International Scholarly Research Network
ISRN Communications and Networking
Volume 2012, Article ID 362603, 23 pages
doi:10.5402/2012/362603

Review Article

Network Topology Models for Multihop Wireless Networks

András Faragó

Department of Computer Science, The University of Texas at Dallas, Richardson, TX 75080, USA

Correspondence should be addressed to András Faragó, farago@utdallas.edu

Received 30 August 2012; Accepted 19 October 2012

Academic Editors: A. Pacheco, T. Sakamoto, Y.-J. Suh, H. Vu, and W.-D. Zhong

A typical feature of huge, random network topologies is that they are too large to allow a fully detailed description. Such enormous, complex network topologies are encountered in numerous settings and have generated many research investigations. Well-known examples are the Internet and its logical overlay networks, such as the World Wide Web as well as online social networks. At the same time, extensive and rapidly growing wireless ad hoc and sensor networks also lead to hard topology modeling questions. In the current paper, we primarily focus on large, random wireless networks but also consider Web and Internet models. We survey a number of existing models that aim at describing the network topology. We also exhibit common generalizations of various sets of models that cover a number of known constructions as special cases. We demonstrate that higher levels of abstraction, despite their very general nature, can still be meaningfully analyzed and offers quite useful and unique help in solving certain hard networking problems. We believe that this research area can and will generate further significant contributions to the analysis of very large networks.

1. Introduction

Many of the communication networks that we use today, or expect to use in the future, have enormous size. This applies not only to the physical networks, including the Internet as well as emerging ubiquitous wireless networks and large scale sensor networks, but also, or even more, to logical overlay networks, such as the World Wide Web. For example, the number of web pages, according already to a 2006 article [1], was as high as 53.7 billion, already at the time of writing that study. Out of the 53.7 billion, 34.7 billion web pages were indexed by Google. Since then, these numbers grew even further. Beyond the sheer size, the usage of these networks is also expected to be extremely heterogeneous, encompassing a huge number of different applications, traffic patterns, diverse requirements, and areas, including business, science, learning, entertainment, and social networking. At the same time, their physical basis is also heterogeneous, including wired, wireless, and optical subnetworks. All this is expected to eventually merge into a ubiquitous, global sociotechnical infrastructure.

To understand and reason about huge socio-technical networks, including methods for designing/optimizing them, the traditional network analysis and modeling approaches are generally insufficient, due to their *limited scalability*. Simulation is usually feasible only up to a rather limited network size. Conventional analysis methods, such as teletraffic theory queuing network modeling, and also face an uphill battle, quickly losing ground in huge networks. At the same time, modeling and analysis are still indispensable, since one may not be able to experiment with the different variants of a new solution via large-scale practical deployment, as it can have a prohibitive cost.

This situation, in which one deals with networks of practically infinite size, has naturally led to the emergence of novel analysis and modeling approaches. They can generally be characterized by having a more abstract, "bird's eye" view of the network and often relying on *asymptotic analysis* on the mathematical side. The special advantage of the asymptotic analysis is that it converts the growing size from a foe to a friend: the larger, the better, from the asymptotic point of view. While it is clear that such methods cannot help much in *local* technical tasks, such as configuring a specific router, they have their important place in the higher layers of the network modeling hierarchy. We survey the models and results in this area, with primary focus on large wireless networks with no infrastucture (ad hoc and sensor networks), based on the extensive literature and the author's

own earlier work [2–26]. As a first step, in the next section we briefly survey how the random network modeling approach emerged.

2. Historical Retrospective: Emergence of Random Network Models

2.1. Internet and Web Topology Models. The first major wave of work in the considered direction was the statistical analysis of the Web graph, in which the nodes are web pages and the edges are the hyperlinks. Similar investigations were made regarding the physical Internet topology as well. Below, we summarize some of the major directions.

2.1.1. Experimental Statistical Analysis. Several research groups in the late 90s independently observed that the node degrees in this graph are distributed according to a *power law* (Kumar et al. [27], Barabási et al. [28, 29], and Broder et al. [30]). Similar phenomena were observed by Faloutsos et al. [31] in the physical Internet topology.

The power law degree distribution means that if n_k denotes the number of nodes with degree k, and $n = \sum n_k$, then

$$\frac{n_k}{n} \sim k^{-\beta}, \tag{1}$$

where $\sim$ denotes asymptotic proportionality, and β is called the *power law exponent*. The typical measured values of β are between 2 and 3. All this was in sharp contrast with traditional random graph models that have independent edges and exhibit (asymptotically) Poisson node degree distributions. The latter models are known as Erdős-Rényi random graphs.

To describe the observed network structure, Barabási and Albert [28] coined the term "*scale-free network*", based on the observation that in a power law distribution the rescaling of the considered quantity preserves the same power law, changing it only with a constant factor. This quickly became very popular and triggered the statistical analysis of "scale-freeness" of network topologies not only in (physical or logical) communication networks, but also in networks that arise in biology, genetics, epidemiology, linguistics, electric power distribution, social sciences, and in many other areas; see, for example, the books [32–36] and hundreds of further references therein.

In retrospect, one may say that "scale-free networks" generated somewhat more hype than substance. It was rightfully pointed out, for example, by Li et al. [37] and Alderson et al. [38] that the power law degree distribution alone can easily fall short from adequately modeling the Internet topology, if no other domain specific knowledge is applied.

2.1.2. Generative Models. A parallel major wave of research was to create generative models. In contrast to experimental statistical analysis, generative models aim at explaining the observed network structures, providing algorithmic approaches to generate them, and also offering the opportunity for in-depth mathematical analysis. The first such model that became well known was the *Preferential Attachment* model of Barabási and Albert [28]. This model generates a graph such that new nodes are more likely to get connected to those that already have a higher degree. Although the authors did not provide a rigorous analysis, only an approximate reasoning, based on the mean-field approach of physics, the model certainly had intuitive appeal ("the rich get richer" principle). This model had an explanatory power and generated scale-free graphs, so it triggered many follow-up investigations. The first truly rigorous formulation and deep analysis of a preferential attachment model, called Linearized Chord Diagram (LCD) model, was provided by Bollobás et al. [39].

Since then, numerous static and evolving graph models of networks have been proposed and analyzed, both experimentally and with serious mathematical rigor, primarily focusing on asymptotic properties. A few examples are: the ACL model of Aiello et al. [40]; the copying model of Kleinberg et al. [41]; the growth-deletion model of Chung and Lu [42]; the self-similar Kronecker-graph model of Leskovec et al. [43]; the compressible Web model of Chierichetti et al. [44]; the forest fire model of Leskovec et al. [45]; the geometric preferential attachment model of Flaxman et al. [46]; the spatial preferential attachment model of Aiello et al. [47]; the random perturbation model of Flaxman [48]; a large number of other models and variants, with a lot of intellectual power in their analysis.

As an example of such models, consider the following, which is due to Chung and Lu [49]. It is motivated by the fact that since the degree distribution may vary, and it appears a significant issue in characterizing a random graph model, therefore, it is important that *any* feasible degree distribution can be efficiently generated. This is accomplished by the model of Chung and Lu [49], which is a static (i.e., not evolving) random graph model, having the advantage of high flexibility, in the sense that it can exhibit any possible degree distribution (including power law with any exponent) in the expected node degrees. The model is built as detailed below.

Let $\mathbf{d} = (d_1, \ldots, d_n)$ be any degree sequence. For example, it may be chosen to exhibit power law with exponent β. In that case, the number k is included in the sequence n_k times such that $n_k/n \sim k^{-\beta}$. Then between any nodes i and j, an edge is drawn independently with probability

$$p_{ij} = \frac{d_i d_j}{\sum_{k=1}^{n} d_k}. \tag{2}$$

It is not hard to show that in the arising random graph, the expected degree of node i will indeed be d_i. In this sense, any possible degree distribution can be represented. Note that in this model, the degree values are preselected, only the edges are random. The random edges are drawn such that the arising *expected* degrees obey the given degree sequence.

As a highly nontrivial feature of this model, Chung and Lu [49] prove, among other things, the following property. Let G be the arising random graph, and let $L(G)$ denote the *average hop distance* between two randomly chosen nodes. Then (under some minor technical conditions for which we refer to the original article [49]) the following holds.

If the expected node degrees are distributed by a power law distribution with exponent $2 < \beta < 3$, then the average hop distance is bounded as

$$L(G) \leq (2 + o(1)) \frac{\log \log n}{\log(1/(\beta - 2))}, \quad (3)$$

where $o(1)$ means a function of n that tends to 0, as $n \to \infty$.

This is a useful result, as it provides a nontrivial upper bound on expected hop distances that otherwise would be hard to obtain. For example, assume we consider a piece of the Web graph with 10 million web pages, and we measured that it has a power law degree distribution with exponent $\beta = 2.5$. From this information alone, can we say something about the average hop distances between web pages? By the cited result (assuming that with $n = 10^7$ we can approximately take $o(1) = 0$), we have

$$L(G) \leq 2 \frac{\log \log 10^7}{\log(1/(2.5 - 2))} \approx 8, \quad (4)$$

which is a reasonably low-specific bound on the average hop distance. It would clearly be *much* harder to obtain the same via direct computations on a graph of 10 million nodes.

2.2. Scalability of Large Wireless Networks. About the same time when the Web and Internet topology investigations began, another independent wave of asymptotic network modeling was initiated by Gupta and Kumar [50]. This direction focused on analyzing the scalability of large wireless networks, primarily ad hoc and sensor networks, from the viewpoint of fundamental limits for transport capacity and related properties. This line of research also attracted much attention. Interestingly, and unfortunately, most of the results are negative. Specifically, they show under various conditions that the achievable per node throughput tends to 0 with growing network size, which we call the *vanishing throughput effect*. Even maintaining global network connectivity is impossible under rather general conditions, if we want to apply nodes with finite processing capacity, see Faragó [6].

