) + s_n(1 - s_n)^{\sqrt{n}} \right]^{\sqrt{n}}.$$

Hence, with a union bound, the probability of no empty squarelets satisfies

$$\begin{aligned} \mathbb{P}[\text{no empty squarelets}] &\leq \frac{1}{s_n^2} \left[ (1 - s_n) + s_n(1 - s_n)^{\sqrt{n}} \right]^{\sqrt{n}}, \\ &\leq \frac{1}{s_n^2} e^{-\sqrt{n}(s_n - s_n(1 - s_n)^{\sqrt{n}})}, \quad \text{since } 1 + x \leq e^x, \\ &\leq \frac{1}{s_n^2} e^{-s_n \sqrt{n}(1 - e^{-s_n \sqrt{n}})} \quad \text{using } 1 - x \leq e^{-x} \quad \text{again.} \end{aligned}$$

If we take  $s_n = 6 \log n / \sqrt{n}$ , then  $e^{-s_n \sqrt{n}} = 1/n^6$ , so that

$$\begin{aligned} \mathbb{P}[\text{no empty squarelets}] &\leq \frac{n}{(36 \log n)^2} \left( \frac{1}{n^6} \right)^{(1 - \frac{1}{n^6})} \\ &\leq \frac{1}{36n^4 (\log n)^2} \quad \text{for } n \geq 2 \end{aligned}$$

Hence, by the Borel-Cantelli lemma, almost surely no squarelet is empty for sufficiently large  $n$ .

**Crowding Factor** We fix  $n$  and focus on one specific squarelet. The number of nodes that fall in this squarelet is

$$Z_n = \sum_{i=1}^X Y_i,$$

where  $X$  is a Binomial random variable with parameter  $(s_n, \sqrt{n})$  denoting the number of lines (both horizontal and vertical) that pass through this squarelet.  $Y_1, \dots, Y_{\sqrt{n}}$  are i.i.d. random variables that are Binomial distributed with parameter  $(s_n, \sqrt{n})$ .  $Y_i$  denotes the number of nodes in line  $i$  that fall in a squarelet through which the line  $i$  passes. From the structure of the node locations, these random variables are jointly independent of  $X$ . As before, by a simple Chernoff bound we have, for all  $a > 0, \theta > 0$ ,

$$\mathbb{P} \left[ Z_n > a (\log n)^2 \right] \leq \frac{\mathbb{E} [\exp(\theta Z_n)]}{\exp(\theta a (\log n)^2)}. \quad (5)$$

Now,

$$\begin{aligned}\mathbb{E}[\exp(\theta Z_n)] &= \left(1 + s_n \left[ \left(1 + (e^\theta - 1) s_n\right)^{\sqrt{n}} - 1 \right]\right)^{\sqrt{n}}, \\ &\leq \exp\left(\sqrt{n} s_n \left[ \left(1 + (e^\theta - 1) s_n\right)^{\sqrt{n}} - 1 \right]\right).\end{aligned}\quad (6)$$

Now, we have

$$\left(1 + (e^\theta - 1) s_n\right)^{\sqrt{n}} \leq \exp\left(\sqrt{n} s_n (e^\theta - 1)\right).\quad (7)$$

Choosing

$$\theta = \log\left(1 + \frac{1}{\log n}\right),$$

in (7) and substituting in (6) we arrive at

$$\mathbb{E}[\exp(\theta Z_n)] \leq \exp\left(6e^6 \log n\right).\quad (8)$$

Here we used the value of  $s_n$  to be  $\frac{6 \log n}{\sqrt{n}}$ . Since

$$\log n \log\left(1 + \frac{1}{\log n}\right) \rightarrow 1, \quad \text{as } n \rightarrow \infty,$$

by choosing  $a = 7e^6 > 3(1 + 2e^6)$  we have with the value of  $\theta$  above that

$$a\theta \log n \geq 3(1 + 2e^6),\quad (9)$$

for all large enough  $n$ . Substituting (8) and (9) in (5) we have, for all large enough  $n$ ,

$$\mathbb{P}\left[Z_n > 7e^6 \log n\right] \leq \frac{\exp(6e^6 \log n)}{\exp(3(1 + 2e^6) \log n)} = \frac{1}{n^3}.$$

Using a simple union bound we see that

$$\mathbb{P}\left[c_n > 7e^6 (\log n)^2\right] \leq \frac{1}{n^3 s_n^2} \leq \frac{1}{n^2}.$$

Using the Borel-Cantelli lemma, we conclude that, almost surely, the crowding factor is no more than  $7e^6 (\log n)^2$ .

