

A Deterministic Approach to Throughput Scaling in Wireless Networks

Sanjeev R. Kulkarni, *Fellow, IEEE*, and Pramod Viswanath, *Member, IEEE*

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 $\Omega(\sqrt{n})$ throughput for a general class of network configurations in a fixed area. Results for random node locations in a fixed area 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 in a fixed area obtained by Gupta and Kumar. Results for some other interesting non-independent and identically distributed (i.i.d.) node distributions are also provided.

Index Terms—Ad hoc networks, capacity, deterministic, individual sequence, multihop, random, scaling, throughput, wireless networks.

I. 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 very large scale integration (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], [17], [18], [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 toward 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, where n is the number of nodes. 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 $\sqrt{n}/\sqrt{\log n}$, 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 transmitters are close to their corresponding receivers, then reliable communication will cause little interference to the other nodes and the scenario is essentially that of a set of noninterfering 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], [12].

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 in a fixed area.

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

Manuscript received August 13, 2002; revised October 4, 2003. This work was supported in part by the National Science Foundation under Grants ECS-9873451, CCR-0312413, and CCR-0237549, by the Army Research Office under Grant DAAD19-00-1-0466, by Draper Laboratory under IR&D 6002 Grant DL-H-546263, and through a grant from Motorola Inc. as part of the Motorola Center for Communication.

S. R. Kulkarni is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: kulkarni@princeton.edu).

P. Viswanath is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: pramodv@uiuc.edu).

Communicated by L. Tassiulas, Associate Editor for Communication Networks.

Digital Object Identifier 10.1109/TIT.2004.828055

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 independent and identically distributed (i.i.d.) uniform node locations in a fixed area 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.

II. 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 A_n 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 $\sqrt{nA_n}$ bit-meters per second and also demonstrate a specific instance of physical node locations and traffic pattern where the growth is indeed of that order.

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, Δ 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 Δ 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).

A. 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 but far from other transmitter–receiver pairs, the communication problem reverts to a set of noninterfering point-to-point communication problems with a well-understood solution. The

results in [2] can be rephrased to conclude that if every transmitter–receiver distance is of the order $O(\frac{\sqrt{A_n}}{\sqrt{n}})$, and far from other transmitter–receiver pairs, then we are in the point-to-point noninterfering 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 $\Omega(n\sqrt{A_n})$. We will show that this condition is satisfied almost surely when the transmitters and receivers are placed randomly according to various distributions. We write $O(n)$ to denote the set of functions $f(n)$ that satisfy the property $f(n) \leq c_1 n$ for a fixed constant $c_1 > 0$ and for all n large enough. Similarly, we write $\Omega(n)$ to denote the set of functions $f(n)$ that satisfy the property $f(n) \geq c_2 n$ for a fixed constant $c_2 > 0$ and for all n large enough.

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 $O(n)$ is a uniform upper bound for every configuration of the nodes. We could construct configurations with transmitters arbitrarily close to the corresponding receivers yet far from other nodes, and thereby achieve $\Omega(n)$ bits per second 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 (scaled down by $\sqrt{A_n}$) 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 $\Omega(n\sqrt{A_n})$. Here we use bit-meters per second 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 $\frac{A_n}{s_n^2}$ 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 $\frac{A_n}{s_n^2}$ 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 $\frac{c_n A_n}{s_n^2} \geq n$.

We summarize the parameters and conditions on the structure as follows.

- 1) **Tx–Rx Distances:** The sum of transmitter–receiver distances grows at least as $\Omega(n\sqrt{A_n})$.
- 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.

