Communications in Information and Systems
Vol. 3, No. 4, pp. 325-348, September 2004 005

A "SMALL WORLD" APPROACH TO HETEROGENEOUS NETWORKS*

ALEX REZNIK†, SANJEEV R. KULKARNI‡ , AND SERGIO VERDÚ‡

Dedicated to Sanjoy Mitter on the occasion of his 70th birthday.

Abstract. We consider a heterogeneous (also called "hybrid") ad-hoc network with wired and wireless links. This type of network was previously considered by Kulkarni and Viswanath in [9] where achievable transport capacity growth rates were demonstrated for a structured wired infrastructure. The present paper improves on this work by demonstrating that efficiency can be increased significantly if the wired links are introduced at random.

1. Introduction. In this paper we consider the effect of adding a wired infrastructure to an unstructured (ad-hoc) wireless network. The impact of such a modification on the original wireless network can be tremendous. Unlike the wireless channel, which suffers both from interference issues as well as from path loss over large distances, the wired infrastructure can provide low-cost transport of a significant amount of data over large distances without interference with other simultaneous communications. However, the existence of such an infrastructure often carries a significant cost in and of itself. Thus, in particular, one would like to use as little wired infrastructure as necessary.

A systematic study of scaling laws in ad-hoc wireless networks has been initiated in the work of Gupta, Kumar and Xie [6],[21] and [5]. In particular, Xie and Kumar showed that for a large range of path loss models the transport capacity (the distance - bandwidth product) of a purely wireless network scales no better than $\Theta(\sqrt{n})$ when the network size is fixed. Furthermore, Gupta and Kumar demonstrated how such scaling laws are achievable for all path loss models using an interference-avoiding communications protocol.

The recent work by Kulkarni and Viswanath [10] gives a simple deterministic protocol that achieves the scaling laws of Gupta and Kumar [6] eliminating much of the complexity of the original routing protocol. The protocol in [10] is based on the related problem of packet routing on a square grid of parallel processors [7],[11]. The protocol is straightforward and deterministic and thus provides a natural starting point to include a wired infrastructure into the ad-hoc wireless network. The reference

*Accepted for publication on June 29, 2004. This research was supported in part by the Army Research Office under grant number DAAD19-00-1-0466, Draper Laboratory under IR&D 6002 grant DL-H-546263, and the National Science Foundation under grant CCR-0312413. It was presented in part at the 2004 International Symposium on Information Theory [16].

†Alex Reznik is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, and Incubation Center, InterDigital Communications Corp., King of Prussia, PA 19406.

‡Sanjeev R. Kulkarni and Sergio Verdú are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544.

[9] studies the effect of this infrastructure on the resulting heterogenous network. A similar model is studied in [12].

The infrastructure considered in [9] is highly structured. It involves a square grid of wired access points overlayed on top of the wireless network. All the wired access points are interconnected and the bandwidth of the overall wired system is assumed to be large enough so that any data rate required by the wireless network can be effectively supported. The wireless users incur no penalty for using the wired infrastructure per se, however transmission (and therefore throughput) delays are incurred because contention arises when packets need to enter and exit the wired infrastructure. Using this approach and arguments from [10], it is shown in [9] that the wired infrastructure described improves the scaling law of the wireless network if there are at least $\Omega(\sqrt{n})$ access points located at most $O(\sqrt[4]{n})$ apart from each other.

We continue building on [10] and [9] by considering an ad-hoc (or more specifically random) wired infrastructure placed on top of a wireless ad-hoc network. The wireless transmission protocol we assume is that of [10] and therefore the results of [10] are directly applicable here. However, unlike [9], our wired infrastructure is comprised of point-to-point links placed in a random fashion. This particular approach is motivated by the concept of a "small world network" (SWN) which has received significant attention in recent years in the study of interactions that occur in social networks.

The rest of the paper is organized as follows. In Section 2 we provide a short overview of some concepts and literature in the "small world network" realm. In Section 3 we derive the scaling laws of an ad-hoc heterogenous network with random point-to-point wired links. As a by-product of our investigation we obtain certain results about random graphs that may be of interest beyond our specific application.

In Section 4 we compare our results with those of [9] and observe that we are able to achieve significantly better transport capacity scaling laws for the same growth rate in point-to-point wired links.

Before we proceed, it is helpful to discuss the "rate of growth" notation used in this paper. This notation is meant to be consistent with the standard computer science notation as defined in e.g. [2]. Let $f(n)$ and $g(n)$ be two positive functions. Then

- We write $f(n) = O(g(n))$ if $f(n)$ grows no faster then $g(n)$. Strictly, $f(n) = O(g(n))$ if there exist positive constants c, n_0 such that $0 \leq f(n) \leq cg(n)$ for all $n \geq n_0$.
- We write $f(n) = \Omega(g(n))$ if $f(n)$ grows at least as fast as $g(n)$. Strictly, $f(n) = \Omega(g(n))$ if there exist positive constants c, n_0 such that $0 \leq cg(n) \leq f(n)$ for all $n \geq n_0$.
- We write $f(n) = \Theta(g(n))$ if $f(n)$ grows exactly at the same asymptotic rate as $g(n)$. Strictly, $f(n) = \Theta(g(n))$ if there exist positive constants c_1, c_2, n_0 such that $0 \leq c_1 g(n) \leq f(n) \leq c_2 g(n)$ for all $n \geq n_0$. We note that $f(n) =$

$\Theta(g(n)) \Leftrightarrow f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$

- Finally, the notation $f(n) = o(g(n))$ is used if $f(n)$ grows strictly slower then $g(n)$, i.e. for any positive c there exists a positive n_0 such that $0 \leq f(n) \leq cg(n)$ for all $n \geq n_0$.

2. A Small World Network Overview. To introduce *small world networks*, we start with two concepts that are central to the study of these networks: the mean path length on a graph, and *clustering*. Roughly speaking, *clustering* describes the extent to which two neighboring (connected) nodes share other neighbors. Regular networks, such as lattices, have a high degree of clustering, however the mean path length scales linearly with the number of nodes. On the other hand, a purely random network has a low degree of clustering since the edges are assigned at random. However, it is known that the mean path length scales logarithmically with the network size, provided the edge density is sufficiently high to ensure connectivity with high probability. Thus, it is intuitive to think of mean path length and clustering as dual concepts: reducing one increases the other.

A *small world network* is a network that maintains the high clustering typical of a highly structured network while exhibiting the logarithmic mean-path weight scaling typical of a random graph. Such networks turn out to be excellent models in several real-world scenarios. Some examples from [19] include:

- Social networks of friends. Networks of friends exhibit very high clustering. We tend to be members of "circles of friends:" groups of people where everyone is friends with everyone else. We also typically have a few friends outside of such circles where none of our other friends are friends of these people. However such relationships are usually in the minority.

  On the other hand, as has been popularized by the "six degrees of separation" principle, we are but a few (colloquially 6) "hand-shakes" away from almost anyone else in the world. It turns out that the few friends we tend to have outside of tight-knit circles are the key to this property.
- Continental electric power grids which exhibit highly localized networks (around population centers) connected by a very small number of long links.
- Neural networks of worms and, by extension, possibly neural networks of more complex animals.

While it seems that "small-world" ideas have been appearing in various literature for a long time, the first clear identification of this effect is apparently due to Watts and Strogatz [20]. In particular, this work demonstrated that a SWN can be built by taking a simple highly-structured graph, such as a 1-D circular grid and introducing a small number of randomly placed shortcuts. The ideas leading to this work seem to originate from the dissertation research of Watts [18].

The initial work by Watts and Strogatz generated significant interest, particularly in the study of infectious disease spread and other percolation phenomena. Some of the references which may be of particular interest are as follows:

- [8] and [13] provide analysis which is helpful in establishing an "expected" minimum distance for SWNs.
- [15] discusses in detail the scaling of the minimum distance in SWN and in particular shows that to get SWN behavior the average number of shortcuts per node should be on the order of unity.
- [3] considers a modified concept which may be of particular interest to us because it incorporates the ideas of "broadcasting."
- A recent overview of the research may be found in [17].
- Finally, we note that while most of these papers address the original "ring" network, a two-dimensional rectangular grid is addressed in [14, 15].