**Tx-Rx Distances** We begin with the observation that since all distances are less than  $\sqrt{2}$ , the distance squared is less than the product of  $\sqrt{2}$  and the distance. The marginal distribution of this squared distance is either that of  $(x_1 - y_1)^2$  or that of  $(x_1 - y_1)^2 + (x_2 - y_2)^2$  with  $x_1, x_2, y_1, y_2$  being i.i.d. uniform on  $[0, 1]$  random variables depending on

whether the transmitter and receiver lie on the same line (horizontal or vertical) or on different ones. A small technical complication, as compared to the i.i.d. node location case, arises since the distances are not jointly independent. To overcome this, we first observe that each distance squared term contains at least one term of the form  $(x - y)^2$  (for  $x, y$  i.i.d. uniform on  $[0, 1]$ ). Further, conditioned on every realization of source and destination pairs being located on either the horizontal or vertical line) there is a choice of these terms for each source and destination pair that are jointly independent. One way to see this choice is by considering the horizontal (vertical) distance squared from source to destination if the source lies on a horizontal (vertical) line. Hence the probability that the sum of the  $\frac{n}{2}$  squared distances is smaller than  $\frac{n}{a}$  is upper bounded by

$$\mathbb{P} \left[ \sum_{i=1}^{\frac{n}{2}} (x_i - y_i)^2 < \frac{n\sqrt{2}}{a} \right]$$

for every  $a > 0$  with  $x_i, y_i$  being i.i.d. uniform on  $[0, 1]$  for  $i = 1, \dots, n/2$ . Now using the upper bound technique of (2) and the upper bound of (3) and choosing  $a = \sqrt{2}\theta$  and  $\theta = 2\pi e^4$  we arrive at (analogous to (4))

$$\mathbb{P} \left[ \text{sum of distances between transmitter-receiver pairs} < \frac{n}{2\pi e^4} \right] < \frac{1}{2^{n/4}}.$$

An appeal to the Borel-Cantelli lemma concludes the proof of our claim that the sum of the distances between transmitter-receiver pairs is  $O(n)$  almost surely.  $\square$

## 4.2 Users in Neighborhoods

Now, consider the  $n$  node locations drawn randomly as follows. Fix a “neighborhood” size  $0 < b < 0.5$ . In a  $(1 - b) \times (1 - b)$  square concentric with the original unit square, independently draw  $\sqrt{n}$  locations. These locations will be the “centers” of neighborhoods around which users will be clustered. For each neighborhood center,  $\sqrt{n}$  nodes are placed uniformly and independently in a  $b \times b$  square about the neighborhood center. Thus, the  $\sqrt{n}$  neighborhood locations are randomly chosen, and  $\sqrt{n}$  users are randomly placed in each neighborhood, for a total of  $n$  users.

As with users on roads, we might expect the throughput to suffer somewhat since the users will be more concentrated in some sense. However, it turns out that we don’t lose any throughput as compared to the i.i.d. node location case, as the result below shows. In our considerations, the squarelet size  $s_n \rightarrow 0$  while the neighborhood size  $b$  is fixed. Since we are interested in properties for sufficiently large  $n$ , we assume without loss of generality that  $s_n < \min\{b, 1 - b\}$ .

**Proposition 4.2** *For nodes distributed in neighborhoods as described above, a throughput  $O(\sqrt{n}/\sqrt{\log n})$  is achievable almost surely.*

As before, this result will follow immediately from Proposition 2.1 once we show the following claim.