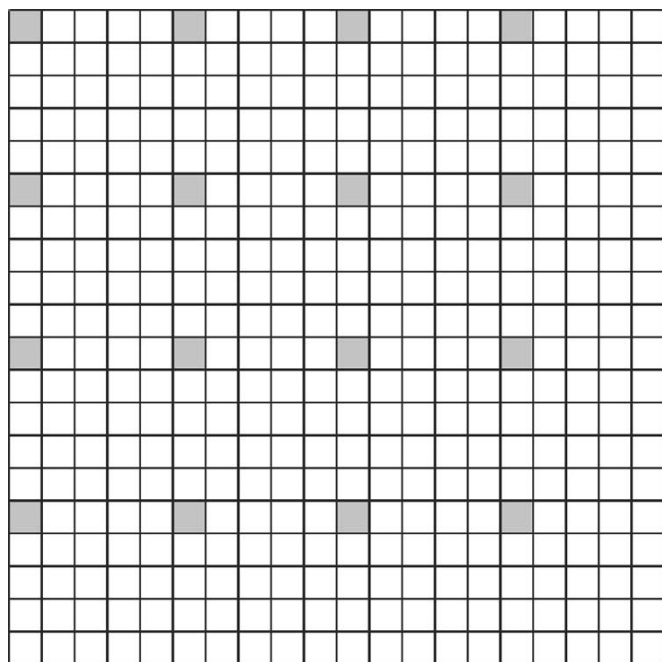


Fig. 1. An example of an equivalence class of squarelets with $K = 5$. The shaded squarelets all belong to the same equivalence class.

B. Equivalence Classes of Squarelets

To maximize throughput, we would like to schedule many simultaneous transmissions. Toward 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 Δ as discussed later) 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 Fig. 1 represent one such equivalence class.

The distance needed to allow squarelets in the same equivalence to simultaneously transmit without interference can be computed in terms of Δ 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 right of Fig. 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).

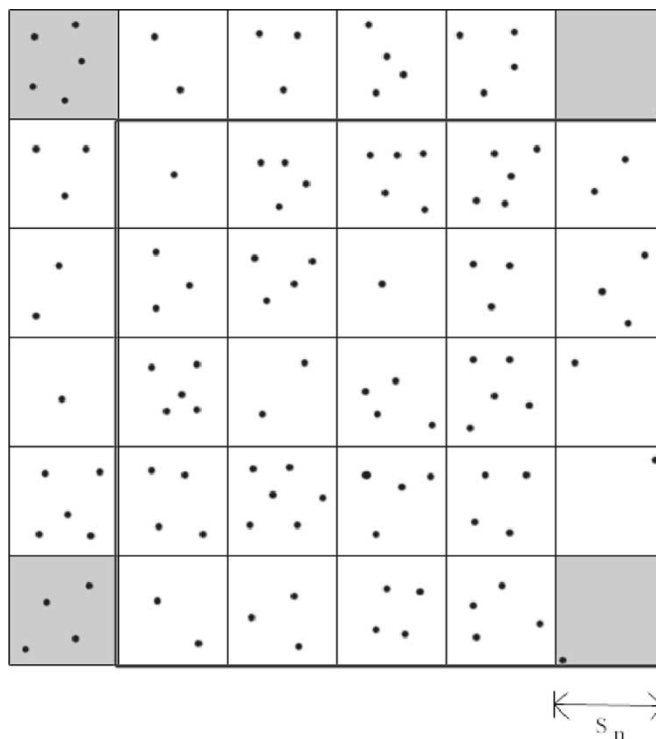


Fig. 2. A zoomed in version showing several squarelets and nodes within the squarelets. The arrangement resulting in maximum transmission distance of a single hop is shown in the bottom right. This together with the node in the top right shows the arrangement governing the bound on K .

C. Scheduling Algorithm and Throughput

As a prelude to describing the scheduling algorithm, we review a closely related problem studied extensively 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 four) simultaneously. Here we are continuing the assumptions of slotted time, communication occurring over slots, communication occurring between neighboring PUs, and each transmission comprising a *packet* of some fixed number of bits. 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 write $o(n)$ to denote the set of functions $f(n)$ that satisfy the property $f(n)/n \rightarrow 0$ as $n \rightarrow \infty$.

Lemma 2.1: [11], [9] $k \times k$ 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.

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 $l = \sqrt{A_n}/s_n$. Next, we let each user have m packets and since a squarelet can have no