In the rest of this paper we build a small-world network from a square grid. This model allows us to directly import from [10] the wireless communication protocol for ad-hoc nets, which is built up on a square-grid partition of the network. However, in the process of turning this grid into a SWN, some results are obtained which may be of general interest to the study of SWN's built on a square grid.

3. A Scaling Law for a Heterogeneous Network.

3.1. A Basic Square Grid with Shortcuts. Consider a square $K \times K$ grid, which is a graph $\mathcal{G}_K = (V_K, E_K)$. The nodes of the grid are integer pairs from the set $V_K \stackrel{\text{def}}{=} \{1, \ldots, K\} \times \{1, \ldots, K\}$. If $a \equiv (x_1, y_1)$ and $b \equiv (x_2, y_2)$ are nodes in $\mathcal{G}_K$, then the node distance between a and b in $\mathcal{G}_K$ is defined as

$$d(a, b) \stackrel{\text{def}}{=} |x_1 - x_2| + |y_1 - y_2|. \tag{1}$$

There is an edge between a and b if and only if $d(a, b) = 1$, i.e. $(a, b) \in E_K \Leftrightarrow d(a, b) = 1$.

The average number of hops between any two randomly selected nodes in $\mathcal{G}_K$ is just the average node distance. This can be found by evaluating

$$\mathsf{E}[d(A, B)] = \mathsf{E}[|X_1 - X_2|] + \mathsf{E}[|Y_1 - Y_2|] = 2\mathsf{E}[|X_1 - X_2|]. \tag{2}$$

Carrying out the computation, one finds that

$$\begin{aligned} \mathsf{E}[d(A, B)] &= \frac{K(K+2)}{2(K+1)} \\ &= \frac{1}{2}\left((K+1) - \frac{1}{K+1}\right) \approx \frac{1}{2}(K+1) \end{aligned} \tag{3}$$

where the approximation is quite accurate even for relatively low values of K, e.g. for $K = 10$ it holds to within 1%. Clearly, on a square grid, the average distance between nodes scales as $\Theta(K)$.

To turn the square grid into a small world network we introduce *shortcuts* into this otherwise very locally connected network. A shortcut is an edge (a, b) such that $d(a, b) > 1$. Our hope is that the scaling law of the average number of hops between two arbitrarily selected nodes can be reduced significantly by addition of relatively few shortcuts.

Developing an exact answer for the average number of hops in a graph with shortcuts is difficult. However, we can simplify the problem significantly by ignoring "edge effects," that is by considering only nodes sufficiently in the middle of a graph. In fact, as the network grows, most of the nodes do wind up "in the middle," provided that "the middle" is appropriately defined. Thus, given that we are interested in the scaling laws of the network as it grows, this approach yields the correct answer. An important consequence of this assumption is that the network appears to be the same from any node that we are considering and we may therefore utilize symmetry to simplify our arguments. Therefore we proceed through most of this section by ignoring edge effects. In the last section, we return to this issue and argue that indeed these do not affect the validity of the results that we obtained.

Additionally, we restrict ourselves to using at most one shortcut for any source-destination pair and we use a shortcut if and only if this reduces the number of hops needed to travel between the two selected nodes. Otherwise, we assume that the regular edges on the square grid are used and the cost of travel is equal to the distance between the nodes.

One can take many approaches to populate the graph with shortcuts. Our approach differs from what is generally done in the SWN literature, but we believe that it lends itself better to the type of analysis required for our problem. As an initial step, let Φ be the set of all possible shortcuts. Let each shortcut in Φ be introduced to the graph equiprobably and independently with probability ϕ.

Fix D, and fix a pair of nodes a and b "sufficiently inside the grid" such that $d(a, b) = D$. Let us find $l(a, b)$, the expected number of hops needed to get from a to b.

There is only one way that $l(a, b) = 1$ and that is if there is a shortcut between a and b. By definition this happens with probability ϕ. If $l(a, b) = 2$ then there must be a shortcut between a and one of the 4 nodes adjacent to b, or between b and one of the 4 nodes adjacent to a. The probability that none of the 8 possible shortcuts exist is $(1 - \phi)^8$ and thus the probability that at least 1 of them is present is $1 - (1 - \phi)^8$. Clearly, to find the expected value of $l(a, b)$ we need to carry this process out through D. To do so, we make the following definitions and observations.

- Given a node a sufficiently inside the grid, the number of nodes b such that $d(a, b) = n$ is $4n$ if $n > 0$ and 1 if $n = 0$.
- The number of possible shortcuts between a set of n nodes and a set of m nodes is given by nm.

- Given that we want $l(a,b) = L$ let $T(L)$ be the number of possible shortcuts that would supports this. Then,

$$T(L) = \begin{cases} 1 & \text{if } L = 1 \\ 8(L-1) + 16\sum_{i=1}^{L-2} i(L-1-i) & \text{if } L > 1. \end{cases} \tag{4}$$

  We also let $T(0) \equiv 0$ for convenience. We note that for $L \geq 2$, $T(L)$ is specified by the polynomial

$$T(x) = \frac{8}{3}(x^3 - 3x^2 + 5x - 3) = \frac{8}{3}(x-1)(x^2 - 2x + 3). \tag{5}$$

  We will find it useful to utilize (5) as an approximation to $T(L)$ for all values of L. We note that $T(L)$ is a strictly increasing function. This is easily seen by noting that the derivative of $T(x)$ is

$$T'(x) = 8x^2 - 16x + \frac{40}{3} > 0 \ \ \forall x. \tag{6}$$

With these observations, we can now compute the expected value of $l(a,b)$. Since $l(a,b)$ takes values on positive integers, we have

$$\mathsf{E}[l(a,b)] = \sum_{L=1}^{\infty} \mathsf{P}(l(a,b) \geq L) = 1 + \sum_{L=2}^{D} \mathsf{P}(l(a,b) \geq L). \tag{7}$$

The event $(l(a,b) \geq L)$ means that no shortcuts that support $l(a,b) < L$ are available, therefore

$$\mathsf{P}(l(a,b) \geq L) = \prod_{m=1}^{L-1} (1-\phi)^{T(m)} = (1-\phi)^{\sum_{m=1}^{L-1} T(m)}. \tag{8}$$

Thus,

$$\mathsf{E}[l(a,b)] = 1 + \sum_{L=2}^{D} (1-\phi)^{\sum_{m=1}^{L-1} T(m)} = \sum_{L=0}^{D-1} (1-\phi)^{\sum_{m=0}^{L} T(m)} \tag{9}$$

where we changed the summation limits by substituting $(L-1) \Rightarrow L$ and noting that the term inside the sum evaluates to 1 if $L = 0$. Let

$$\begin{aligned} S(L) &\stackrel{\text{def}}{=} \sum_{m=0}^{L} T(m) = 1 + \sum_{m=2}^{L} T(m) \\ &= \frac{1}{3}\left(2L^4 - 4L^3 + 10L^2 - 8L + 3\right), \quad L \geq 1 \end{aligned} \tag{10}$$

with $S(L) \equiv 0$ for $L \leq 0$. Additionally, as we did with $T(x)$, we will find it useful to extend $S(L)$ to the reals by defining

$$S(x) \stackrel{\text{def}}{=} \frac{1}{3}\left(2x^4 - 4x^3 + 10x^2 - 8x + 3\right), \quad x > 0 \tag{11}$$

with $S(x) \equiv 0$ for $x \leq 0$. We note that in this case the extension is exact for all positive integers.

Clearly, as long as both a and b are sufficiently inside the grid, the expected value of $l(a,b)$ is a function of only ϕ and D. That is, we may define

$$\bar{l}(\phi, D) \stackrel{\text{def}}{=} \mathsf{E}l(a,b) \text{ for } a, b \text{ such that } d(a,b) = D. \tag{12}$$

Moreover, as D increases, $\bar{l}(\phi, D)$ converges to a finite value for all $\phi \in (0,1)$ and therefore we may define

$$\bar{l}(\phi) \stackrel{\text{def}}{=} \lim_{D\to\infty} \bar{l}(\phi, D) = \sum_{L=0}^{\infty} (1-\phi)^{S(L)}. \tag{13}$$

The convergence of $\bar{l}(\phi, D)$ is shown rigourously in Appendix A.