- Claim 4.2**
1. *For squarelet size  $s_n = \frac{b\sqrt{3\log n}}{\sqrt{n}}$ , no squarelet is empty almost surely.*
  2. *For  $s_n$  as above, the crowding factor satisfies  $c_n \leq 3e \log n$  almost surely.*
  3. *The sum of the distances between transmitter-receiver pairs is of the order  $n$  meters almost surely.*

**Nonempty Squarelets** As before, we first bound the probability that a fixed squarelet is empty, and then use a union bound to bound the probability of no empty squarelets. Unlike the case of users on roads, here there are edge effects. The four corner squarelets have the largest probability of being empty. Thus, an upper bound for these will also provide an upper bound for all other squarelets.

The probability that none of the  $\sqrt{n}$  nodes from a fixed neighborhood falls in one of the corner squarelets is

$$\left(1 - \frac{s_n^2}{(1-b)^2}\right) + \frac{s_n^2}{(1-b)^2} \left(1 - \frac{s_n^2}{b^2}\right)^{\sqrt{n}},$$

where the first term is the probability that the neighborhood does not intersect the squarelet and the second term is the probability of no nodes in the squarelet in the case that they intersect. We conclude that the probability of no node from one neighborhood falls in any fixed squarelet is upper bounded by

$$\left(1 - \frac{s_n^2}{(1-b)^2}\right) + \frac{s_n^2}{(1-b)^2} \left(1 - \frac{s_n^2}{b^2}\right)^{\sqrt{n}}.$$

Therefore, with  $\sqrt{n}$  independent neighborhoods, the probability that a fixed squarelet is empty satisfies

$$\mathbb{P}[\text{a fixed squarelet is empty}] \leq \left[ \left(1 - \frac{s_n^2}{(1-b)^2}\right) + \frac{s_n^2}{(1-b)^2} \left(1 - \frac{s_n^2}{b^2}\right)^{\sqrt{n}} \right]^{\sqrt{n}}.$$

Hence, with a union bound, the probability of no empty squarelets satisfies

$$\begin{aligned}
\mathbb{P}[\text{no empty squarelets}] &\leq \frac{1}{s_n^2} \left[ \left( 1 - \frac{s_n^2}{(1-b)^2} \right) + \frac{s_n^2}{(1-b)^2} \left( 1 - \frac{s_n^2}{b^2} \right)^{\sqrt{n}} \right]^{\sqrt{n}} \\
&\leq \frac{1}{s_n^2} \left( 1 - \frac{s_n^2}{b^2} \right)^n \quad \text{since } s_n < 1-b \text{ so that } \frac{s_n^2}{(1-b)^2} < 1 \\
&\leq \frac{1}{s_n^2} e^{-\frac{ns_n^2}{b^2}}
\end{aligned}$$

If we take  $s_n = b\sqrt{3\log n}/\sqrt{n}$ , then  $e^{-ns_n^2/b^2} = 1/n^3$ , so that

$$\mathbb{P}[\text{no empty squarelets}] \leq \frac{1}{3b^2n^2 \log n}$$

Hence, by the Borel-Cantelli lemma, almost surely no squarelet is empty for sufficiently large  $n$ .

**Crowding Factor** We fix  $n$  and focus on one specific squarelet. The central squarelet has the largest probability of crowding due to the nature of node locations. Thus an upper bound on the crowding probability in the central squarelet also provides a corresponding upper bound for any other squarelet. Now consider any specific neighborhood with respect to the central squarelet. Depending on the location of this neighborhood, there may be no intersection with the central squarelet and no node in this neighborhood will contribute to crowding in the central squarelet. Hence an upper bound to the probability of any one node in the specific neighborhood under consideration lying in the central squarelet is provided by assuming that the central squarelet is a subset of the neighborhood. This upper bound is equal to  $\frac{s_n^2}{b^2}$ , the ratio of the area of the squarelet to that of the neighborhood. Since there are  $\sqrt{n}$  nodes uniformly distributed in every neighborhood, the probability that  $k$  of them lie in the central squarelet is upper bounded by