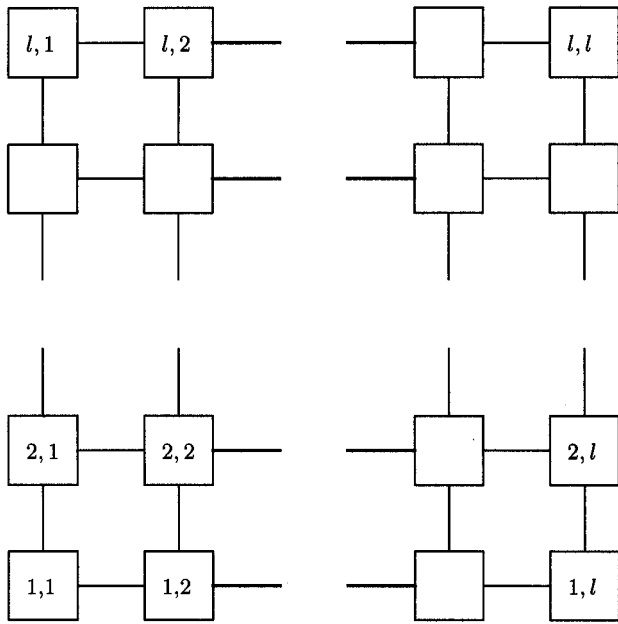


Fig. 3. A mesh network of PUs each connected to its vertical and horizontal neighbors.

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 four 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 each 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 four 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

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

Further, due to our condition on Tx–Rx distances, the total distance traveled by the $mn/2$ packets is at least order $\Omega(mn)$. 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 $\Omega(\frac{ns_n}{c_n})$.

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

III. 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. The almost sure statement is made with respect to the probability space of the infinite sequence of node locations.

Proposition 3.1: A throughput $\Omega(\sqrt{nA_n}/\sqrt{\log n})$ 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

$$s_n = \frac{\sqrt{3A_n \log n}}{\sqrt{n}}$$

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 $\Omega(n\sqrt{A_n})$ 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 $\frac{\sqrt{nA_n}}{\sqrt{\log n}}$ bit-meter per second. Since the average distance between a transmitter–receiver pair is of the order $\sqrt{A_n}$, 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 Claim 3.1.

Nonempty Squarelets In an n -node system, the probability that any fixed squarelet is empty is equal to $(1 - \frac{s_n^2}{A_n})^n$. From a simple union bound and using the fact that there are $\frac{A_n}{s_n^2}$ squarelets, we conclude that the probability that at least one squarelet is empty is upper-bounded by

$$p_n \stackrel{\text{def}}{=} \frac{A_n \left(1 - \frac{s_n^2}{A_n}\right)^n}{s_n^2}.$$

Using the fact that $1 - x \leq \exp(-x)$ and with the value of s_n equal to $\frac{\sqrt{3A_n \log n}}{\sqrt{n}}$, we have that

$$p_n < A_n \frac{\exp(-ns_n^2/A_n)}{s_n^2} \leq \frac{1}{n^2}$$

and hence, that

$$\sum_{n=1}^{\infty} p_n < \infty.$$

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 $(\frac{s_n^2}{A_n}, n)$. By a standard 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

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

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

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

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. First, let us scale down all distances by a factor of $\sqrt{A_n}$. Now each scaled distance is less than $\sqrt{2}$ and the square of the scaled distance is less than the product of $\sqrt{2}$ and the scaled distance. Further, the sum of scaled distances is lower-bounded by the sum of squared scaled distances. The squared scaled 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 i.i.d. Hence, we have, for $\theta > 0, a > 0$

$$\mathbb{P} \left[\begin{array}{l} \text{sum of distances between transmitter–receiver pairs} \\ < \frac{n\sqrt{A_n}}{a} \end{array} \right] \leq \frac{\left(\mathbb{E} \left[\exp \left(-\theta \left((x_1 - y_1)^2 + (x_2 - y_2)^2 \right) \right) \right] \right)^{\frac{n}{2}}}{\exp \left(-\frac{\theta \sqrt{2n}}{a} \right)} \quad (2)$$

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

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

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

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

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 $n\sqrt{A_n}$. This concludes the proof of the claim. \square

IV. 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 $\Omega(n\sqrt{A_n})$. In the following examples, the lack of independence only slightly complicates the analysis.

A. Users on Roads

Consider the n node locations drawn randomly as follows. In the square of area A_n , 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 do not lose much in the throughput as the following result shows.

Proposition 4.1: For nodes distributed along lines as described above, a throughput $\Omega(\sqrt{nA_n}/(\log n))$ 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

$$s_n = \frac{6 \log n \sqrt{A_n}}{\sqrt{n}}$$

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 $\Omega(n\sqrt{A_n})$ 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

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

since with probability $1 - \frac{s_n}{\sqrt{A_n}}$ 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 $1 - \frac{s_n}{\sqrt{A_n}}$ of not falling in the squarelet.