Let us demonstrate some of the properties of $\bar{l}(\phi, D)$. If we fix ϕ, then $\bar{l}(\phi, D)$ rapidly reaches its asymptotic value as D grows. This is illustrated by the two plots in Figure 1. Thus, for sufficiently large values of ϕ we can plot $\lim_{D\to\infty} \bar{l}(\phi, D)$ by plotting $\bar{l}(\phi, D)$ for $D = 100$. This is shown in Figure 2.

3.2. Extending the basic model. A significant problem with the above formulation is that while the expected number of hops between two nodes is upper bounded by a constant, the expected number of shortcuts originating at each node grows in proportion to the number of nodes and the total number of shortcuts grows in proportion to the square of the number of nodes. Indeed, the expected number of shortcuts of radius no more than R originating at any given node is given by

$$\sum_{k=2}^{R} 4k\phi = \Theta(R^2) = \Theta(n) \tag{14}$$

where n is the total number of nodes at most R away from a given node.

In order to study the scaling laws of our heterogeneous network, we would like to introduce some control over the rate of growth of the number of shortcuts with the size of the network. One method for doing so is to weigh shortcuts of different length differently. In particular, let us assume that the probability that a shortcut of length k is present in the network is given by $\frac{\phi}{k^p}$ where $0 < \phi < 1$ and $p > 0$ are parameters of the network. This modification provides us with significant control. Indeed, the expected number of shortcuts of radius no more than R originating at any node is now given by

$$\sum_{k=2}^{R} 4k\frac{\phi}{k^p} = \begin{cases} O(1) & \text{if } p > 2 \\ \Theta(\log R) = \Theta(\log n) & \text{if } p = 2 \\ \Theta(R^{2-p}) = \Theta(n^{1-\frac{p}{2}}) & \text{if } p < 2. \end{cases} \tag{15}$$

To understand how this modification affects the expected number of hops between two randomly selected nodes, we re-trace the steps we used to get to (9) and study

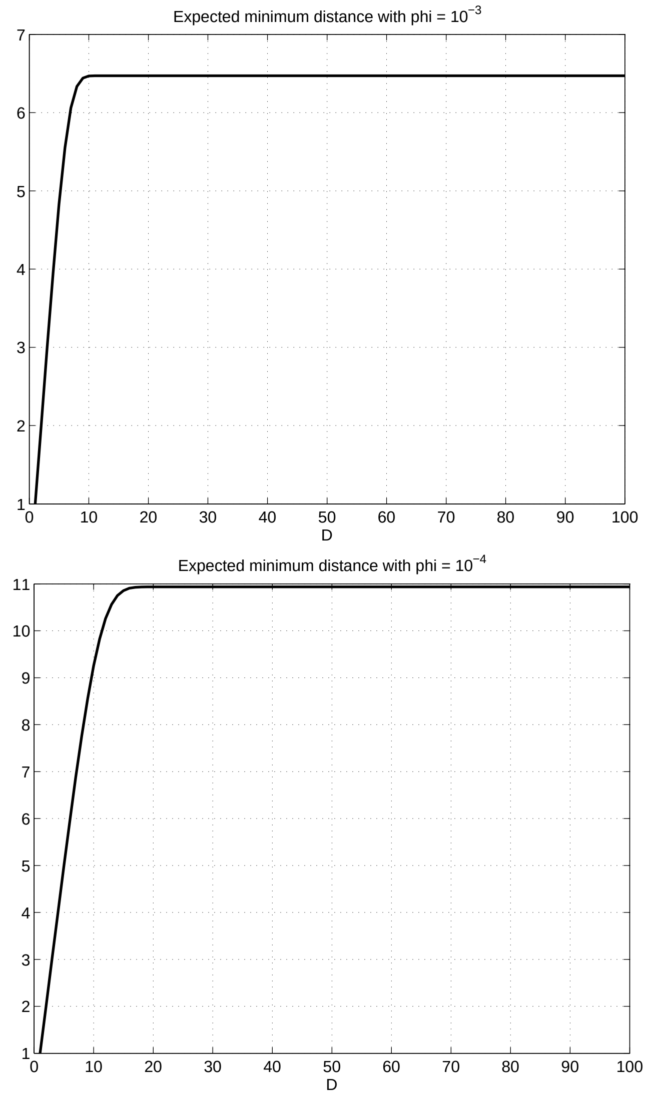

FIG. 1. $\bar{l}(\phi, D)$ **for** $\phi = 10^{-3}$ **and** $\phi = 10^{-4}$

$\mathsf{E}l(a, b)$ as a function of ϕ and D. We begin by randomly selecting two nodes (a and b) at a distance D away from each other and considering a shortcut that reduces the number of hops between a and b to L. Let us bound the length of the shortcut. If the shortcut connects two nodes in a shortest path from a to b, then it's length is $(D - L)$. This is the lower bound on the length and we can expect that usually a shortcut would have to span a longer distance. In this case we have

$$\mathsf{P}(l(a, b) \geq L) \geq \prod_{m=1}^{L-1} \left(1 - \frac{\phi}{(D - m)^p}\right)^{T(m)}. \tag{16}$$

The upper bound on the graph distance spanned by a shortcut is $2D$. If a shortcut

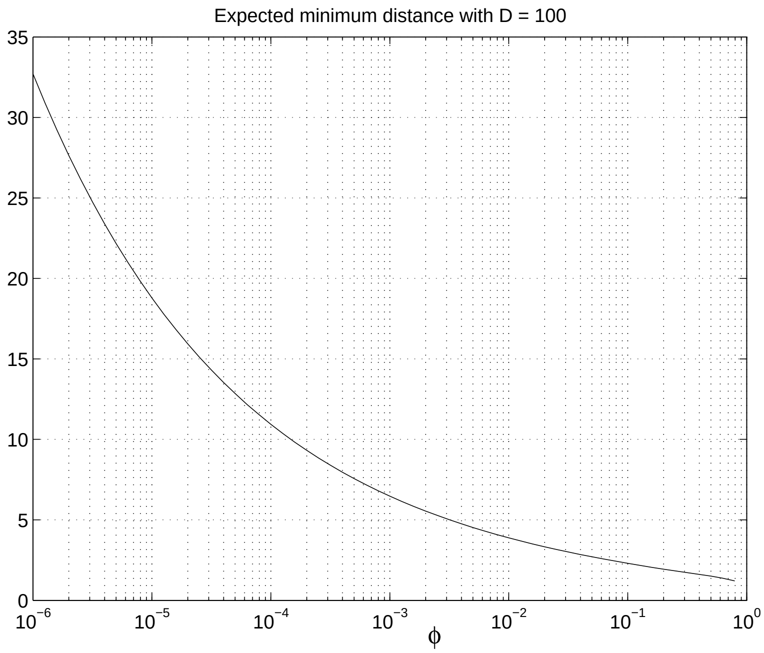

FIG. 2. $\bar{l}(\phi, D)$ **for** $D = 100$

spans such a distance, the total graph distance from the source to the shortcut entry point and from shortcut to destination must be at least D. Since this is the graph distance between the nodes, any shortcuts that are longer than this would not be used. In this case we have

$$\mathsf{P}(l(a,b) \geq L) \leq \prod_{m=0}^{L-1} \left(1 - \frac{\phi}{(2D)^p}\right)^{T(m)} = \left(1 - \frac{\phi}{2^p D^p}\right)^{S(L-1)}. \tag{17}$$

In order to keep the dependence on ϕ and D explicit denote

$$\bar{l}_p(\phi, D) \stackrel{\text{def}}{=} \mathsf{E}[l(a,b)]. \tag{18}$$

Then, using the property (7) and retracing the steps used to get to (9) we get

$$\begin{aligned} \bar{l}_p(\phi, D) &\geq 1 + \sum_{L=2}^{D} \prod_{m=1}^{L-1} \left(1 - \frac{\phi}{(D-m)^p}\right)^{T(m)} \\ &= \sum_{L=0}^{D-1} \prod_{m=0}^{L} \left(1 - \frac{\phi}{(D-m)^p}\right)^{T(m)} \end{aligned} \tag{19}$$

and

$$\begin{aligned} \bar{l}_p(\phi, D) &\leq \sum_{L=0}^{D-1} \prod_{m=0}^{L} \left(1 - \frac{\phi}{(2D)^p}\right)^{T(m)} \\ &= \sum_{L=0}^{D-1} \left(1 - \frac{\phi}{2^p D^p}\right)^{S(L)}. \end{aligned} \tag{20}$$

On one hand, the problem has now become more complicated, since, as we show in Appendix A, if $p > 0$, $\bar{l}_p(\phi, D) \to \infty$ as $D \to \infty$. However, as we demonstrate next, this modification gives us a fairly precise control over the scaling law of the number of hops in the network.