2.2.1. The Vanishing Throughput Effect. To explain this phenomenon, let us review some fundamental results in this direction. The most important result in the paper of Gupta and Kumar [50] considers the achievable throughput per source-destination (S-D) pair in a large random network, where n nodes are placed independently and uniformly at random in a planar disk of unit area. Thus, the network topology is random but static. The nodes have the same transmission radius, which can be chosen arbitrarily. They can transmit with a fixed maximum rate (bit/sec), but they are allowed to divide the channel into subchannels without any constraint, in any domain (e.g., frequency, time, and code). The nodes communicate using arbitrary protocols for channel access and routing, with the only restriction that there is a minimal requirement of interference avoidance in the same (sub)channel to ensure successful receptions. The considered traffic pattern is that each node has a randomly chosen destination in the network. The authors prove that in this general model, the achievable throughput per S-D pair is

$$\Theta\left(\frac{1}{\sqrt{n \log n}}\right). \quad (5)$$

Regarding scalability, the key message of this result is that the achievable throughput per S-D pair tends to zero as the network size grows to infinity. This is what we call the *vanishing throughput effect*. It means, the network is fundamentally not scalable, since it becomes unable to usefully operate when it grows extremely large.

It is then a natural question to investigate how robust is the vanishing throughput effect under changing the model. Is it merely an artifact due to some special modeling assumptions, or is it a fundamental phenomenon that remains true in any reasonable model?

Already in their original paper [50], Gupta and Kumar point out that if the area where the nodes are located is the surface of a sphere rather than a planar disk, the results still remain in effect. They also show that no essential change results even if instead of random node placement, the nodes are located optimally and their transmission radiuses are also chosen optimally. For 3-dimensional network topology (random node placement in a unit cube), Toumpis and Goldsmith [51] show that, although the rate of vanishing is different, the per node throughput still tends to zero. Similar results are obtained by Gupta and Kumar in [52].

Various other static network scenarios were also analyzed by different authors. For example, Yi et al. [53] show that the use of directional antennas, although can increase the capacity, still does not invalidate the vanishing throughput effect. Peraki and Servetto [54] prove that even if the nodes can generate multiple beams in parallel, aimed at multiple receivers, the achievable improvement is still insufficient to sustain positive per node throughput in the limit. Kozat and Tassiulas [55] consider the scenario when the ad hoc network has infrastructure support, that is, besides the ad hoc operation, it can also use a cellular-like infrastructure. They show that although this support can bring an exponential improvement in the capacity, yet the per node throughput still asymptotically vanishes, although at a much slower rate. Infrastucture-based hybrid scenarios were also considered by Liu et al. [56] and Toumpis [57], all pointing at the direction that vanishing throughput remains in effect, given that the infrastructure size is not overwhelming compared to the ad hoc network size.

2.2.2. Escapes from the Vanishing Throughput Effect. At this point one may wonder: has any model been found at all for large random ad hoc networks in which constant (i.e., nonvanishing) per node throughput can be maintained in the asymptotic regime? The answer is yes but, unfortunately with serious reservations.

One possibility is to consider the effect of mobility, since all the above results refer to random but static networks. Grossglauser and Tse [58] analyze the mobile version of the Gupta-Kumar [50] model, assuming that the nodes move

randomly and independently with uniform stationary distribution over the unit disk. The key advantage of mobility in this setting is that either a source node can get close to its destination and send the data directly, or it can spread the packets among other nodes, which relay the data to the destination whenever they get close to it. A typical general state of this network is that each node buffers many packets destined to different destinations, and whenever a node gets close to another node, it sends the data, if any, waiting for that destination. Grossglauser and Tse [58] prove that in this setting, it is possible to maintain asymptotically constant per session throughput. The problem, however, is that it comes at the price of infinitely growing delay. El Gamal et al. [59] analyzed the trade-off between throughput and delay and found that with bounded node movement speed (it is certainly bounded at least by the speed of light) the average delay grows to infinity with the network size. Note that each packet travels at most 2 hops in this network (either directly to the destination or through a single-relay node), so the delay is due to waiting for getting close to the destination. Another problem is that mobility in the Grossglauser-Tse model is not an *option*, it is an *obligation*: essentially, the nodes *must* be on the move all the time to make the solution workable.

Another possible escape from the vanishing throughput effect is to restrict the traffic pattern in the random static network. Li et al. [37] argue that the random destination scenario in Gupta-Kumar model can be regarded as a worst case traffic distribution. The authors find that if the traffic tends to be local, in the sense that the expected route hop length is bounded, then constant throughput can be maintained in the static random network. This is still not sufficiently reassuring, however, since it would be desirable that the network can handle any traffic pattern, rather than only a small subset of the possibilities. Furthermore, if we do not want to replace the network with a multitude of small local networks that are disconnected from each other, then the problem of the infinitely growing node degrees still remains there.

Another approach to improve the asymptotic throughput scaling is to apply more sophisticated physical layer techniques such as MIMO. A paper of Özgür et al. [60] shows that with intelligent, hierarchical node cooperation and MIMO techniques, it is possible to achieve an aggregate throughput of $\Omega(n^{1-\epsilon})$ for n randomly placed nodes, where $\epsilon > 0$ is an arbitrarily small, but positive, constant. This translates into a per S-D pair throughput of $\Omega(1/n^{\epsilon})$. If ϵ is chosen small enough, then it is close to the ideally sought constant throughput per node. Nevertheless, since ϵ must be a fixed positive constant, therefore, no matter how small it is, the per node throughput will still vanish as $n \to \infty$.

2.2.3. Why Does the Throughput Vanish? To better understand a fundamental reason for the vanishing throughput, let us briefly discuss what causes it and how in the Gupta-Kumar model [50].

One may observe that the transmission radius r of nodes has two opposite effects. If r is small, then more hops are needed to deliver a packet to its destination, since in each hop the packet can advance at most a distance of r towards its destination. Therefore, small r causes a growing burden of nodes to serve as relays, which decreases the end-to-end throughput. This would justify choosing r as large as possible. On the other hand, large transmission radius increases the interference, which causes the throughput to decrease again, so interfence reduction would require choosing r as small as possible.

The detailed analysis of the above conflicting tendencies in [50] shows that the forwarding burden is proportional to $1/r$, since the average route hop length is proportional to $1/r$. On the other hand, the interference increases quadratically with the radius, due to the area involved, which is proportional to r^2. Balancing the two effects yields the conclusion that r has to be chosen as small as possible, as the interference reduction, due to its quadratic nature, brings more benefit than what is lost by longer routes. Specifically, the joint effect results in the formula

$$\lambda(n) = O\left(\frac{1}{nr(n)}\right), \tag{6}$$

where n is the number of nodes, $\lambda(n)$ is the throughput in bit/sec for each S-D pair, and $r(n)$ is the transmission radius (as a function of n). It is clear from (6) that if we want to avoid $\lambda(n) \to 0$, then $r(n)$ has to tend to 0 at least as fast as $1/n$, so that $nr(n)$ does not grow to infinity.

There is, however, an effect that does not allow to choose the transmission radius arbitrarily small. This is the requirement that the network topology must be *connected*, since otherwise communication between certain endpoints becomes impossible. The discussed model uses an earlier result of the same authors [61] about the needed transmission radius for asymptotic connectivity. They show that if the nodes are placed uniformly at random in a unit disk, then the network is connected with probability approaching 1 if and only if the transmission radius satisfies

$$\pi r^2 = \frac{\ln n + c(n)}{n}, \tag{7}$$

with $c(n) \to \infty$. According to this, it is necessary for connectivity that

$$r > \sqrt{\frac{\ln n}{\pi n}} \tag{8}$$

holds for all large enough values of n. Combining it with (6), we obtain

$$\lambda(n) = O\left(\frac{1}{n\sqrt{(\ln n)/n}}\right) = O\left(\frac{1}{\sqrt{n \ln n}}\right), \tag{9}$$

clearly showing the vanishing throughput effect.

Nevertheless, as we have already mentioned earlier, there are notable exceptions, utilizing various effects, such as mobility (Grossglauser and Tse [58]), restricted traffic pattern (Li et al. [62]), using infrastructure (Liu et al. [56]), or relaxing the condition of full connectivity (Dousse et al. [63] and Faragó [10]), to recall only a few examples. Therefore,

the issue of wireless network scalability is still under further research.

It is worth mentioning, that the graph models that are used in the wireless network investigations are very different from the Internet and Web models. The random graphs in wireless network analysis are based on geometric considerations and termed geometric random graphs. They also have a rich set of analytical results, see, for example, the books of Franceschetti and Meester [64] and Penrose [65]. In a sense, geometric random graphs are between the classical Erdős-Rényi model and those graph models that are used to describe the Internet and Web topologies. Specifically, geometric random graphs have (asymptotically) Poisson node degree distributions, just like the Erdős-Rényi random graphs. That is, geometric random graphs (modeling wireless network topologies) do *not* exhibit scale-free behavior. On the other hand, their edges are not independent, just like in the Internet/Web models, so they have many properties that are distinctively different from the Erdős-Rényi random graphs.

2.3. The Current Situation. As briefly reviewed above, there exists a vast and rather diverse body of various graph based network models that are mostly analyzed from the viewpoint of asymptotic properties. Note that beyond the theoretical advances, they also have emerging practical applications, such as Internet topology generators, search engine optimization, protocol design, and optimization in wireless networks.

The current situation on the model development and analysis side (which is our primary interest) can be characterized with the following.

(i) The diversity of models also led to the diversity of analysis methods. With minimal exaggeration, one can say that a new analysis method has to be invented almost each time when a new model is proposed. There is a sense of missing unification and a lack of general methods that apply to large families of different models.

(ii) The analysis is often very hard and typically cannot rely on the well-developed methods of classical random graph theory, as pointed out by leading experts in the theory of random graphs (Bollobás and Riordan [66]).

(iii) Despite the existence of emerging applications, there is still a large gap between analysis results of descriptive nature and methodology/algorithms that provide meaningful help in network design problems.

(iv) Validation of models is a problem. As pointed out by Flaxman [48]: "Unfortunately, it is much easier to propose a generative model than to refute one."

3. Large, Multihop Wireless Networks

Wireless networks of large size, random topology and no supporting infrastructure, such as ad hoc and sensor networks, are expected to play an important role in the future. The random network topology of these systems is frequently described by various random graph models, most often by some variants of *geometric random graphs*. First we review some of the typical classes of graphs that are used in this context.