$$\mathbb{P}[Y_1 \geq k],$$

where  $Y_1$  is a Binomial random variable with parameters  $\left(\frac{s_n^2}{b^2}, \sqrt{n}\right)$ . Since each neighborhood is chosen according to an i.i.d. distribution, the probability that the crowding factor in the central squarelet is larger than  $a \log n$  is upper bounded by

$$\mathbb{P}\left[\sum_{i=1}^{\sqrt{n}} Y_i \geq a \log n\right],$$

where  $Y_1, \dots, Y_{\sqrt{n}}$  are i.i.d. Binomial random variables with parameters  $\left(\frac{s_n^2}{b^2}, \sqrt{n}\right)$ . Using a simple Chernoff bound, we now have, for all  $a > 0, \theta > 0$ ,

$$\mathbb{P}[\text{Number of nodes in central squarelet} > a \log n] \leq \frac{(\mathbb{E}[\theta Y_1])^{\sqrt{n}}}{\exp(\theta a \log n)}. \quad (10)$$

As before, we have

$$\mathbb{E}[\exp(\theta Y_1)] = \left(1 + \frac{(e^\theta - 1) s_n^2}{b^2}\right)^{\sqrt{n}}.$$

Substituting this in (10), we arrive at

$$\begin{aligned} \mathbb{P}[\text{Nodes in central squarelet} > a \log n] &\leq \frac{\left(1 + \frac{(e^\theta - 1) s_n^2}{b^2}\right)^n}{\exp(\theta a \log n)}, \\ &\leq \frac{\exp\left(\frac{(e^\theta - 1) n s_n^2}{b^2}\right)}{\exp(\theta a \log n)}, \end{aligned} \quad (11)$$

$$= \exp\left(\left(3(e^\theta - 1) - \theta a\right) \log n\right). \quad (12)$$

We used the inequality  $e^x \geq 1 + x$  in arriving at (11) and substituted the value of  $s_n$  in arriving at (12). Choosing  $\theta = 1$  and  $a = 3e$  in (12), we conclude that

$$\mathbb{P}[\text{Number of nodes in any fixed squarelet} > a \log n] \leq \frac{1}{n^3}.$$

As before, using a simple union bound we have

$$\mathbb{P}[c_n > 3e \log n] \leq \frac{1}{n^3 s_n^2} \leq \frac{1}{n^2}.$$

Using the Borel-Cantelli lemma, we conclude that, almost surely the crowding factor is no more than  $3e \log n$ .

**Tx-Rx Distances** A random variable  $X$  is stochastically larger than another random variable  $Y$  if for every  $a$ , the probability that  $X$  is larger than  $a$  is an upper bound to the probability that  $Y$  is larger than  $a$ , see [13]. We first observe that the distance between a transmitter and receiver node pair is *stochastically larger* when the two nodes are in different neighborhoods as compared to the situation when they lie in the same neighborhood. Hence the probability that the sum of the distances between transmitter-receiver pairs is less than  $d_n$  is upper bounded by the probability of the same event conditioned on the situation when the transmitter and receiver lie in the same neighborhood. This latter probability is the calculation of i.i.d. node locations in an area of size  $b$ . This calculation was done in Section 3 for  $b = 1$  but since  $b$  is fixed (w.r.t. the scaling in  $n$ ) we conclude, using the calculation of (4), that almost surely the sum of transmitter-receiver pair distances is  $O(n)$ .  $\square$

## 5 Conclusions

In this paper, we have demonstrated the power of a deterministic approach to studying the scaling of throughput with number of nodes in a wireless network. This deterministic structure combined with a simple scheduling algorithm allows us to easily obtain achievability results on throughput for general configurations. From the deterministic results, we can also recover results for random node locations by verifying that appropriate conditions hold almost surely. For example, we get a simple derivation of the achievability results of [3], and we also that our conditions can be useful in calculating the throughput scaling of networks where the initial node locations are not uniformly and independently located in the network. For example, the initial node locations could be along straight lines (modeling train lines or streets/highways), or nodes could be more densely located in particular parts of the network (modeling neighborhoods or other high density areas). In each case, verifying some simple conditions leads to corresponding lower bounds on the throughput scaling. The deterministic view taken in this paper is in line with recent work on individual sequences in information theory, statistics and learning theory (cf. [6, 9, 15, 14]).