THEOREM 1. *Fix $p \geq 0$. Then for any $0 < \phi < 1$,*

$$\bar{l}_p(\phi, D) = \begin{cases} O\left(D^{p/4}\right) & \text{if } p < 4 \\ O(D) & \text{if } p \geq 4. \end{cases} \tag{21}$$

Proof. The second line of (21) is trivial. We simply note that for all p

$$\bar{l}_p(\phi, D) \leq \sum_{L=0}^{D-1} 1 = D. \tag{22}$$

Next, we start with (20) and write

$$\bar{l}_p(\phi, D) \leq \sum_{L=0}^{D-1} \exp\left(S(L) \log\left(1 - \frac{\phi}{2^p D^p}\right)\right) \tag{23}$$

$$\leq \sum_{L=0}^{D-1} \exp\left(-\frac{\phi S(L)}{2^p D^p}\right). \tag{24}$$

Now, $S(L)$ is a 4^{th} order polynomial with the leading coefficient equal to $\frac{2}{3}$. We may therefore write for any $\epsilon > 0$,

$$\bar{l}_p(\phi, D) \leq C + \sum_{L=0}^{D-1} \exp\left(-\frac{\phi}{2^p}\left(\frac{2}{3} - \epsilon\right)\frac{L^4}{D^p}\right), \tag{25}$$

where C is a finite constant that absorbs the difference from some initial and finite number of terms where $S(L) < \left(\frac{2}{3} - \epsilon\right) L^4$. Specifically, let $\tilde{L}$ be such that for all $L > \tilde{L}$, $\left(\frac{2}{3} - \epsilon\right) L^4 < S(L)$. Then $\tilde{L}$ is necessarily finite and

$$C \stackrel{\text{def}}{=} 1 + \sum_{L=0}^{\tilde{L}} \exp\left(-\frac{\phi S(L)}{2^p D^p}\right) - \sum_{L=0}^{\tilde{L}} \exp\left(-\frac{\phi}{2^p}\left(\frac{2}{3} - \epsilon\right)\frac{L^4}{D^p}\right). \tag{26}$$

Finally, for $p < 4$ we have from Lemma A.4

$$\sum_{L=0}^{D-1} \exp\left(-\frac{\phi}{2^p}\left(\frac{2}{3} - \epsilon\right)\frac{L^4}{D^p}\right) = \Theta\left(D^{p/4}\right), \tag{27}$$

where the lemma can be applied by taking

$$\alpha \equiv \exp\left(-\frac{\phi}{2^p}\left(\frac{2}{3} - \epsilon\right)\right). \tag{28}$$

This completes the proof of the theorem. □

A critical consequence of Theorem 1 is that not only does the expected number of hops scale as $D^{\min(\frac{p}{4}, 1)}$, but as D grows large so does the maximal number of hops that we expect to observe. This statement is a direct consequence of Markov's Inequality [1] which states that for any non-negative r.v. X and any $\lambda > 0$

$$\mathsf{P}\left(X \geq \lambda \mathsf{E} X\right) \leq \frac{1}{\lambda}. \tag{29}$$

Now let $\omega(\cdot)$ denote any function such that $\omega(\cdot) \to +\infty$. We note that the rate of growth of ω may be extremely slow, for example we may take $\omega(x) = \log(1 + \log(1 + \cdots \log(1 + x)))$. Then, by Markov's Inequality we immediately have that as $D \to \infty$

$$l(a,b) < \omega(D) D^{\min(\frac{p}{4},1)} \text{ in probability} \tag{30}$$

or, alternatively,

$$\mathsf{P}\left[l(a,b) \geq \omega(D) D^{\min(\frac{p}{4},1)} \text{ i.o.}\right] = 0. \tag{31}$$

The probability in (30) and (31) arises as we consider the random placement of shortcuts according to our imposed probability law as parameterized by p and ϕ. In other words the above statement means the following:

> The probability that the placement of shortcuts is such that for any choice of nodes a and b such that the grid distance between them is D and some $\omega(\cdot) \to \infty$ $l(a,b)$ exceeds $\omega(D) D^{\min(\frac{p}{4},1)}$ *infinitely often* is 0.

We note here that since we have not placed any restrictions on how slowly $\omega(\cdot)$ may grow towards infinity, we will at times simply state that $\min(\frac{p}{4},1)$ is the asymptotic rate of growth of the $l(a,b)$ and it is always to be understood that the actual growth rate is "just a little faster" than that.

The convergence in probability demonstrated in (30) may not always be sufficient and a stronger probabilistic statement may be desired. This can be achieved by placing additional restriction on $\omega(\cdot)$, as demonstrated in the following lemma.

LEMMA 1. *For any choice of nodes a and b with grid distance D*

$$\forall \epsilon > 0 \quad P\left[\limsup_{D\to\infty} \frac{l(a,b)}{\sqrt[4]{\log(D)} D^{\min(\frac{p}{4},1)}} > \epsilon\right] \to 0, \tag{32}$$

which may be alternatively stated as

$$l(a,b) < C_{as} \sqrt[4]{\log(D)} D^{\min(\frac{p}{4},1)} \textit{ almost surely }, \tag{33}$$

where C_{as} is an appropriately chosen constant that does not depend on D.

Proof. The necessary result can be demonstrated using (30) by additionally showing that the Borel-Cantelli Lemma [1] holds. To show the latter, we need to demonstrate that the series

$$\sum_D \mathsf{P}\left(l(a,b) \geq \omega(D) D^{\frac{p}{4}}\right) \tag{34}$$

converges, where $\omega(D)$ is an increasing function of D.

Examining (34) we can use (17) to write

$$\sum_D \mathsf{P}\left(l(a,b) \geq \omega(D)D^{\frac{p}{4}}\right) \leq \sum_D \left(1 - \frac{\phi}{2^p D^p}\right)^{S\left(\omega(D)D^{\frac{p}{4}}\right)} \tag{35}$$

$$= \sum_D \left(1 - \frac{\phi}{2^p D^p}\right)^{\tilde{\omega}(D)\tilde{S}(D)}, \tag{36}$$

where we recall that $S(\cdot)$ is a fourth order polynomial and therefore in decomposing $S\left(\omega(D)D^{\frac{p}{4}}\right)$ into $\tilde{\omega}$ and $\tilde{S}$ we may write $\tilde{S}(D) = \Theta(D^p)$ and $\tilde{\omega}(D) = \Theta\left((\omega(D))^4\right)$. We then continue

$$\sum_D \mathsf{P}\left(l(a,b) \geq \omega(D)D^{\frac{p}{4}}\right) \leq \sum_D \exp\left(\tilde{\omega}(D)\tilde{S}(D)\log\left(1 - \frac{\phi}{2^p D^p}\right)\right) \tag{37}$$

$$\leq \sum_D \exp\left(-\frac{\phi\tilde{\omega}(D)\tilde{S}(D)}{2^p D^p}\right). \tag{38}$$

We now note that since $\tilde{S}(D) = \Theta(D^p)$, $\frac{\phi\tilde{S}(D)}{2^p D^p} = \Theta(1)$. Let C_{conv} be a constant chosen so that $\frac{\phi\tilde{S}(D)}{2^p D^p} = C_{conv} + o(\tilde{\omega}(1))$. Then, if we take $\tilde{\omega}(D) \stackrel{\text{def}}{=} \frac{2}{C_{conv}}\log(D)$ we have

$$\sum_D \mathsf{P}\left(l(a,b) \geq \omega(D)D^{\frac{p}{4}}\right) \leq \sum_D \exp\left[-2\log(D)\right] \tag{39}$$

$$= \sum_D D^{-2} < \infty. \tag{40}$$

We conclude that (34) converges if $\tilde{\omega}(D) = \Theta(\log(D))$ provided that the constant factor is large enough. Finally we have $\tilde{\omega}(D) = \Theta\left((\omega(D))^4\right) \Rightarrow \omega(D) = \Theta\left((\tilde{\omega}(D))^{\frac{1}{4}}\right)$ which proves the lemma. The constant factor C_{as} depends on C_{conv} and the polynomial $S(\cdot)$ in a highly non-trivial way and we do not derive it here explicitly. For our purposes, it is sufficient to note that C_{as} is finite and positive, and although it depends on p and ϕ, it is independent of D. □

3.3. A Scaling Law on a Square Grid with Shortcuts. We are now ready to proceed with the derivation of the scaling law for a particular heterogeneous network with a wired infrastructure. We make the following assumptions on our network.