3.1. Frequently Used Graph Classes. An important class is the *Unit Disk Graph (UDG)* [67] model of the network topology. A UDG is a graph that is defined by the (planar) geometry of node positions. It is assumed that each node has the same transmission radius r, and two nodes are connected by a link if and only if they are within distance r (which is often normalized to $r = 1$, hence the name). In other words, the radio range of each node is just a circular disk. As a critical difference from the physical model, in a UDG it does not matter where the rest of the nodes are located and how much interference they generate.

A clear advantage of UDGs is that a number of important algorithmic problems that are NP-complete for general graphs become solvable in polynomial time for this special class [68], thus allowing much more efficient protocols.

Unfortunately, the UDG model is quite simplistic, it is rather far from accurately reflecting the actual radio network topology. A refinement is the *Quasi-Unit Disk Graph (Q-UDG)* model [69], in which a *shrink factor* ρ is added, with $0 < \rho < 1$, for describing the radio range of a node by two concentric circular disks, the outer one with radius r and the inner one shrunk by the factor ρ, yielding radius ρr. If two nodes are at most ρr distance apart, then they are always connected by a link. If they are more than r apart, then they are never connected. Finally, if the distance is between ρr and r, then the link may or may not exist. Geometrically, this means that the radio range of a node can have arbitrary shape but moderated by the requirement that it should be between a circumscribed circle of radius r and an inscribed circle of radius ρr.

A nice feature of Q-UDGs is that, while providing a more general network topology model, they still preserve the algorithmic advantages of UDGs, at the price of an additional $1/\rho^2$ factor in complexity [69]. Thus, if the shrink factor ρ is a not too small constant, then most of the UDG advantages carry over, with only a constant factor penalty in complexity.

Another natural issue is that different nodes may transmit with different power or have different spectrum-dependent attenuation of the transmission signal [70]. This leads to the concept of *Disk Graph (DG)*, which differs from the UDG in that each node i has its own, possibly different, transmission radius r_i, and two nodes are connected by an undirected link if they are mutually in each other's range. DGs are somewhat less friendly from the algorithmic point of view than UDGs and Q-UDGs but still better than general graphs and still allow efficient solutions or approximations for a number of algorithmic problems, as we investigated in [26].

Similarly to the generalization that leads to the Q-UDG concept, one can also introduce *Quasi-Disk Graphs (Q-DG)*, by adding a shrink factor ρ that allows to refine the radio range description as for Q-UDG.

All the above graph models can naturally be extended to higher dimensions, replacing the disks by balls in the appropriate space.

A common nontrivial generalization of all these graphs, the *Bounded Independence Graph (BIG)* model, is also worth mentioning [68]. (It is also referred to as *Bounded Growth Graph* [71].) This class is defined by the requirement that the maximum number of independent nodes (a set of nodes in a graph is called independent if there is no edge between them) within the k-hop neighborhood $\mathcal{N}_k(v)$ of any node v is bounded by a polynomial of k. Although this definition is based purely on the graph structure and does not have a direct geometric meaning, it can still be related to geometry through the concept of *doubling metric spaces* [68]. These are metric spaces in which any ball of radius r can be covered by a finite number of balls of radius $r/2$. This property does not hold for all metric spaces, although it holds for Euclidean spaces of any finite dimension (radio propagation properties may lead to a "radio-distance" that is quite different from Euclidean). It can be shown that if a geometric graph is defined in a doubling metric space, in analogy with UDG or DG, then it is always a bounded independence graph [68]. A nice feature of this class is that a number of hard algorithmic problems become efficiently solvable in it [71].

So far, we have described these classes deterministically, ignoring randomness. Of course, from each graph class one can generate random members, according to various probability distributions. These are usually defined indirectly, through some generating mechanism. For example, if we pick the node positions uniformly at random in a planar domain, for example, a square, and then construct a UDG over these nodes, then we get a random unit disk graph.

All these graph classes are related to some kind of geometric insight. It is not surprising, since geometry and distance play a key role in forming the radio network topology. On the other hand, radio propagation (with possible obstacles and other irregularities) can induce much more complicated distances that may not satisfy the mathematical distance axioms, primarily the triangle inequality. Nevertheless, even in this more complicated situation, it is still possible to meaningfully analyze geometric-like graphs and prove nontrivial results about important properties, such as connectivity, as we are going to see in connection with our *Abstract Geometric Random Graphs*.

3.2. Some Results on Geometric Random Graphs. In our context, the typical model in the literature is the following. n nodes are placed randomly and independently in a regularly shaped domain, which is most often a square or a disk in the Euclidean plane. Each node has a transmission radius of r, and two nodes are connected with a link if and only if their distance is at most r. The radius is typically a function of n in the asymptotic analysis. Note that this *geometric random graph* model is different from the more traditional independent-edge random graph model (often called Erdős-Rényi random graphs), since in the geometric setting the edges are correlated. While many deep results are available for the traditional model (see, e.g., [39]), they are not directly applicable to ad hoc networks.

As one of the early results on ad hoc network connectivity, Philips et al. [72] proved the following. Assume the nodes are chosen in a square of area A from a planar Poisson process of constant density D. (This is asymptotically equivalent to choosing the points independently at random from the uniform distribution over the domain.) The area A of the square grows to infinity, while the density remains constant, so the number of nodes also grows to infinity. A result of [72] is that for any $\epsilon > 0$, if

$$r \leq \sqrt{(1-\epsilon)\ln A/\pi D}, \tag{10}$$

then the network gets disconnected with probability approaching 1, as $A \to \infty$. It implies that the transmission radius r cannot remain bounded if we want connectivity with a growing number of nodes that are placed with constant density in a growing area. If we want to keep the entire domain constant, it can be achieved by normalizing the quantities with the domain size.

Gupta and Kumar [61] analyzed the situation when the nodes are placed uniformly at random in a unit disk in the plane. They investigated the *critical transmission range*, that is, the minimum transmission radius needed for connectivity. They obtained that the network is asymptotically connected with probability 1 if and only if the transmission radius satisfies

$$\pi r^2 = \frac{\ln n + c(n)}{n} \tag{11}$$

with $c(n) \to \infty$. Wan and Yi [73] extended the results to k-connectivity on a unit disk and square.

The above result of Gupta and Kumar [61] can also be deduced from Penrose's formula [74, 75] on the asymptotic distribution of the longest edge of a minimum spanning tree in a random geometric graph, since the longest edge length in a minimum spanning tree is known to be equal to the critical transmission range [75]. Penrose analyzes a rather general setting, including a large class of probability density functions, allowing that the nodes are placed in an irregularly shaped area, in a (possibly) high-dimensional space. It is assumed, however, that it is a normed space, thus the considered distance is not arbitrary. It is derived from a norm, which does not apply to every distance metric.

The node degree required for connectivity was analyzed by Xue and Kumar [76] in a somewhat different model. They place n nodes uniformly at random in a unit square in the plane, but the edges are not defined via a transmission radius. They consider the graph that arises if each node is connected to its ϕ_n nearest neighbors. That is, an edge between nodes i, j exists, if either i is among the ϕ_n nearest neighbors of j or j is among the ϕ_n nearest neighbors of i. The arising graph is denoted by $G(n, \phi_n)$. Note that this model is somewhat different from the transmission radius-based geometric random graph model. For example, in the latter it may happen with positive probability that a node is isolated. On the other hand, by definition, no node can be isolated in $G(n, \phi_n)$ for $\phi_n > 0$. This captures the situation when each node adjusts its transmission radius individually until it has the desired number of neighbors. It is shown in [76] that

$G(n, \phi_n)$ is asymptotically almost surely connected (i.e., the probability that it is connected tends to 1 as $n \to \infty$), if $\phi_n > 5.1774 \ln n$. On the other hand, if $\phi_n < 0.074 \ln n$, then the graph will be asymptotically almost surely disconnected. Thus, the critical number of neighbors is on the order of $\ln n$, but the gap between the upper and lower bounds is significant, their ratio is close to two orders of magnitude. The upper bound was improved by Wan and Yi [73] to $\alpha e \ln n$, where $e \approx 2.718$ is the base of the natural logarithm and $\alpha > 1$ is any real constant.

The issue of connectivity is closely related to the minimum degree in geometric random graphs. Penrose [77] shows that if the points are chosen uniformly at random in a d-dimensional unit cube ($d \geq 2$) and the edges are added one by one in the order of increasing length, then the graph becomes k-connected almost surely whenever all nodes reach degree $\geq k$. This can be interpreted that for large geometric random graphs, almost surely, the sole reason for not being k-connected is the trivial reason that at least one node does not have sufficient degree. For simple connectivity, it means that the graph is connected asymptotically almost surely if and only if the transmission radius is chosen large enough so that there are no isolated nodes. All this holds not only with the Euclidean distance, but also for all $l_p, 1 < p \leq \infty$, distances.

For the l_∞ distance, Appel and Russo [78] determined the exact asymptotics of the minimum transmission radius r_n needed to eliminate isolated nodes. They obtain

$$r_n \sim \left(\frac{\ln n}{2dn}\right)^{1/d}, \tag{12}$$

where $d \geq 1$ is the dimension of the space in which the nodes are chosen uniformly at random from the unit cube. The combination of this with the above mentioned results of Penrose on the degrees yields precise asymptotics for the needed transmission radius for k-connectivity, if the nodes are placed uniformly at random in a d-dimensional unit cube ($d \geq 2$) and the distance is based on the maximum norm (l_∞ distance). For Euclidean distance in the plane, the asymptotics for the needed transmission radius for simple connectivity ($k = 1$) was derived earlier by Dette and Henze [79].

4. Generalizations of Geometric Random Graphs

4.1. Premetric Random Graphs. In real networks, the radio coverage area is often more complicated than a simple disk, since signal propagation may depend on various factors, such as terrain, foliage, and weather and atmospheric conditions. Therefore, the range can be direction and location dependent. In other words, the existence of a link between two nodes is not necessarily determined by their distance alone. Moreover, the status of a node can possibly be characterized not only by its random spatial coordinates, it may include other random variables, as well. For example, one may also include the traffic load, processor load, queue length, jamming level, battery level, or any other parameter. If the condition of link existence is expressed such that a distance function based on the variables does not exceed a threshold, then it is not surprising that this (abstract) distance will not be a metric in the traditional sense. To capture this complexity, a generalized model was proposed in [8]. It is based on a space that is much more general than what is used in other models.