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

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

Hence, with a union bound, the probability of at least one squarelet being empty satisfies

$$\begin{aligned} \mathbb{P}[\text{at least one squarelet empty}] \\ \leq \frac{A_n}{s_n^2} \left[\left(1 - \frac{s_n}{\sqrt{A_n}}\right) + \frac{s_n}{\sqrt{A_n}} \left(1 - \frac{s_n}{\sqrt{A_n}}\right)^{\sqrt{n}} \right]^{\sqrt{n}} \\ \leq \frac{A_n}{s_n^2} e^{-\sqrt{n/A_n}(s_n - s_n(1 - \frac{s_n}{\sqrt{A_n}})^{\sqrt{n}})}, \text{ since } 1 + x \leq e^x \\ \leq \frac{A_n}{s_n^2} e^{-s_n \sqrt{n/A_n} (1 - e^{-s_n \sqrt{n/A_n}})}, \text{ using } 1 - x \leq e^{-x} \text{ again.} \end{aligned}$$

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

$$\begin{aligned} \mathbb{P}[\text{at least one squarelet empty}] &\leq \frac{n}{(36 \log n)^2} \left(\frac{1}{n^6}\right)^{(1 - \frac{1}{n^6})} \\ &\leq \frac{1}{1296 n^5 (\log n)^2}, \text{ for } n \geq 2. \end{aligned}$$

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 $(\frac{s_n}{\sqrt{A_n}}, \sqrt{n})$ 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 $(\frac{s_n}{\sqrt{A_n}}, \sqrt{n})$. 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 standard Chernoff bound we have, for all $a > 0, \theta > 0$

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

Now

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

Now, we have

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

Choosing

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

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

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

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

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

Using a simple union bound we see that

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

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 Let us first scale down all distances by a factor of $\sqrt{A_n}$. Then 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))

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

An appeal to the Borel–Cantelli lemma concludes the proof of our claim that the sum of the distances between transmitter–receiver pairs is $\Omega(n\sqrt{A_n})$ almost surely. \square

B. Users in Neighborhoods

Now consider the n node locations drawn randomly as follows. Fix a “neighborhood” size $b\sqrt{A_n}$ with $b \in (0, 0.5)$. In a $\sqrt{A_n}(1 - b) \times \sqrt{A_n}(1 - b)$ 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 $\sqrt{A_n}b \times \sqrt{A_n}b$ 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 do not lose any throughput as compared to the i.i.d. node location case, as the following result 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 $s_n < b\sqrt{A_n}$.

Proposition 4.2: For nodes distributed in neighborhoods as described above, a throughput $\Omega(\sqrt{nA_n}/\sqrt{\log n})$ 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

$$s_n = \frac{b\sqrt{3A_n \log n}}{\sqrt{n}}$$

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 $\Omega(n\sqrt{A_n})$ 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

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

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 that no node from one neighborhood falls in any fixed squarelet is upper-bounded by

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

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

$\mathbb{P}[\text{a fixed squarelet is empty}]$

$$\leq \left[\left(1 - \frac{s_n^2}{A_n(1-b)^2}\right) + \frac{s_n^2}{A_n(1-b)^2} \left(1 - \frac{s_n^2}{A_nb^2}\right)^{\sqrt{n}} \right]^{\sqrt{n}}.$$

Hence, with a union bound, the probability of at least one empty squarelet satisfies

$$\begin{aligned} \mathbb{P}[\text{at least one empty squarelet}] &\leq \frac{A_n}{s_n^2} \left[\left(1 - \frac{s_n^2}{A_n(1-b)^2}\right) + \frac{s_n^2}{A_n(1-b)^2} \left(1 - \frac{s_n^2}{A_nb^2}\right)^{\sqrt{n}} \right]^{\sqrt{n}} \\ &\leq \frac{A_n}{s_n^2} \left(1 - \frac{s_n^2}{A_nb^2}\right)^n \\ &\leq \frac{A_n}{s_n^2} e^{-\frac{ns_n^2}{A_nb^2}}. \end{aligned}$$

In the intermediate step above, we used

$$\frac{s_n^2}{A_n(1-b)^2} < 1$$

which follows from $s_n \leq A_n b < A_n(1-b)$. If we take $s_n = b\sqrt{3A_n \log n}/\sqrt{n}$, then $e^{-ns_n^2/A_n b^2} = 1/n^3$, so that

$$\mathbb{P}[\text{at least one empty squarelet}] \leq \frac{1}{3b^2 n^2 \log n}.$$