- The wireless protocol operates according to the "protocol" interference model of [6]. This is the same assumption as is made in [10].
- A network of n nodes is divided into $\frac{n}{2}$ point-to-point source-destination pairs. The selection of node-destination pairs is arbitrary. This is done prior to the beginning of any transmission and remains fixed. When the selection of source-destination pairs is arbitrary (as opposed to random) we assume, as does [10], that sum-distance between all source-destination pairs scales as

$\Theta(n)$. When the selection is random, we use the conditions on the growth of c_n (described below) under which [10] demonstrates that the sum-distance between all source-destination pairs grows as $\Theta(n)$.

- The wired infrastructure is setup prior to the beginning of transmission and remains fixed while the network is active.
- Prior to the beginning of any transmission, each source-destination pair selects the best route utilizing at most a single wired link.

To set up our communication protocol, we proceed as in [10]. We consider a square of area 1 with n nodes in it. Each node has at most m packets to transmit. We divide the area into small squares called *squarelets* having side length (size) s_n and let c_n denote the maximum number of nodes in each squarelet, which in [10] is called the *crowding factor*. We further impose the restriction that each squarelet must contain at least one node. For a purely wireless network, [10] shows the following result

THEOREM 2. *The throughput capacity in bit-meters per second (and bits per second) for a purely wireless network with squarelet size s_n and crowding factor c_n is $\Omega\left(\frac{ns_n}{c_n}\right)$. Moreover, the maximum total number of packets at each squarelet is given by mc_n.*

It will be useful for us to outline the proof of this statement since we use it to prove our own results. The proof relies on the following result from the parallel processing community. Consider a square grid of $l \times l$ processors each with k packets to transmit to one of the other processors. The problem is referred to as $k \times k$ *permutation routing*. The following result is shown in [7], [11].

LEMMA 2. *$k \times k$ permutation routing in a $l \times l$ mesh can be performed deterministically in $\frac{kl}{2} + o(kl)$ steps with maximum queue size at each processor equal to k. Further, every routing algorithm takes at least $\frac{kl}{2}$ steps.* □

We now return to the proof of Theorem 2 as provided in [10]

Proof. Note the following equivalences between our wireless network and the $k \times k$ permutation routing problem:

- Each squarelet is equivalent to a processor: $l \equiv \frac{1}{s_n}$
- The total number of packets that each squarelet needs to transmit is upper bounded by mc_n, thus $k \equiv mc_n$.
- The squarelets are divided into K^2 equivalence classes, where K is a constant that is determined by the parameters of the protocol model. See [10] for details about how it is obtained. It is assumed that the nodes within a squarelet time-slot their transmissions.

Under these assumptions, we can use Lemma 2 and argue (see [10] for details) that all the packets can get to their destination in a number of time slots which is at most

$$O\left(4K^2\frac{k}{2l}\right) = O\left(\frac{2K^2mc_n}{s_n}\right). \tag{41}$$

Due to the condition we imposed on the source-destination pair distance, the total distance travelled by the packets is $\Theta(mn)$ giving us the desired transport capacity result.

The maximal size of the queue per squarelet follows directly from the equivalence of k in the $k \times k$ permutation routing problem and mc_n in our problem. □

In particular, with the best possible node configuration, we have $s_n = \frac{1}{\sqrt{n}}$ and $c_n = O(1)$, in which case $\Omega(\sqrt{n})$ transport capacity is attained. Furthermore, the following is proven in [10].

LEMMA 3. *Let the nodes of the network be i.i.d. distributed over the unit square and the source-destination pairs be selected at random. Then, as $n \to \infty$ the following hold.*

- *With a squarelet size of $s_n = \frac{\sqrt{3\log n}}{\sqrt{n}}$ no squarelet is empty almost surely.*
- *$c_n \leq 3e\log n$ almost surely.*
- *The sum of the distances between sources and their respective destinations grows as $\Theta(n)$ almost surely.*

Thus, [10] concludes that the transport capacity scales as $\Omega\left(\frac{\sqrt{n}}{\sqrt{\log n}}\right)$ almost surely.
□

Let us now take the network above and add wired shortcuts. This is done as follows. Consider a square grid of squarelets, i.e. let each squarelet be a node on the square grid. The wired link representing the shortcut may be physically connected to any of the nodes in a squarelet. It is added according to the probabilistic model that we studied. The shortcuts are added before any transmission starts and remain fixed while the network is active. The source-destination pairs always choose the route that involves the lowest number of wireless hops.

Because the assumption that the grid appears identical to all nodes is essential to our results on the square grid with shortcuts, it is necessary to impose a "symmetry" assumption on the distribution of nodes among the squarelets of our wireless network. We make this assumption precise:

> Consider a collection of N random variables indexed according to the index set $\mathcal{I} \stackrel{\text{def}}{=} \{1, \ldots, N\}$. If $\mathcal{J} \subset \mathcal{I}$ then let $P_{\mathcal{J}}$ denote the marginal joint distribution of the random variables whose indices are in $\mathcal{J}$. Then the random variables are said to be "symmetrically distributed" if for any $\mathcal{I}' \subset \mathcal{I}$ and $\mathcal{I}'' \subset \mathcal{I}$ such that $|\mathcal{I}'| = |\mathcal{I}''|$ we have $P_{\mathcal{I}'} = P_{\mathcal{I}''}$ and moreover $P_{\mathcal{I}'}$ is a symmetric function of its arguments - i.e. it is invariant to permutations of the arguments.

With this definition, we can make the following statement about the number of shortcuts originating or terminating at a node that actually wind up being used. For the remainder of this section let v_i denote the number of utilized shortcuts that either originate or terminate in a squarelet i. Then we have the following

THEOREM 3. *For any squarelet sufficiently inside the grid*

$$v_i = O(\omega(n) n s_n^2) \text{ in probability} \tag{42}$$

where $\omega(\cdot)$ is any positive function such that $\omega(\cdot) \to \infty$. The probability is with respect to the random placement of shortcuts onto the grid.

Proof. Recall that we have exactly n source-destination pairs and that each source-destination pair may use no more then one shortcut to communicate. Therefore, the total number of used shortcuts is upper bounded by n. Let $\mathcal{S}$ be the set of squarelets. Then we have

$$\sum_{i \in \mathcal{S}} v_i \leq n \Rightarrow \sum_{i \in \mathcal{S}} \mathsf{E} v_i \leq n. \tag{43}$$

Since we are interested only in nodes that are "sufficiently inside the grid," it follows by symmetry that the marginal probability law of each v_i should be the same, and in particular,

$$\mathsf{E} v_i = \mathsf{E} v_j \quad \forall i, j \in \mathcal{S}. \tag{44}$$

Given that the total number of squarelets is s_n^{-2}, we conclude that

$$\mathsf{E} v_i \leq n s_n^2. \tag{45}$$

Finally, since $v_i \geq 0$, we apply Markov's Inequality to the above and conclude

$$v_i = O(\omega(n) n s_n^2) \text{ in probability} \tag{46}$$

for any node sufficiently inside the grid. This proves the first assertion of the theorem.
□

Using the results above we can now make a statement about the achievable transport capacity in our network. Because Theorem 3 provides results for asymptotic growth *in probability* only, we only make the following statement *in probability* as well.