Definition 1 (premetric space). A premetric space is a pair $\mathcal{M} = (S, \rho)$, where S is a set and $\rho : S \times S \mapsto \mathbf{R}$ is a real-valued function, such that for every $x, y \in S$ the following hold:

(i) $\rho(x, y) \geq 0$,

(ii) $\rho(x, y) = 0$ if and only if $x = y$,

(iii) $\rho(x, y) = \rho(y, x)$.

The difference between a pre-metric space and the usual concept of a metric space is that in the pre-metric space, the well-known *triangle inequality* $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ does not have to hold. Thus, $\rho(x, y)$ is not necessarily a *distance* in the usual sense. Still, to help visualization, we refer to it as distance, keeping in mind that it can be a more general function.

The nodes in this model are chosen randomly and independently from a pre-metric space. Their common probability distribution, however, does not have to be uniform at all. In principle, it can be arbitrary. A totally arbitrary distribution over an abstract set, however, can have pathological mathematical properties. Therefore, to avoid unnecessary technical difficulties, it is assumed that all considered sets are measurable with respect to the probability measure and all considered expected values exist. Otherwise, the distribution is arbitrary. Note that while it is mathematically possible to create pathological examples that do not satisfy these conventions, they never occur in practically meaningful cases.

Let us now define the *pre-metric random graph model.* In the definition $\mathbf{N}$ and $\mathbf{R}_+$ denote the set of natural numbers and positive reals, respectively.

Definition 2 (premetric random graph model). The model is given by a triple $\mathcal{G} = (\mathcal{M}, r, P)$, where $\mathcal{M} = (S, \rho)$ is a premetric space, $r : \mathbf{N} \mapsto \mathbf{R}_+$ is a function, and P is a probability measure on $\mathcal{M}$. A random graph is generated by the model in the following way. For some $n \in \mathbf{N}$ random points $X_1, \ldots, X_n$ are drawn from the pre-metric space independently at random, according to the probability measure P. These points are the nodes of the graph. The edges of the graph are defined by connecting any two different points X_i, X_j if and only if

$$\rho\left(X_i, X_j\right) \leq r(n). \tag{13}$$

In the above definition, the function $r(n)$ represents the transmission radius, which may depend on the number of nodes for normalization purposes, as it is typical in most models. The arising grap intends to capture the network topology in a general setting. The interesting thing is that even in this general model, one can prove highly nontrivial properties. Since the details are rather technical, we skip them here and refer to the article Faragó [8].

4.2. Abstract Geometric Random Graphs. We have analyzed an even more general model in [6, 10]. Since it allows proving surprising results, we present it in more detail.

4.2.1. Starting with the Most General Setting. In order to build up our modeling approach, let us first explain what we mean by random graphs and a random graph model in the possibly most general sense.

In full generality, by a *random graph* on a fixed number of nodes (n), we mean a random variable that takes its values in the set of all undirected graphs on n nodes. We are going to denote by G_n a random graph on n nodes. At this point, it is still completely general, possibly generated by any mechanism, with arbitrary dependencies among its parts, it is just *any* graph-valued random variable, taking its values among undirected graphs on n nodes.

Having introduced general random graphs, a *random graph model* is given by a sequence of graph-valued random variables, one for each possible value of n:

$$\mathcal{M} = (G_n;\ n \in \mathbf{N}). \tag{14}$$

Next, we introduce some general features that apply to any random graph model.

4.2.2. Degrees and Connectivity. Let G_n be any random graph on n nodes and let us denote by $e(G_n)$ the number of edges in the graph. We characterize the degrees of G_n by the expected degree of a randomly chosen vertex, which we call he *expected average degree* of G_n. It is denoted by $\overline{d}(n)$ and defined by

$$\overline{d}(n) = \frac{2E(e(G_n))}{n}. \tag{15}$$

It is based on the fact that the actual average degree in any graph G on n nodes is $2e(G)/n$. Often the expected degree of each individual node is also equal to $\overline{d}(n)$ but in a general model it may not hold. (Note that even if the expected degree of each node is equal to the expected average degree, it does not mean that the *actual* random degrees are also equal.)

Ideally, we would like a random graph model, in which $\overline{d}(n)$ remains bounded by a constant, and the model is *asymptotically almost surely (a.a.s.)* connected, meaning

$$\lim_{n\to\infty} \Pr(G_n \text{ is connected}) = 1. \tag{16}$$

Note: whenever we write down a limit, such as the one above, we also assume that the limit exists.

Since asymptotic connectivity in most models requires unbounded degrees, therefore, one may hope that if we accept less than full connectivity, then there is a better chance to keep the node degrees bounded. To this end, let us define a weaker version of connectivity.

Definition 3 (β-connectivity). For a real number $0 \le \beta \le 1$, a graph G on n nodes is called β-connected if G contains a connected component on at least βn nodes.

When we consider a sequence of graphs with different values of n, then the parameter β may depend on n. When this is the case, we write β_n-connectivity. Note that even if $\beta_n \to 1$, this is still weaker than full connectivity in the limit. For example, if $\beta_n = 1 - 1/\sqrt{n}$, then we have $\beta_n \to 1$, but each graph on n nodes can still have $n - \beta_n n = \sqrt{n}$ nodes that are not part of the largest connected component.

4.2.3. Building the Model. Let us now introduce a very general model class that reflects a typical feature of geometric random graph models. This feature is that in geometric random graphs, the primary random choice is picking random nodes from some domain, and then the edges are already determined by some geometric property (typically some kind of distance) of the random nodes. We elevate this approach to an abstract level that includes many special cases of interest. The most general version of our abstract geometric model is built using the components detailed below.

Representing the Nodes: Node Variables. The nodes are represented by an infinite sequence $X_1, X_2, \ldots$ of random variables, called *node variables*. They take their values in an arbitrary (nonempty) set S, which is called the *domain* of the model. In most practical cases, the domain is a simple subset of the Euclidean plane or of the 3-dimensional space. In general, however, S can be any abstract set from which we can choose random elements (to avoid mathematical complications that would only obscure the main message, we assume that all considered sets, functions, etc., are measurable with respect to the used probability measures and all considered expected values exist. This is satisfied in every practically relevant model). When we want to generate a random graph on n nodes, then we use the first n entries of the sequence, that is, $X_1, \ldots, X_n$ represent the nodes in G_n. It is important to note that we do not require the node variables to be independent.

Representing the Links: Edge Functions. We denote by $Y_{ij}^{(n)} \in \{0, 1\}$ the indicator of the edge between nodes X_i, X_j in the random graph G_n. Since loops are not allowed, we always assume $i \neq j$, without repeating this condition each time. The (abstract) geometric nature of the model is expressed by the requirement that the random variables $Y_{ij}^{(n)}$ are determined by the nodes $X_1, \ldots, X_n$, possibly with additional independent randomization. Specifically, we assume that there exist functions $f_{ij}^{(n)}$, such that

$$Y_{ij}^{(n)} = f_{ij}^{(n)}\left(X_1, \ldots, X_n, \xi_{ij}\right), \tag{17}$$

where ξ_{ij} is a random variable that is uniformly distributed on $[0, 1]$ and is independent of all the other defining random variables of the model (i.e., the node variables and all the other ξ_{kl} variables). Henceforth, the role of ξ_{ij} is referred to as *independent randomization* (note that the specified distribution of ξ_{ij} does not impose a restriction, since the functions $f_{ij}^{(n)}$ are arbitrary). The undirected nature of the graph is expressed by the requirement $Y_{ij}^{(n)} = Y_{ji}^{(n)}$, which can simply be enforced by computing all values for $i < j$ only and defining the $i > j$ case by exchanging i and j.

4.3. Specific Classes within Abstract Geometric Random Graphs. Before turning to results, let us present some examples to show the usefulness and comprehensiveness of the generalization provided by our abstract geometric random graphs. These examples illustrate that most practically relevant models for wireless network topologies fit in the common generalization that we provided by introducing abstract geometric random graphs.

Geometric Random Graphs. All the usual geometric random graph models fit naturally in our general framework. For example, the base set S can be chosen as a unit disk or square in the plane or a unit ball or cube (or any other domain) in higher dimension. Let us choose i.i.d. points $X_1, X_2, \ldots$ from S, according to some probability distribution. Let $\rho(x, y)$ denote the distance of the points $x, y \in S$, it can be any distance function. Finally, let $r > 0$ be a radius (possibly depending on n). Then the edge function

$$f^{(n)}\left(X_i, X_j, \xi_{ij}\right) = \begin{cases} 1 & \text{if } \rho\left(X_i, X_j\right) \le r \\ 0 & \text{if } \rho\left(X_i, X_j\right) > r \end{cases} \tag{18}$$

defines a geometric random graph in the usual sense. (The independent randomization is not used here, so the edge function does not depend on ξ_{ij}.) It is clear that this includes all the usual geometric random graph models, allowing any metric space as the basis. Moreover, we can also use non-independent points, such as the "clustered uniform" example in the previous section, as long as the distribution is exchangeable.

Erdős-Rényi Random Graphs. The by now classical random graph model of Erdős and Rényi (see, e.g., [80, 81]), where each possible edge is included independently with some probability, p is also included as a direct special case. We can set $S = \{1, \ldots, n\}$ and for $X_i, X_j \in S$

$$f^{(n)}\left(X_i, X_j, \xi_{ij}\right) = \begin{cases} 1 & \text{if } \xi_{ij} \le p \\ 0 & \text{if } \xi_{ij} > p. \end{cases} \tag{19}$$

Note that now the edge function depends only on the independent randomization, so indeed each edge is included independently with probability p.