Fundamental upper bounds to throughput (as those promised by network information theory) are an important area of research that will shed insight into how good the current achievability results are. Recent results in [12] show upper bounds for certain configurations of node locations, but closing the gap with lower bounds remains open.

## 6 Acknowledgements

This work was initiated when P. Viswanath was at Flarion Technologies and S. R. Kulkarni was visiting there. The authors thank Rajiv Laroia at Flarion technologies for providing the conducive environment.

## References

- [1] J. Broch, D. Johnson and D. Maltz, “The dynamic source routing protocol for mobile ad hoc networks”, Dec. 1998, IETF Internet draft.
- [2] M. Grossglauser and D. Tse, “Mobility increases the capacity of ad hoc wireless networks”, *IEEE INFOCOM*, 2001.

- [3] P. Gupta and P. R. Kumar, "Capacity of wireless networks", *IEEE Trans. on Information Theory*, Vol. 46(2), pp. 388-401, Mar 2000.
- [4] P. Gupta, "Design and Performance Analysis of Wireless Networks", *PhD Thesis*, Aug. 2000, University of Illinois Urbana-Champaign.
- [5] P. Gupta and P. R. Kumar, "Internets in the sky: The capacity of three dimensional wireless networks," *Communications in Information and Systems*, Vol. 1(1), pp. 33-50, January 2001.
- [6] D. Haussler, J. Kivinen, and M. Warmuth, "Tight worst-case loss bounds for predicting expert advice", *Proc. European Conf. on Computational learning theory*, 1994.
- [7] IEEE Computer Society LAN MAN Standards committee, "Wireless LAN medium access control and physical layer specifications", IEEE Standard 802.11, 1997.
- [8] M. Kaufamn, J. F. Sibeyn and T. Suel, "Derandomizing routing and sorting algorithms for meshes", *Proc. 5th Symp. on Discrete Algorithms*, LNCS, Springer-Verlag, 1995.
- [9] S. R. Kulkarni and S. E. Posner, "Rates of convergence for nearest neighbor estimation under arbitrary sampling", *IEEE Trans. Inform. Theory*, Vol. 41, pp. 1028-1039, 1995.
- [10] M. Kunde, "Block gossiping on grids and tori: deterministic sorting and routing match the bisection bound", *Proc. European Symp. on Algorithms*, pp. 260-271, 1991.
- [11] J. Li, et. al, "Capacity of Ad Hoc Wireless Networks", *Proc. ACM MobiCom*, pp. 61-69, Oct. 2001.
- [12] L-L. Xie and P. R. Kumar, "A network information theory for wireless communication: scaling laws and optimal operation", submitted to *IEEE Trans. Inform. Theory*, Apr. 2002.
- [13] A. W. Marshall and I. Olkin, "Majorization: a theory of inequalities", Academic Press, 1971.
- [14] N. Merhav, M. Feder and M. Gutman, "Universal prediction of individual sequences", *IEEE Trans. Inform. Theory*, Vol. 38, pp. 1258-1270, 1992.
- [15] A. B. Nobel, G. Morvai and S. R. Kulkarni, "Density estimation from an individual numerical sequence", *IEEE Trans. Inform. Theory*, Vol. 44(2), pp. 537-541, Mar. 1998.
- [16] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers", *Computer Communications Review*, Vol. 24(4), pp. 234-244, 1994. (ACM SIGCOMM 1994).

- [17] R. Rozovsky and P. R. Kumar, “SEEDEX: A MAC protocol for ad hoc networks”, *Proc. ACM MobiCom*, pp. 67-75, Oct. 2001.