THEOREM 4. *Consider a network where the distribution of nodes between squarelets is symmetric. Then* in probability

i) *The throughput in bit-meters per second (and in bits per second) in the resulting heterogeneous network grows as*

$$\Omega\left(\frac{\min\left[n s_n^{\min(p/4,1)}, s_n^{-2}\right]}{\omega(n) c_n}\right). \tag{47}$$

ii) *The required capacity of any wired link is upper bounded by*

$$O(\omega(n) n s_n^2 m c_n). \tag{48}$$

Here, as before, $\omega(\cdot)$ is any function that grows to $+\infty$.[1] *The probability is with respect to the random placement of shortcuts onto the grid.*

Proof. Clearly, the total distance travelled by all the packets is still $\Theta(mn)$. Thus, we need to understand how the time that it takes a packet to get from source to destination scales. Each packet can travel in one of two ways: either purely through the wireless channel or using one, but only one, of the shortcuts.

Let the maximal number of hops between any two nodes on this grid grow as M_n. If the packet travels through the wireless channel only, the total number of hops taken by any packet is upper bounded by M_n by assumption. Thus, it takes at most $O(M_n m c_n)$ slots to get to the destination.

If the packet uses a shortcut, it takes at most M_n hops to get to the entry point for the shortcut and at most M_n hops to get from the exit point to its destination. As in [9], we need to be concerned about the queueing at the exit point of a shortcut. To upper bound the amount of queuing, we need to upper bound the number of used shortcuts that originate or terminate at each node. We note up front, and this will become evident shortly, that most shortcuts are not actually used.

The deduction above immediately leads to the following conclusions. *In probability*:

- The delay sustained in getting a packet across grows as the faster growing of the two contributions: the delay required to get the packet across the wireless network, which is $O(M_n m c_n)$; and the delay sustained in accessing/exiting the wireline network, which is $O(\omega(n) n s_n^2 m c_n)$. The transport capacity therefore grows as

$$\Omega\left(\frac{n}{\max\left[M_n c_n,\ \omega(n) n s_n^2 c_n\right]}\right). \tag{49}$$

- The number of packets carried on any wired link is $O(\omega(n) n s_n^2 m c_n)$, which proves the second statement of the theorem.

To complete the proof, we need to express M_n in terms of s_n. To do so, recall that we have n (or $\Theta(n)$) source-destination pairs whose sum-distance is $\Theta(n)$. Recall that none of the source-destination pair distances may exceed 1: 1 is the length of the side of our square network. It follows that there is a non-vanishing portion of the source-destination pair with distances $\Theta(1)$. If the latter statement were not true, we could not have the sum of source-destination pair distances be $\Theta(n)$.

[1] We note that under the assumption of a symmetric distribution of nodes over squarelets, it may be possible to express the rate of growth of c_n in terms of s_n. We leave our results in terms of c_n for consistency with [9].

Take a pair of nodes such that the number of squarelet-hops required to get from one to another is D. Our results in Section 3.2 indicate that *in probability* the presence of shortcuts between squarelet reduces the number of hops required to $O\left(\omega(n) D^{\min(p/4,1)}\right)$.

We note that $\frac{1}{s_n}$ is the number of squarelets along the side of our unit-square. Moreover, recall that there is a non-vanishing proportion of source-destination pairs with distances $\Theta(1)$. Substituting the number of squarelets for the distance units we get that there is a non-vanishing proportion of source-destination pairs with distances $\Theta\left(\frac{1}{s_n}\right)$. Thus, we may substitute as $D = \Theta\left(\frac{1}{s_n}\right)$ and write the maximal number of hops as

$$M_n = O\left(\frac{\omega(n)}{s_n^{\min(p/4,1)}}\right). \tag{50}$$

We can substitute for M_n into (49) to obtain a transport capacity of

$$\Omega\left(\frac{\min\left[n s_n^{\min(p/4,1)}, s_n^{-2}\right]}{\omega(n) c_n}\right) \tag{51}$$

and the theorem follows. □

We call attention to a fact that is not readily apparent from the statement and proof of the theorem, but should become apparent when it is applied to some specific situations below. In the majority of cases the transport capacity in our heterogeneous network as given by (47) is determined by the time it takes for packets to traverse the wireless network and not the access/exit queue wait times in the wired infrastructure. This is in fact the primary cause of our ability to improve upon the results in [9] where the wait times to enter/exit the network limited the achievable transport capacity.

Using the results of [10] as cited above, and noting that both the "best" and the uniformly random allocation of nodes satisfy our symmetry assumption, we have the following immediate corollary

COROLLARY 1. *Under the best node allocation, transport capacity of*

$$\Omega\left(\frac{n^{\max\left(1-\frac{p}{8},\frac{1}{2}\right)}}{\omega(n)}\right) \tag{52}$$

is attainable in probability.

Using the random uniform allocation of nodes, transport capacity of

$$\Omega\left(\frac{n^{\max\left(1-\frac{p}{8},\frac{1}{2}\right)}}{\omega(n) \log n^{\min\left(1-\frac{p}{8},\frac{1}{2}\right)}}\right) \tag{53}$$

is attainable in probability.

Proof. The results above are readily observed by substituting for s_n and c_n from [10] into (47). We note that in both cases the resulting transport capacity growth rate is determined by the wireless transmission time and not by the access/exit waits for the wired infrastructure. □

To conclude this discussion, we return now to the problem of obtaining the results in the stronger, *almost sure* sense. Although we were not able to show non-trivial *almost sure* convergence for Theorem 3, we conjecture that it holds if $\omega(\cdot)$ in the statement of the theorem is replaced by $[\log(\cdot)]^r$ for some r, $0 < r < \infty$. This was certainly the case for the number of hops, as shown in Lemma 1. If this is so, then the results of Theorem 4 and Corollary 1 also follow with $\omega(\cdot)$ replaced by a positive power of log.

3.4. Edge effects. We now return to the issue of edge effects. There are two places in the development above where edge effects need to be addressed. The first is in counting the number of shortcuts that may be used from node a to node b to attain a particular number of hops (equation (4)). Let both the source and the destination nodes be the corner nodes on the square grid. Certainly, this is the worst case scenario in terms of breaking down the "middle of the grid" assumption. In this case the number of shortcuts available is reduced by approximately a factor of 16 (only a quarter of the neighborhood is available for each node). However, the rate of polynomial growth of available shortcuts is not affected. This still grows as the cube of the required number of hops. Since of all of our results depend only of the rate of polynomial growth of $T(L)$, these are not effected.

The second place where the "middle of the graph" assumption is used is in the proof of Theorem 4. Specifically, the symmetry argument used in (44) relies on this assumption. We show that the results obtained under this assumption are not affected by edge effects so long as $p < 4$.

To do so, consider adding a strip of squarelets of width $O(s_n^{p/4})$ (recall that s_n is the size of the squarelet which decreases with n) around the edges of the original network (of width 1). These squarelets are used as fake wireline access points and nodes within a strip of edge squarelets of width $O(s_n^{p/4})$ inside the original grid may select them.

Consider now squarelets that are located at least $O(s_n^{p/4})$ *inside* the original grid. To wired access points inside these squarelets, the network appears completely symmetric since

- They each have a neighborhood of possible sources/destination of radius $O(s_n^{p/4})$.
- Each of the possible sources/destination in the neighborhood of such a wired access point is equally likely to use it since it has a neighborhood of other possible candidates of radius $O(s_n^{p/4})$ around it.
- Neighborhoods of radii larger the $O(s_n^{p/4})$ are irrelevant since no packet uses more then $O(s_n^{p/4})$ hops to get from the source to the destination.