A Geometric but Nonmetric Example: Battery Levels. In geometric random graph models, ρ satisfies the triangle inequality. This, however, cannot capture all situations that occur in ad hoc or sensor networks. As an example, assume the nodes are located in the plane. Let x_i, y_i be the coordinates of the ith node. Furthermore, we also characterize a node with its battery level $E_i > 0$. E_i represents the remaining energy, assuming the node is not fully out of energy. Thus, a node is represented by a triple $X_i = (x_i, y_i, E_i)$. Let $d(E_i)$ be the distance over which a node can communicate, given its energy level E_i. (The function $d(E_i)$ can be derived from the physical characteristics of the node and from radio propagation conditions.) Now, a possible example of such a "distance" function is

$$\rho_1\left(X_i, X_j\right) = \frac{\sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}}{\min\left\{d(E_i), d\left(E_j\right)\right\}}. \tag{20}$$

If we take $r = 1$ and use the above ρ_1 function in (18), then it expresses the condition that a link exists if and only if its end nodes are at most at a distance that can be bridged by the energy levels of both nodes. Note that the above function ρ does not satisfy the triangle inequality, so it does not lead to a geometric random graph model in the usual sense. On the other hand, it still fits in our framework, as in (18) we did not require the triangle inequality to hold for ρ.

Another Non-Metric Example: Link Blocking. We can capture some features of traffic-dependent network characteristics as well. Let each node i be characterized by a triple $X_i = (x_i, y_i, \lambda_i)$, where x_i, y_i are planar coordinates and λ_i is the traffic demand of the node. Let B_{ij} be the blocking probability of the link (i, j), given that the link exists. We may compute B_{ij} as a function of λ_i, λ_j from some traffic models. For example, if we use Erlang's well-known formula, assuming a capacity of C units on the link and its load is taken as the sum of its end nodes' traffic load $\lambda_i + \lambda_j$, then we obtain

$$B_{ij} = \frac{\left(\lambda_i + \lambda_j\right)^C / C!}{\sum_{i=0}^{C} \left(\lambda_i + \lambda_j\right)^i / i!}. \tag{21}$$

(Of course, we may use other traffic models, as well; this is just an example.) Now we can take the "distance" function

$$\rho_2\left(X_i, X_j\right) = \frac{1}{1 - B_{ij}} \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2} \tag{22}$$

and use it in (18) with some radius r. We can observe that for small blocking probability ($B_{ij} \ll 1$) $\rho_2(X_i, X_j)$ will be approximately the same as the Euclidean distance. On the other hand, as B_{ij} approaches 1, the factor $1/(1 - B_{ij})$ tends to infinity, and, therefore, high blocking probability makes the existence of the link in the model less likely, even if the physical distance is small. This example also violates the triangle inequality, so it is not a geometric random graph.

Log-Normal Shadowing. A typical phenomenon in the radio environment is *fading*. An example of fading is a relatively slow random fluctuation in the signal strength, which occurs even if the locations are fixed. Measurements show that this random variation can be accurately modeled by a log-normal distribution (see, e.g., [82]). Hence, the name *log-normal shadowing*, is widely used for this phenomenon. A way to capture it in our model is this. Let us characterize a node i by a triple $X_i = (x_i, y_i, \eta_i)$, where x_i, y_i represent a random position in the plane, and each η_i is an infinite sequence of independent, log-normally distributed random variables:

$$\eta_i = \left(\eta_j^{(i)};\ j = i, i+1, i+2, \ldots\right). \tag{23}$$

The "distance" is defined as

$$\rho_3\left(X_i, X_j\right) = \eta_b^{(a)} \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}, \quad (24)$$

where $a = \min\{i, j\}$ and $b = \max\{i, j\}$. (The reason for we need an infinite sequence of log-normal random variables is that this way we can have independent log-normal shadowing for every link.) This distance can express the fact that from the radio communication point of view, we really perceive an "effective distance", which is a log normally modulated random variant of the physical distance. Using this ρ_3 in (18) leads again to a random graph that is not geometric, as ρ does not satisfy the distance axioms.

Directional Antennas. We can also represent directional antennas in the model. As a simple example, let Y_i be the position of a node in the Euclidean plane, α_i be the angle (with respect to some fixed coordinate axis) at which its antenna is directed, and δ_i be the angular width of the beam (assuming an idealized directional antenna). Let us represent the node by the variable $X_i = (Y_i, \alpha_i, \delta_i)$. Let $S(X, \alpha, \delta)$ denote the planar angular sector pointed at X, with its axis of symmetry directed at α and of angular width δ. Further, let $\|\cdot\|$ denote the Euclidean norm. Then we can introduce the following "distance":

$$\rho_4\left(X_i, X_j\right) = \begin{cases} \left\|Y_i - Y_j\right\| & \text{if } X_i \in S\left(X_j, \alpha_j, \delta_j\right), \\ & X_j \in S(X_i, \alpha_i, \delta_i), \\ \infty & \text{otherwise.} \end{cases} \quad (25)$$

If we use this function $\rho_4(X_i, X_j)$ in (18), then we get a model of a random ad hoc network topology with directional antennas.

Terrain Variations, Obstacles. Another example is to take into account uneven radio propagation characteristics due to terrain variations or propagation obstacles. For example, let us assume that the nodes operate in a frequency band in which only line of sight communication is possible (such as infrared). Then two nodes can only communicate if there is no obstacle covering them from each other. This feature can also be built into the model. Let X_i be the plane position of a node. Assume there exists a set $\mathcal{R} = \{R_1, R_2, \ldots\}$ of random obstacles in the area. Let $s(X_i, X_j)$ be the line segment connecting the points X_i, X_j, and let $L(X_i, X_j, \mathcal{R})$ be the "line of sight" function:

$$L\left(X_i, X_j, \mathcal{R}\right) = \begin{cases} 1 & \text{if } s\left(X_i, X_j\right) \cap R_k = \varnothing \ (\forall k) \\ \infty & \text{otherwise.} \end{cases} \quad (26)$$

To express that only those nodes can communicate that are in line of sight of each other, let us introduce the "distance"

$$\rho_5\left(X_i, X_j\right) = \left\|X_i - X_j\right\| L\left(X_i, X_j, \mathcal{R}\right). \quad (27)$$

If this is used in (18), then we get a network topology model that can deal with radio propagation obstacles.

Combinations. The various conditions in the preceding examples can be combined into more complex models. For example, if we want that all the conditions expressed by the $\rho_1, \ldots, \rho_5$ functions are satisfied, then we can use

$$\rho\left(X_i, X_j\right) = \max\left\{\rho_1\left(X_i, X_j\right), \ldots, \rho_5\left(X_i, X_j\right)\right\} \quad (28)$$

in (18).

5. Results on Abstract Geometric Random Graphs

5.1. Restrictions. Regarding the abstract geometric random graph model in the presented very general form, it is clear that allowing *totally arbitrary* node variables and edge functions offers little hope for meaningful analysis. Therefore, next we introduce some restricting conditions. Later we are going to see that one has to make only surprisingly mild restrictions to meaningfully analyze the trade-off between node degrees and β-connectivity.

Locality. Up to now, we allowed that an edge in G_n can depend on all the nodes, and the dependence expressed by the $f_{ij}^{(n)}$ functions can be arbitrary and different for each edge. To get a little closer to the usual geometric random graph model, let us introduce the following property, called *locality.* Informally, it restricts the dependence of an edge to its endpoints, in a homogeneous way but still via an *arbitrary* function.

Definition 4 (locality). An abstract geometric random graph model is called local, if for every n and $i, j \le n$ the existence of an edge between X_i, X_j depends only on these nodes. Moreover, the dependence is the same for every i, j, possibly with independent randomization. That is, there are functions $f^{(n)}$ such that the edge indicators are expressible as

$$Y_{ij}^{(n)} = f^{(n)}\left(X_i, X_j, \xi_{ij}\right), \quad (29)$$

where ξ_{ij} represents the independent randomization.

Name Invariance. Our second condition called *name invariance* refers to the joint distribution of nodes. If we allow totally arbitrary joint distribution, then it offers little chance for meaningful analysis. On the other hand, restricting ourselves only to independent, identically distributed (i.i.d.) node variables would exclude important cases, such as clustering. For the aforementioned reasons, we use a more general condition, rather than assuming independent, identically distributed node variables. To introduce it, let us first recall a useful concept from probability theory, called exchangeability.

Definition 5 (exchangeable random variables). A sequence of random variables is called exchangeable if for any $k \ge 1$, it holds that if we select any k of the random variables, the joint distribution of the selected random variables depends only on k, but is independent of which particular k variables are selected, and in which order.

Note that i.i.d. random variables are always exchangeable, but the converse generally does not hold, so exchangeable random variables form a larger family.

Now let us introduce the condition that we use to restrict the arbitrary dependence of node variables.

Definition 6 (name invariance). An abstract geometric random graph model is called name invariant, if its node variables are exchangeable.

We call it the *name invariance* of the model because it means the names (the indices) of the nodes are irrelevant in the sense that the joint probabilistic behavior of any fixed number of nodes is invariant to renaming (reindexing) the nodes. In particular, it also implies that the individual node variables are identically distributed, but they do not have to be independent.

Name invariance is naturally satisfied with the most frequently used random node choices, such as uniform independent random points in a planar domain, a Poisson point process in the plane, or in higher dimension. We allow, however, much more complex node generation (over an arbitrary set!) since dependencies are not excluded by name invariance.

A simple example for a dependent, yet still name invariant, node generation process is a "clustered uniform" node generation. As an example, let S be a sphere in 3-dimensional space, that is, the surface of a 3-dimensional ball. Let R be the radius of the ball. Let us first generate a pivot point Y uniformly at random from S. Then generate the nodes $X_1, X_2, \ldots$ uniformly at random and independently of each other from the neighborhood of radius $r \ll R$ of the random pivot point Y (on the sphere). It is directly implied by the construction that exchangeability holds. Moreover, any particular X_i will be uniformly distributed over the *entire* sphere, since Y is uniform over the sphere. On the other hand, the X_i are far from independent of each other, since they cluster around Y, forcing any two of them to be within distance $2r$. The setting can be generalized to applying several pivot points and nonuniform distributions, creating a more sophisticated clustering.

5.2. Threshold Function for Partial Connectivity. We define a concept that will characterize the trade-off between node degrees and the type of partial connectivity that we introduced as β-connectivity in Definition 3. For notational convenience, the set of nonnegative real numbers, extended with ∞, will be denoted by $\mathbf{R}_0^\infty$. Real functions are also extended to ∞ by $f(\infty) = \lim_{x\to\infty} f(x)$, whenever the limit exists (it will always exist in our cases). The value of β is always assumed to be in $[0, 1]$.