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 $\frac{s_n^2}{A_n b^2}$, 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 $\left(\frac{s_n^2}{A_n b^2}, \sqrt{n}\right)$. 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 $\left(\frac{s_n^2}{A_n b^2}, \sqrt{n}\right)$. Using a standard Chernoff bound, we now have, for all $a > 0, \theta > 0$

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

As before, we have

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

Substituting this in (10), we arrive at

$$\begin{aligned} \mathbb{P}[\text{Nodes in central squarelet} > a \log n] &\leq \frac{\left(1 + \frac{(e^\theta - 1)s_n^2}{A_n b^2}\right)^n}{\exp(\theta a \log n)} \\ &\leq \frac{\exp\left(\frac{(e^\theta - 1)ns_n^2}{A_n b^2}\right)}{\exp(\theta a \log n)} \\ &= \exp((3(e^\theta - 1) - \theta a) \log n). \end{aligned} \quad (11)$$

$$(12)$$

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

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

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 [14]. 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 $b^2 A_n$. This calculation was done in Section III for $b = 1$ but since b is fixed (with respect to the scaling in n) we conclude, using the calculation of (4), that almost surely the sum of transmitter–receiver pair distances is $\Omega(n\sqrt{A_n})$. \square

V. CONCLUSION

In this paper, we have demonstrated the power of a deterministic approach to studying the scaling of throughput with the 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 show 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], [10], [16], [15]).

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 [13], [8] show upper bounds that are tight in the scaling law sense (for the scenario when A_n is growing linearly with n).

ACKNOWLEDGMENT

This work was initiated when P. Viswanath was at Flarion Technologies and S. R. Kulkarni was visiting there. The authors wish to 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," IETF Internet Draft, 1998.
- [2] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," in *Proc. IEEE INFOCOM*, 2001.
- [3] P. Gupta and P. R. Kumar, "Capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, pp. 388–401, Mar 2000.
- [4] P. Gupta, "Design and performance analysis of wireless networks," Ph.D. dissertation, Univ. Illinois Urbana-Champaign, 2000.
- [5] P. Gupta and P. R. Kumar, "Internets in the sky: The capacity of three dimensional wireless networks," *Commun. Inform. and Syst.*, vol. 1, no. 1, pp. 33–50, Jan. 2001.
- [6] D. Haussler, J. Kivinen, and M. Warmuth, "Tight worst-case loss bounds for predicting expert advice," in *Proc. European Conf. Computational Learning Theory*, 1994.
- [7] *Wireless LAN Medium Access Control and Physical Layer Specifications*, IEEE Std. 802.11, 1997.
- [8] A. Jovicic, S. R. Kulkarni, and P. Viswanath, "Upper bounds to transport capacity of wireless networks," in *Proc. IEEE Conf. Decision and Control*, 2003.
- [9] M. Kaufmann, J. F. Sibeyn, and T. Suel, "Derandomizing routing and sorting algorithms for meshes," in *Proc. 5th Symp. Discrete Algorithms (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1995.
- [10] 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, July 1995.
- [11] M. Kunde, "Block gossiping on grids and tori: Deterministic sorting and routing match the bisection bound," in *Proc. European Symp. Algorithms*, 1991, pp. 260–271.
- [12] J. Li *et al.*, "Capacity of ad hoc wireless networks," in *Proc. ACM MobiCom*, Oct. 2001, pp. 61–69.
- [13] L.-L. Xie and P. R. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Trans. Inform. Theory*, submitted for publication.
- [14] A. W. Marshall and I. Olkin, *Majorization: A Theory of Inequalities*. New York: Academic, 1971.
- [15] N. Merhav, M. Feder, and M. Gutman, "Universal prediction of individual sequences," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1258–1270, July 1992.
- [16] A. B. Nobel, G. Morvai, and S. R. Kulkarni, "Density estimation from an individual numerical sequence," *IEEE Trans. Inform. Theory*, vol. 44, pp. 537–541, Mar. 1998.
- [17] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," *Comput. Commun. Rev.*, vol. 24, no. 4, pp. 234–244, 1994.
- [18] R. Rozovsky and P. R. Kumar, "SEEDEX: A MAC protocol for ad hoc networks," in *Proc. ACM MobiCom*, Oct. 2001, pp. 67–75.