Finally, let us discard any source-destination pairs that do not use a wired access point that is at least $O(s_n^{p/4})$ inside the original grid. Clearly, from the point of view of any remaining wired access points the network is completely symmetric. It remains to verify that the scaling laws have not been affected. To do so, we simply note that we have removed only some of the source-destination pairs with at least one of the two end-nodes located in $O(s_n^{-p/4})$ squarelets from the network. Since the distribution of nodes among the squarelets is by assumption symmetric, we may apply Markov's Inequality to conclude that we removed no more then $O(\omega(n)ns_n^{p/4})$ source-destination pairs. Since $s_n = o(1)$, as long as $p < 4$, the number of source destination pairs removed is $o(1)$ and the number of source destination pairs remaining is $\Theta(1) - o(1) = \Theta(1)$. Since there are $\Theta(1)$ pairs of distance 1 (as we argued in the proof of Theorem 4), the scaling law of the transport capacity is not affected.

4. Summary and Conclusions. To summarize the results obtained in this paper and put them into proper perspective it is useful to start by comparing our results with those in [9]. The simplest scenario for such an illustration is the best possible node distribution among squarelets. Recall that in this case [9] achieves a transport capacity of

$$\Theta(a_n + \sqrt{n}), \tag{54}$$

where a_n is the number of access terminals as a function of n. In our context, it is not useful to talk of "access points" since every node is a potential access point. However, we can compare the two results in terms of point-to-point links. Using the "cellular" wired architecture of [9] each access point is connected to all other access points and therefore W_n, the number of wired links grows as $W_n = O(a_n^2)$. Thus, in terms of wired links, the architecture of [9] can deliver at best

$$\Theta(\sqrt{W_n} + \sqrt{n}) \tag{55}$$

growth in transport capacity. In fact, when the average number of links per node grows as $O(1)$ with the number of nodes, we are only able to "break even" in terms of the usefulness of the wired infrastructure.

In contrast, up to an additional factor of $\omega(n)$, our approach can deliver growth rates of at least

$$\Theta\left(W_n^{\max(1-p/8,1/2)} + \sqrt{n}\right). \tag{56}$$

Additionally, with p in the range between 2 and 4, the expected number of wired links per node is $O(1)$, however, an overall improvement in the scaling law of the transport capacity is observed.

We believe that a key contribution of our work stems from the fact that the improvements over [9] are due exclusively to the fact that the queueing at the wired link access point is reduced. Whereas in [9] each access point has to support a potentially large number of links, in our scenario each access point supports only $O(\omega(n)m)$ number of links or just slightly worse than constant per packet per node. It is by reducing the queueing load at the access points that the randomized wired infrastructure enables us to improve upon the more structured approach. This intuition may lead to alternative structured approaches to introduce wired links into wireless networks that provide performance similar to the one described in this paper. Moreover, we believe that this intuition can be exported to networks where node distributions are not symmetric (see [9] for some examples) and should lead to a guiding principle for populating such networks with wired access point as well.

This paper is intended as a introductory step towards the study of how wired links may assist in wireless ad-hoc networks. While it demonstrates that this approach may yield fruitful results, it leaves several important issues unresolved. First, it would be desirable to demonstrate *almost sure* results as opposed to the *in probability* statements made here. Additionally, one may wish to remove the rather constraining restriction of a single hop and consider the scaling laws of a network where any arbitrary number of hops is permitted. Other approaches towards creating a wired infrastructure should be studied. In particular it is interesting to see whether a simple deterministic scheme can achieve or exceed the scaling laws delivered by our randomly created network.

In a broader sense, one may wish to get away from the protocol model of [6] and introduce multi-user communication approaches into the network. While the work of [21] showed that this does not affect the scaling laws in the wireless network, it is not clear that the same should hold for a heterogeneous network.

Finally, the connection between packet passing schemes in discrete processor architectures and ad-hoc networks needs deeper exploration and one may consider whether the results obtained here have applications of interest to the processor community.

Appendix

A. Supporting results for convergence properties of $\bar{l}_p(\phi, D)$. The convergence properties of $\bar{l}_p(\phi, D)$ are demonstrated in Section 3 using the following series of results.

LEMMA A.1. *For all $\phi \in (0,1)$ and any $S(L) = s_0 L^t + o(L^t)$ for $t \geq 1$ and $s_0 > 0$*

$$\lim_{D\to\infty} \sum_{L=0}^{D-1} (1-\phi)^{S(L)} < \infty. \tag{57}$$

Additionally,

$$\lim_{\phi\to 0} \lim_{D\to\infty} \sum_{L=0}^{D-1} (1-\phi)^{S(L)} = \infty. \tag{58}$$

Proof. Choose M and $0 < \epsilon < s_0$ such that $S(L) \geq (s_0 - \epsilon)L$ for all $L > M$. Thus, we have for $K \stackrel{\text{def}}{=} \sum_{L=0}^{M-1} (1-\phi)^{S(L)} < \infty$,

$$\sum_{L=0}^{\infty}(1-\phi)^{S(L)} \leq K + \sum_{L=M}^{\infty}(1-\phi)^{(s_0-\epsilon)L} \tag{59}$$

$$= K + \frac{(1-\phi)^{(s_0-\epsilon)M}}{1-(1-\phi)^{(s_0-\epsilon)}} < \infty. \tag{60}$$

This proves the first part of the lemma. To prove the second part, we would like to switch the order of the limits, i.e. we would like to write

$$\lim_{\phi\to 0}\sum_{L=0}^{\infty}(1-\phi)^{S(L)} = \sum_{L=0}^{\infty}\lim_{\phi\to 0}(1-\phi)^{S(L)} = \sum_{L=0}^{\infty} 1 = \infty. \tag{61}$$

To make the statement above rigorous, define ϵ_L such that

$$1 - \epsilon_L = (1-\phi)^{S(L)}. \tag{62}$$

Clearly, for all L, $\epsilon_L \to 0$ as $\phi \to 0$. If this convergence were *uniform* then the statement in (61) would follow directly. However, as we show next, uniform convergence is not required. Fix any positive integer N and select ϕ such that $\epsilon_L \leq \frac{1}{N+1}$ for all $L \leq N$. Then for such a choice of ϕ we have

$$\sum_{L=0}^{\infty}(1-\phi)^{S(L)} \geq \sum_{L=0}^{N}(1-\phi)^{S(L)} \tag{63}$$

$$\geq \sum_{L=0}^{N}(1-\epsilon_L) \tag{64}$$

$$\geq (N+1) - (N+1)\frac{1}{N+1} \tag{65}$$

$$= N.$$

We conclude that $\sum_{L=0}^{\infty}(1-\phi)^{S(L)}$ may be made arbitrarily large for a sufficiently small choice of ϕ and the lemma follows. □

Lemma A.2. *For all $\phi \in (0,1)$, all $p > 0$ and any $T(L) = t_0 L^t + o(L^t)$ for $t \geq 0$ and $t_0 > 0$*

$$\lim_{D\to\infty}\sum_{L=0}^{D-1}\prod_{m=0}^{L}\left(1 - \frac{\phi}{(D-m)^p}\right)^{T(m)} = \infty. \tag{66}$$

Proof. Fix $\tilde{D} < \infty$. Then, for any such choice of $\tilde{D}$,

$$\lim_{D\to\infty} \sum_{L=0}^{D-1} \prod_{m=0}^{L} \left(1 - \frac{\phi}{(D-m)^p}\right)^{T(m)}$$

$$\geq \lim_{D\to\infty} \sum_{L=0}^{D-1} \prod_{m=0}^{L} \left(1 - \frac{\phi}{(\min(D-m, \tilde{D}))^p}\right)^{T(m)} \tag{67}$$

$$= \lim_{D\to\infty} \left[\sum_{L=0}^{D-\tilde{D}} \prod_{m=0}^{L} \left(1 - \frac{\phi}{\tilde{D}^p}\right)^{T(m)} \right.$$

$$+ \sum_{L=D-\tilde{D}+1}^{D-1} \left(\prod_{m=0}^{D-\tilde{D}} \left(1 - \frac{\phi}{\tilde{D}^p}\right)^{T(m)} \right. \tag{68}$$

$$\left.\left. \times \prod_{m=D-\tilde{D}+1}^{L} \left(1 - \frac{\phi}{(D-m)^p}\right)^{T(m)} \right) \right]$$

$$\geq \lim_{D\to\infty} \sum_{L=0}^{D-\tilde{D}} \prod_{m=0}^{L} \left(1 - \frac{\phi}{\tilde{D}^p}\right)^{T(m)} \tag{69}$$

$$= \lim_{D\to\infty} \sum_{L=0}^{D} \left(1 - \frac{\phi}{\tilde{D}^p}\right)^{\sum_{m=0}^{L} T(m)}. \tag{70}$$