Let us first explain the threshold function concept informally. We define a threshold for β-connectivity, such that whenever β is above the threshold, then it is impossible to achieve a.a.s. β-connectivity for any model in the considered family of random graph models. On the other hand, if β is below the threshold, then this is not the case anymore, that is, there is at least one model in the family that achieves a.a.s β-connectivity with this β. Thus, the threshold separates the cases when a.a.s. β-connectivity is impossible, from the cases when it is possible. Since the threshold will depend on the expected average degree, we call it threshold function.

Now let us present the formal definition. Recall that the expected average degree in a random graph G_n is defined as $\overline{d}(n) = 2E(e(G_n))/n$.

Definition 7 (threshold for β-connectivity). Let $\mathcal{F}$ be a family of random graph models. For any model $\mathcal{M} \in \mathcal{F}$ let G_n denote the random graph on n nodes generated by $\mathcal{M}$ and let

$$D_{\mathcal{M}} = \limsup_{n\to\infty} \overline{d}(n) \tag{30}$$

be the limiting expected average degree. A function $f : \mathbf{R}_0^\infty \mapsto [0, 1]$ is called a β-connectivity threshold function for $\mathcal{F}$ if the following two conditions are satisfied.

(i) For any model $\mathcal{M} \in \mathcal{F}$ and for every $\beta > f(D_{\mathcal{M}})$

$$\lim_{n\to\infty} \Pr(G_n \text{ is } \beta\text{-connected}) < 1 \tag{31}$$

holds, where G_n is generated by $\mathcal{M}$.

(ii) If β is below the threshold, then (i) does not hold anymore, in the following sense. For every $\epsilon > 0$, there exists a model $\mathcal{M}_0 \in \mathcal{F}$ and a

$$\beta \le f(D_{\mathcal{M}_0}) - \epsilon, \tag{32}$$

such that

$$\lim_{n\to\infty} \Pr(G_n \textit{ is } \beta\text{-connected}) = 1, \tag{33}$$

where G_n is generated from $\mathcal{M}_0$.

The importance of this concept is the following. If for a considered class $\mathcal{F}$ of random graph models we can find out what the corresponding β-connectivity threshold function is, then we can tell precisely what range of expected average degrees allow a.a.s. β-connectivity for a given β. Or, conversely, if we know the (asymptotic) expected average degree for a particular model $\mathcal{M}$ in the considered class, then we can decide what level of connectivity can be asymptotically achieved for this model.

5.3. Computing the Threshold. Now we state the theorem that conveys the surprising message that for the very general class of abstract geometric random graph models, we can still find the *precise* β-connectivity threshold function, if we assume that the models satisfy the conditions of locality and name invariance. The previously presented examples all satisfy these conditions, so they show that even with these restrictions, we can still include many complex and practically important models. For the proof of the theorem, see [7, 10].

Theorem 8 (threshold function for local and name invariant abstract geometric graphs). *Let $\mathcal{F}$ be the family of local and name invariant abstract geometric random graph models. For any model $\mathcal{M} \in \mathcal{F}$ set*

$$D_{\mathcal{M}} = \limsup_{n\to\infty} \overline{d}(n). \tag{34}$$

Then the β-connectivity threshold function for $\mathcal{F}$ is

$$f(D_{\mathcal{M}}) = 1 - e^{-D_{\mathcal{M}}}. \quad (35)$$

5.4. Consequences for Full Connectivity. It is worth mentioning that the definition of the threshold function and Theorem 8 directly imply that bounded expected average degrees in $\mathcal{F}$ exclude a.a.s. β_n-connectivity when $\beta_n \to 1$. As a result, a.a.s. full connectivity, which corresponds to $\beta = 1$, is also excluded. These claims are formally stated below, the proof is a direct application of Theorem 8.

Theorem 9. *Let $\beta_n \to 1$ be an arbitrary sequence in $[0,1]$. Then for any local and name invariant abstract geometric random graph model $\mathcal{M}$, it holds that if $D_{\mathcal{M}} < \infty$, then the random graphs generated by $\mathcal{M}$ cannot be a.a.s. β_n-connected.*

The interpretation of this result is that (asymptotically) the requirements of full connectivity and bounded degrees are incompatible, in the broad class of models we have considered.

At this point, one may wonder whether there is *any* meaningful random graph model at all, in which a.a.s. full connectivity is possible, yet the node degrees remain bounded. Note that our results do not exclude this, since they only apply to local and name invariant abstract geometric random graph models. Although this class is quite comprehensive, it does not contain *all* meaningful models.

A nontrivial example worth mentioning here is the (random) Euclidean minimum spanning tree (MST). Let us choose n i.i.d. random points in the d-dimensional unit cube and view them as vertices of a complete graph, where each edge is assigned a weight that is equal to the (random) distance of its endpoints. Let T_n be the MST of this graph. Note that T_n is unique with probability 1. It is clear that T_n is connected, as, by definition, it is a spanning tree. Moreover, the following nontrivial fact is known: for every fixed dimension the maximum degree of the Euclidean MST is bounded by a constant, depending only on the dimension, but not on n (see, e.g., [83]). Thus, the model $\mathcal{M} = (T_n;\ n \in \mathbf{N})$ has the property that it is fully connected, yet its node degrees remain bounded.

It is clear that the Euclidean MST model is name invariant, since nothing depends on how the nodes are indexed. Does it then contradict to our results? No, because it does not satisfy the requirement of locality. Of course, the usual definition of the MST is indeed not local. But now our results imply that the nonlocality is *unavoidable* in this case, as long as we want to preserve name invariance. In other words, it is impossible to define the Euclidean MST in a local way, such that, at the same time, the model is also name invariant.

Note that the fact that the Euclidean MST cannot be defined locally, with name invariance, is nontrivial. For example, one might try to define new node variables that contain enough information to decide for any pair whether an MST edge connects them, without looking at other nodes. A possibility is to introduce new node variables $Y_i = (X_i, \dots, X_n, X_1, \dots, X_{i-1})$, with edges that connect two such new nodes if their first components are connected by an MST edge, among the original X_i variables. In this way, we can create a locally defined MST model, since one can decide from Y_i, Y_j alone, whether X_i, X_j are connected by an MST edge, as the information about *all* the original nodes is available in each of the new node variables. Thus, in the transformed domain, we have a *local* model. The MST over the Y_i variables will be isomorphic (with probability 1) to the MST over the X_i node variables, so the new model is equivalent to the original, yet it is local. Then, however, the name invariance would be destroyed. Even though each Y_i individually has the same distribution (since it does not matter in what order the X_i are listed), the joint distribution of Y_1 and Y_2 will not be the same as the joint distribution of Y_1 and Y_3. The reason is that the first coordinate of $Y_1 = (X_1, \dots, X_n)$ is the same as the last coordinate of $Y_2 = (X_2, \dots, X_n, X_1)$, but such a relationship does not hold between Y_1 and $Y_3 = (X_3, \dots, X_n, X_1, X_2)$.

Generally, it follows from our results and from the aforementioned properties of the Euclidean MST that no matter how tricky local definition we invent for this random graph model, it cannot preserve name invariance. The fact that name invariance *excludes* the possibility of a local Euclidean MST definition appears to be hard to prove without our results.

6. Physical Models and Their Relationship to Graph Models

6.1. The SINR Concept. The *Signal to Interference and Noise Ratio (SINR)* model (see [84], and further references therein) captures the physical conditions of receiving the radio signal with a satisfactory quality.

Let $v_1, \dots, v_n$ be radio nodes and, with a slight abuse of notation, let each v_i also represent the position of the node (in the plane, for simplicity). Let $d(x, y)$ be the Euclidean distance and assume that node v_i transmits with power P_i. Then, the *reception zone* of node v_i consists of all those points x in the plane, where $\mathrm{SINR}_i(x) \geq \beta$ holds with some parameter β, where

$$\mathrm{SINR}_i(x) = \frac{P_i d(v_i, x)^{-\alpha}}{N + \sum_{j \neq i} P_j d\left(v_j, x\right)^{-\alpha}}. \quad (36)$$

In the simplest case, the parameters β, N, and α are assumed known constants; α is called *path loss exponent* and it usually falls in the range $2 \leq \alpha \leq 6$. The meaning of the $\mathrm{SINR}_i(x) \geq \beta$ condition is that the received power of v_i at x is at least β times larger than a noise power N, plus the interference from all other nodes at x, assuming that attenuation is proportional to a power of distance.

Now, one may ask the natural question: what kind of network topology arises from this model? We can define it by connecting any two nodes v_i, v_k whenever both $\mathrm{SINR}_i(v_k) \geq \beta$ and $\mathrm{SINR}_k(v_i) \geq \beta$ hold, that is, they are mutually in each other's reception zone, so they can communicate. This generates a graph (network topology) that describes which pairs of nodes are capable of communicating.

The arising (undirected) graph, however, does not seem to belong to any special class with nice properties, since

the existence of any given edge depends, in a complicated, nonlinear way, on *all* the node positions and transmission powers, not only on the end nodes of the particular link. For example, if any single node v_i changes position or power, then all $\mathrm{SINR}_i(x)$ values change, so the *whole* graph may become different. The effect may not remain local, in sharp contrast with the philosophy of graph models. Thus, the SINR approach does not seem to offer any good opportunity to apply the special graph classes and properties, along with the rich treasury of results that build on them. Next, however, we show the surprising and unexpected fact that it is still possible to build bridges between the "different worlds."

6.2. A Bridge between SINR and Graph Models. Building on the involved analysis of [84] about the geometry that the SINR condition (36) generates, we can prove a result (see [20]), which shows the unexpected fact that the "messy" SINR based topology is in fact not too far from a simpler, standard graph model, at least for some choice of the parameters.

Theorem 10. *Assume that all nodes transmit with the same power and the path loss exponent is $\alpha = 2$. Then, for arbitrary $\beta \geq 1$ and for arbitrary node positions, the network topology that arises from the SINR model is a Quasi-Disk Graph with shrink factor*

$$\rho = \frac{\sqrt{\beta} - 1}{\sqrt{\beta} + 1}. \tag{37}$$

This theorem shows the unexpected fact that despite the messy "everything depends on everything" nature of the physical SINR model, the arising graph still belongs to a special class that is part of the toolkit of multihop wireless network modeling.

6.3. Trade-Off between SINR and Global Properties of the Network Topology. Equation (37) implies that with $\beta \to \infty$ the Quasi-Disk Graph in Theorem 10 approaches a Disk Graph, since ρ tends to 1. Depending on the node positions, however, it may not be a Unit Disk Graph (UDG), only a Disk Graph (DG). On the other hand, if the node positions are random (chosen independently from the same distribution), then, by symmetry considerations, each node will have the same *expected* radius r. Due to independence, one can also expect a strong concentration of the actual random r_i values around their common expected value r. Taking into account that with large enough β, the shrink factor will be $\rho \approx 1$, we obtain that, under these conditions, the SINR generated network topology can asymptotically be well approximated by a UDG.

This fact is encouraging from the viewpoint of analysis of network level properties, such as the connectivity of the network topology, since much is known about these issues in the UDG model. Additionally, in the random setting, it is known that if n nodes are placed uniformly at random in a unit disk, then the network will be connected with probability approaching 1, as $n \to \infty$, if and only if the transmission radius r satisfies

$$\pi r^2 = \frac{\log n + c(n)}{n}, \tag{38}$$

where $c(n)$ is an arbitrary function with $c(n) \to \infty$ [61]. Various generalizations are also known in this direction, see, for example, [6, 8]. It is also important that both DGs and UDGs have efficient routing and labeling schemes, as well as spanners and separators, and so forth, with good properties (see [85] for definitions and further references).

These considerations point to the unexpected fact that the "messy" SINR generated network topology and the abstract UDG model are in fact not as far from each other as one might expect. This may allow the extension of a lot of useful analytical and protocol design results from UDGs to SINR based models.

6.4. A Common Generalization of SINR and Graph Models. The SINR model, in particular, its special case used in Theorem 10, is only the simplest version of this type of physical models. In reality, there are a good number of additional complications that need to be taken into account for a faithful representation of the radio environment. For example, the transmission power of each node may be different and randomly changing. The received power may be subject to fading. The noise level may be time and location dependent and randomly changing. The path loss exponent may be different from $\alpha = 2$; it may also depend on direction, time, and location. The distance may not be Euclidean, to reflect additional effects, such as obstacles to radio propagation.

The formal introduction of such effects into the SINR model, in order to making it a more faithful description of the radio network, is not too hard in itself, since one can use the results of the extensive research that has been done in accurately modeling the physical radio environment. It is much more difficult, however, to avoid hopeless messiness from the viewpoint of networking protocols. In other words, it is very desirable to develop the more complex models in a way that still preserves a meaningful relationship with graph models, so that we can build further bridges in the spirit discussed previously.

In a more complex scenario, it seems unlikely that Theorem 10 would directly carry over. Therefore, let us intro-duce a more general graph model that opens a new direction to capture the network topology, and, in a sense, serves as a common generalization of SINR and graph models. We call it *Weight Ratio Graph (WRG)* model. It is a graph characterized by the following parameters.

(i) A set of nodes $\{v_1, \ldots, v_n\}$.

(ii) A positive weight $w_{i,j}$ for each pair (v_i, v_j) of nodes $(i \neq j)$.

(iii) A parameter $\beta > 0$, which will play a similar role to the β parameter of the SINR model.

With these parameters the actual graph is defined in the following way. First, we say that node v_i can receive node v_k's transmission if

$$\frac{w_{k,i}}{\sum_{j\neq k} w_{j,i}} \geq \beta, \tag{39}$$

that is, the weight $w_{k,i}$, perceived as the power that reaches v_i from v_k, is at least β times larger than the power that reaches v_i from all other nodes. Finally, there is a link between two nodes if they can both receive each other's transmission, according to the above definition.

It is not hard to see the connection between the SINR and WRG models. If we take $w_{i,j} = P_i d(v_i, v_j)^{-\alpha}$ and use the same β in both models, then WRG will produce exactly the same network topology as the SINR with $N = 0$. Therefore, at first it seems that WRG can only represent the special case when the noise is zero. A closer look reveals, however, that this is not true. In fact, surprisingly, the WRG model essentially has *universal* expressive power, as shown by the following result [20].

Theorem 11. *Let G be an arbitrary graph with no isolated nodes. Then the parameters of the WRG model can always be chosen such that the model will generate precisely the graph G.*

As we have seen, the WRG model can generate *any* graph with no isolated nodes, which, of course, also includes the SINR topology with nonzero noise, assuming there is no isolated node. As a result, we obtain the following direct, but interesting consequence.

Theorem 12. *For an SINR model, it is always possible to eliminate the effect of noise by transforming it into another SINR model with $N = 0$ that generates the same network topology, given that it contains no isolated node. The transformation can be achieved by adjusting the transmission powers from P_i to some P_i' and possibly the value of β, with no other change in the model. That is, some positive P_i', β' values can always be chosen such that*

$$\frac{P_i d(v_i, v_k)^{-\alpha}}{N + \sum_{j\neq i} P_j d\left(v_j, v_k\right)^{-\alpha}} \geq \beta \tag{40}$$

holds if and only if

$$\frac{P_i' d(v_i, v_k)^{-\alpha}}{\sum_{j\neq i} P_i' d\left(v_j, v_k\right)^{-\alpha}} \geq \beta'. \tag{41}$$

Moreover, this equivalence holds for all i, k simultaneously.

It is worth noting that while Theorem 12 follows Theorem 11 as an immediate consequence, it appears to be harder to prove Theorem 12 *directly*, building only on the SINR model. The achieved shortcut shows the power of Theorem 11.

7. Networks with Multiradio Nodes

The growing importance of wireless networks is a major and lasting trend in the networking landscape. It also comes hand in hand, however, with the increasing *diversity* of wireless networking solutions and standards. Some of them are already widely used, such as variants of IEEE 802.11 wireless LANs (802.11a, 802.11b, and 802.11g), others are emerging, such as IEEE 802.15 personal area networks, Zigbee, Bluetooth, broadband wireless (IEEE 802.16), and a variety of sensor networking solutions.

Another well visible trend is that radio interfaces are rapidly getting more and more inexpensive and physically small. It is now technically and economically quite feasible to equip wireless network nodes with several radio transmitters/receivers. This creates an environment, where the network effectively has *multiple physical layers*. For example, it is already quite common that a laptop or a PDA has both an IEEE 802.11 card and a Bluetooth interface. Given the tendency of decreasing price and shrinking physical size, it is quite likely that the trend of having multiple radio interfaces will get even more prevalent in the future. Thus, many of the ubiquitous wireless network nodes will likely be capable of operating with multiple physical layers. A few examples, present and future.

(i) Multiradio nodes (e.g., laptops with multiple radio interfaces, as mentioned above).

(ii) Multichannel radio environments that logically act as multiple radios. Note that the IEEE 802.11 standard defines multiple channels that are only partly utilized today.

(iii) Multiple antenna systems that implement several independent channels via sophisticated physical layer techniques, such as beamforming, Space Division Multiple Access (SDMA), cooperative coding, and multiple-input/multiple-output (MIMO) systems.

(iv) Combination of radio and infrared interfaces in a node.

(v) Utilization of low-power wireless technologies, such as RFID solutions with energy harvesting and various wireless sensor platforms.

(vi) Combinations of radio and other potentially possible wireless transmission technologies, such as free space optical transmission and laser beams.

Thus, while the technical *possibility* of multiple physical layers is already quite clear today, it is much less obvious how can it be efficiently *utilized* to gain significant improvement in the overall network performance. Or, from the practical/economical point of view, the ultimate question is: Will the multiradio network development lead to sufficient performance improvement that justifies the investment?

We present a general network topology model to analyze multiradio networks. The model is based on the mathematical concept of *multigraphs*. First, however, let us do a short literature review about what is available regarding the multiradio environment.

The availability of multiple radio interfaces and multiple channels per radio in multihop networks have raised significant interest in the academic and industrial community due to the design and deployment of *mesh wireless networks*,

that is, multihop networks where some of the nodes, usually less mobile, provide backbone connectivity to mobile end users [86, 87]. As a consequence, whether related directly to mesh networking or just considering generic multiradio systems, a good number of papers have been published at networking conferences in this area. Apparently, all works on multiradio/multichannel systems suggest and demonstrate the advantage of exploiting the radio diversity provided by multichannel/multiradio systems, and in general, of having multiple radio interfaces for considerable performance improvement.

Various solutions have been proposed for multirate (but singleradio) systems. Some technologies, such as IEEE 802.11a/b/g, have multiple transmission rates, so that they can accommodate multiple channel conditions, enhance throughput, and decrease interference losses. For example, in the work by Awerbuch et al. [88] it is shown that the minimum-hop routing strategy, which is usually deployed in single rate networks, is suboptimal in multirate networks. Therefore, the authors propose a different link selection metric (Medium Time Metric, MTM) that favors shorter, higher throughput, and more reliable links instead of the longer links that regular routing would select. Interference-aware topology control and routing is discussed in [89], where single-radio (IEEE 802.11b) nodes are considered. By taking advantage of the multiple channels of the radio, the authors define a way of choosing channels so that the resulting topology is "interference-minimum" among all possible k-connected topologies. For splitable traffic demands, the authors then present the Bandwidth-Aware Routing (BAR) protocol, which is demonstrated to outperform routing on a single-channel topology. Chandra et al. [90] advocates the use of single-radio systems in a multi-LAN environment. A wireless device could connect to multiple wireless LANs simultaneously despite having a single-radio card. The paper describes a software layer approach to this problem, called MultiNet, that *virtualizes* a single wireless card and facilitates parallel multiple connections.

The *capacity* of wireless multiradio, multichannel networks has been addressed, for example, by Kyasanur and Vaidya in [91]. In networks with n static nodes, each with m radios and c channels, $1 \leq m \leq c$, this work shows that the capacity of the network exhibits bounds that are different from those established by Gupta and Kumar [50] on single-radio/single-channel systems. In particular, the multiradio, multi channel bounds depend on the ratio between c and m. A characterization of the achievable rates in a mesh network with orthogonal channels is given in [92], where the authors determine necessary and sufficient conditions for achieving a rate vector in systems where the channels do not interfere with each other. By extending a former solution of theirs [93], the authors derive joint routing and scheduling algorithms for systems that are full duplex and equipped with multiple radios.