It follows that

$$\lim_{D\to\infty} \sum_{L=0}^{D-1} \prod_{m=0}^{L} \left(1 - \frac{\phi}{(D-m)^p}\right)^{T(m)}$$

$$\geq \lim_{\tilde{D}\to\infty} \lim_{D\to\infty} \sum_{L=0}^{D} \left(1 - \frac{\phi}{\tilde{D}^p}\right)^{\sum_{m=0}^{L} T(m)} \tag{71}$$

$$= \lim_{\phi\to 0} \lim_{D\to\infty} \sum_{L=0}^{\infty} (1-\phi)^{\sum_{m=0}^{L} T(m)} = \infty, \tag{72}$$

where the last equality follows by Lemma A.1. □

Lemma A.3. *For any $\alpha \in (0,1)$ and $\gamma \in (0,1)$*

$$\sum_{L=0}^{D-1} \alpha^{\frac{L}{D^\gamma}} = \Theta\left(D^\gamma\right). \tag{73}$$

Proof. We have

$$\sum_{L=0}^{D-1} \alpha^{\frac{L}{D^\gamma}} = \frac{1 - \left(\alpha^{D^{-\gamma}}\right)^D}{1 - \alpha^{D^{-\gamma}}} = \frac{1 - \alpha^{D^{1-\gamma}}}{1 - \alpha^{D^{-\gamma}}}. \tag{74}$$

Thus, the asymptotic rate of growth of $\sum_{L=0}^{D-1} \alpha^{\frac{L}{D^\gamma}}$ is the same as the asymptotic rate of growth of

$$f(x) \overset{\text{def}}{=} \frac{1 - \alpha^{x^{1-\gamma}}}{1 - \alpha^{x^{-\gamma}}}. \tag{75}$$

To understand the asymptotic rate of growth $f(x)$, first note that

$$\lim_{x\to\infty} \alpha^{x^{1-\gamma}} = 0 \tag{76}$$

and

$$\lim_{x\to\infty} \alpha^{x^{-\gamma}} = 1. \tag{77}$$

Then take some $r > 0$ and consider

$$\lim_{x\to\infty} \frac{f(x)}{x^r} = \lim_{x\to\infty} \frac{x^{-r}}{1-\alpha^{x^{-\gamma}}} \tag{78}$$

$$= \lim_{x\to\infty} \frac{-rx^{-r-1}}{\gamma \ln(\alpha)\alpha^{x^{-\gamma}} x^{-\gamma-1}} \tag{79}$$

$$= \lim_{x\to\infty} \frac{-r}{\gamma \ln(\alpha)} x^{\gamma-r}. \tag{80}$$

Thus,

$$\lim_{x\to\infty} \frac{f(x)}{x^r} = \begin{cases} 0 & \text{if } r > \gamma \\ -\frac{1}{\ln \alpha} & \text{if } r = \gamma \\ \infty & \text{if } r < \gamma \end{cases} \tag{81}$$

which completes the proof of the lemma. □

Lemma A.4. *Consider* $0 < \alpha < 1$, *and* $q > p > 1$ *then*

$$\sum_{L=0}^{D} \alpha^{\frac{L^q}{D^p}} = \Theta\left(D^{\frac{p}{q}}\right). \tag{82}$$

Proof. Let L^* denote the highest integer L such $\frac{L^q}{D^p} \le 1$. Thus,

$$L^* = \lfloor D^{\frac{p}{q}} \rfloor. \tag{83}$$

Then

$$\sum_{L=0}^{D} \alpha^{\frac{L^q}{D^p}} \le L^* + 1 + \sum_{L=L^*}^{D} \alpha^{\frac{L^q}{D^p}}. \tag{84}$$

Recalling now that for $L > L^*$ $\frac{L^q}{D^p} \ge 1$, we have

$$\sum_{L=L^*}^{D} \alpha^{\frac{L^q}{D^p}} \le \sum_{L=L^*}^{D} \alpha^{\frac{L}{D^{p/q}}} = O\left(D^{\frac{p}{q}}\right), \tag{85}$$

where the last equality follows by Lemma A.3 and we use $O(\cdot)$ and not $\Theta(\cdot)$ because the contribution from the missing leading terms grows with D and may reduce the overall rate of growth. The lemma follows since $L^* = \Theta\left(D^{\frac{p}{q}}\right)$. □

REFERENCES

[1] P. Billingsley, *Probability and Measure*, John Wiley & Sons, New York, NY, 3rd edition, 1995.

[2] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, Cambridge, MA, 2nd edition, 2001.

[3] S. N. Dorogovtsev and J. F. F. Mendes, *Exactly solvable analogy of small- world networks*, arXiv: cond-mat, (9907445), July 1999.

[4] I. S. Gradshteyn, I. M. Ryzhik, and A. Jeffrey (Editor). *Tables of Integrals, Series, and Products*, Academic Press, New York, 5th edition, 1994.

[5] P. Gupta, *Design and Performance Analysis of Wireless Networks*, PhD thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois, 2000.

[6] P. Gupta and P. R. Kumar, *The capacity of wireless networks*, IEEE Transactions on Information Theory, IT-46(2000), pp. 388–404.

[7] M. Kaufman, J. F. Sibeyn, and T. Suel, *Derandomizing routing and sorting algorithms for meshes*, In: Proc. 5th Symp. on Discrete Algorithms. LNCS, Springer-Verlag, 1995.

[8] R. V. Kulkarni, E. Almaas, and D. Stroud, *Exact results and scaling properties of small-world networks*, arXiv: cond-mat, (9908216), August 1999.

[9] S. R. Kulkarni and P. Viswanath, *Throuhgput scaling in heterogeneous networks*, In: Proc. 2003 Int. Symp. on Information Theory.

[10] S. R. Kulkarni and P. Viswanath, *A deterministic approach to throughput scaling in wireless networks*, IEEE Transactions on Information Theory, IT-50(2004), pp. 1041–1049.

[11] M. Kunde, *Block gossiping on grids and tori: deterministic sorting and routing match the bisection bound*, In: Proc. European Symp. on Algorithms, pages 260–271, 1991.

[12] B. Liu, Z. Liu, and D. Towsley, *On the capacity of hybrid wireless networks*, In: Proc. INFOCOM 2003.

[13] M. E. J. Newman, C.Moore, and D. J. Watts, *Mean-field solution of the small-world network model*, Physical Review Letters, 80(2000), pp. 3201–3204.

[14] M. E. J. Newman, I. Jensen, and R. M. Ziff, *Percolation and epidemics in a two-dimensional small world*, Physical Review E, 65:021904-1–021904-7, 2001.

[15] M. E. J. Newman and D. J. Watts, *Scaling and percolation in the small- world network model*, Physical Review E, 60(199), pp. 7332–7342.

[16] A. Reznik, S. R. Kulkarni, and S. Verdú, *Scaling laws in random heterogeneous networks*, In: Proc. 2004 Int. Symp. on Information Theory.

[17] S. H. Strogatz, *Exploring complex networks*, Nature, 410(2001), pp. 268–276.

[18] D. J. Watts, *The Structure and Dynamics of Small-World Systems*, PhD thesis, Cornell University, Ithaca, NY, 1997.

[19] D. J. Watts, *Small Worlds. The Dynamics of Networks between Order and Randomness*, Princeton University Press, Princeton, NJ, 1999.

[20] D. J. Watts and S. H. Strogatz, *Collective dynamics of 'small-world' net- works*, Nature, 393(1998), pp. 440–442.

[21] L.-L. Xie and P. R. Kumar, *A network information theory for wireless communication: Scaling laws and optimal operation*, IEEE Transactions on Information Theory, IT-50(2004), pp. 748–767.