Routing for a multiradio network with static nodes (like a *wireless community network* [94]) has been presented in [95]. The authors propose a new metric for routing in networks with multiradio nodes, taking into account loss rate and link bandwidth and define a corresponding path metric that also considers interference among links using the same channel. The performance of the new metric is demonstrated via a 23-node testbed with nodes mounting two IEEE 802.11a wireless cards. The same testbed is also described in [96], where general design guidelines are provided for building multiradio systems. This work also revisits the standard problems in wireless networking in light of having nodes with multiple radio cards. The advantage of using multiradio nodes has also been demonstrated in [97], where a single-hop wireless system is presented, termed MultiRadio Diversity (MDR), which uses path diversity to improve loss resiliency. The paper demonstrates the multiradio advantage by showing throughput gains up to 2.3× over single-radio communication schemes (their testbed is based on nodes supporting multiple IEEE 802.11a wireless cards). Shin et al. [98, 99] advocate the use of multiradio to increase the capacity and performance of cellular networks. The idea in this case is to exploit the best available link to the base station. The authors envision an environment in which relay networks are dynamically formed so that whenever no acceptable direct links are available from a mobile node to the base station, a multihop path can be found. Algorithms for efficient formation of multihop relay networks are presented and evaluated for latency, signal overhead, gateway load (network formation), path length, and link sharing (relay network functions).

Channel assignment and routing for multihop wireless mesh networks is formulated mathematically as a joint optimization problem, by Alicherry et al. in [100]. The authors consider interference constraints, the number of channels per radio, and the number of radios per node. By solving the corresponding mixed integer linear programming model a centralized algorithm is proposed that optimizes the overall network throughput subject to fairness constraints. Distributed channel assignment in multihop, multiradio networks has been explored in [101]. The problem is shown to be *NP*-hard even if the paths between any two nodes are given. A randomized assignment scheme is then proposed (Skeleton Assisted partition FrEe or SAFE) that maintains network connectivity. A method for interference-aware cannel assignment has been proposed by Ramachandran et al. [102]. By modeling interferences between each mesh multiradio router via a newly defined *multiradio conflict graph*, a new channel assignment algorithm is designed that addresses explicitly the interference problem. Simulations and testing on a IEEE 802.11 testbeds are performed, which yield performance gains in excess of 40% with respect to static (non-interference aware) channel assignment solutions.

In order to ease the application access to multiradio cards on a single-node, the Multiradio Unification Protocol (MUP) has been proposed by Adya et al. [103]. The goal of MUP is coordinating the operations of the different (IEEE 802.11) cards tuned to nonoverlapping frequency channels by optimizing local spectrum usage. The advantage of using this kind of software interface has been demonstrated via ns2-based simulations. Results show that TCP throughput and user perceived latency significantly improve under dynamic traffic patterns over realistic topologies.

7.1. The Multigraph Model. Our model of the network topology for a network with multiple physical layers is a *multigraph*. As it is well known, a multigraph differs from an ordinary graph in that multiple edges can connect any pair of vertices. Naturally, the vertices represent network nodes, while the edges represent the links in the various physical layers, generated by the different radio interfaces.

As graphs have long been used to model network topologies, one may rightfully ask the question at this point: Can a multigraph lead to any *essential* new insight? In what follows we show that this model yields interesting and nontrivial novel problems.

Consider a wireless multihop network with omni-directional antennas and let its network topology be represented by a graph G_1. Assume that we equip the same nodes with another physical layer, such as a second radio with beamforming capabilities. This second physical layer generates another network topology, denoted by G_2. If we put them together, we obtain the topology of the combined system, which is a *multigraph* G.

Let us call the above merging operation of graphs the multigraph sum of G_1 and G_2. The operation will be denoted by $\uplus$. Thus, we obtain G as

$$G = G_1 \uplus G_2. \tag{42}$$

The operation can be extended to more components in a natural way, so we can take the multigraph sum of any number of graphs or multigraphs: $\uplus_{i=1}^{N} G_i = G_1 \uplus G_2 \uplus \cdots \uplus G_N$.

Having introduced the multigraph sum, let us now consider a curious property of it, which we call multigraph advantage. We explain it through an important graph parameter, the *edge-connectivity* (for short, we simply call it *connectivity*), denoted by $\lambda(G)$. This is the minimum number of edges that can disconnect the graph, when these edges are deleted. In other words, this is the size of a *minimum cut* in the graph. In case G is a multigraph, the parallel edges are all counted when we consider the size of a cut. The connectivity provides an important characterization of the network topology. For example, it tells how vulnerable the network to link failures. It also shows how rich the network is in link-disjoint routes. Now one can easily see that

$$\lambda(G_1 \uplus G_2) \geq \lambda(G_1) + \lambda(G_2) \tag{43}$$

always holds. More generally, we have

$$\lambda\left(\uplus_{i=1}^{N} G_i\right) \geq \sum_{i=1}^{N} \lambda(G_i). \tag{44}$$

The reason is that the size of any given cut in the multigraph sum is just the sum of the sizes of the corresponding cuts in the components. Thus, the connectivity of the multigraph sum cannot be smaller than the sum of the connectivities of the components. In fact, from the above reasoning, one might first expect that actually the inequalities (43) and (44) always hold with *equality*. This is, however, not the case. It can happen that

$$\lambda(G_1 \uplus G_2) > \lambda(G_1) + \lambda(G_2), \tag{45}$$

that is, the inequality may be strict. (As an exercise, the reader is advised to construct some simple examples!)

This is what we call *multigraph advantage*. Regarding connectivity, it means that the connectivity of the multigraph sum can be *strictly larger* than the sum of the connectivities of the components. Thus, the network topology has a *quantifiable* extra benefit from taking the multigraph sum, representing multiple physical layers.

One may reasonably ask at this point: is this behavior just the consequence of specially chosen examples or is it somehow typical? The following section shows that it is quite typical.

7.2. Analysis of Multigraph Advantage. Mobile wireless networks are often modeled by various types of *random graphs*. In what follows, we first discuss the choice of the considered random graph model, then we formally state the theorem showing that it exhibits significant multigraph advantage.

The type of random graph that is most frequently used to model mobile wireless networks is the *random geometric graph* [65]. In the simplest case, it is generated by randomly placing points in a planar domain, typically a regular one, such as a square. The points represent the randomly positioned network nodes. Any two points are connected with an edge (representing a network link), whenever their distance is at most a given value r, that stands for the transmission radius.

There are various generalizations of the above simple model. For example, the domain can be irregular and it may be part of a higher dimensional space, possibly with a different metric. It can even be elevated to a more general and abstract level, as discussed in earlier sections.

On the other hand, the oldest and most studied random graph model is the one in which all edges are chosen *independently* with the same probability [39]. This model is referred to as independent-edge random graph, Erdős-Rényi random graph, or sometimes Bernoulli random graph. While this model lacks the edge correlations induced by geometry, nevertheless it has a number of important advantages, beyond the fact that it is much more amenable to mathematical analysis.

(i) In situations where the transmission radius is comparable with the domain diameter, the main reason for a missing link is not the distance, it is the presence of random obstacles or random variations in the radio propagation. In such case, an independent edge model can be adequate, or even more accurate than a geometric random graph.

(ii) Some studies have found (see [104, 105]) that if the radio propagation model is more realistic, that is, it takes into account statistical variations around the mean power, then it tends to decrease link correlations, and the random graph becomes similar to the indepedent-edge model.

(iii) If the network applies power control, it tends to counterbalance the effect of distance, especially if the

distances are not too large. This again points to the applicability of the independent edge model.

The above reasons justify that we choose the independent-edge model for investigating the multigraph advantage.

Let us consider the following situation. Let $G_{n,p}$ denote a random graph on n nodes, with edge probability p. Note that typically p is a function of n, that is, $p = p(n)$, and we are interested in asymptotic properties, when $n \to \infty$.

A well-known property (see [39]) is that $G_{n,p}$ is connected asymptotically with probability 1 if and only if

$$p(n) = \frac{\log n + \omega(n)}{n}, \tag{46}$$

where $\omega(n)$ is any function that tends to infinity with n. Informally, this means that the edge probability of $(\log n)/n$ is the minimum requirement for connectivity, as $\omega(n)$ can tend to infinity arbitrarily slowly.

We investigate the multigraph advantage in a situation when the network is moderately dense in the sense that the average degree is constant times higher than the minimum needed for connectivity.

Definition 13. The random graph $G_{n,p}$ is said to be in the moderately dense regime if there exists a constant $c > 1$, such that

$$p(n) = \frac{c \log n}{n} \tag{47}$$

holds.

Now we consider the following situation. Let G_{n,p_1} and G_{n,p_2} be two independently drawn random graphs on the same set of nodes, with edge probabilities p_1 and p_2, respectively. Assume that they are both in the moderately dense regime. Let us compare the connectivities $\lambda(G_{n,p_1})$ and $\lambda(G_{n,p_2})$ and the multigraph connectivity $\lambda(G_{n,p_1} \uplus G_{n,p_2})$. We can prove the following result (see [12]) about the multigraph advantage regarding connectivity.

Theorem 14. *Let G_{n,p_1} and G_{n,p_2} be independently drawn random graphs in the moderately dense regime, on the same set of nodes. Let their edge probabilities be*

$$p_1(n) = \frac{a \log n}{n}, \qquad p_2(n) = \frac{b \log n}{n}, \tag{48}$$

with constants $a, b > 1$. Then there exists a constant $c = c(a, b) > 0$, such that the asymptotic multigraph advantage regarding connectivity is at least $c \log n$. That is,

$$\lim_{n \to \infty} \Pr\left(\frac{\lambda\left(G_{n,p_1} \uplus G_{n,p_2}\right)}{\lambda\left(G_{n,p_1}\right) + \lambda\left(G_{n,p_2}\right) + c \log n} \geq 1 \right) = 1. \tag{49}$$

Thus, as shown by Theorem 14, the multigraph advantage is quite significant in the moderately dense regime, for two reasons. First, it tends to infinity as the graphs grow. Second, since it is known [39] that in the moderately dense regime, the connectivity of the random graph is $O(\log n)$, therefore, the gain in connectivity is of the same order of magnitude as the connectivity of the component graphs itself. In other words, there is a guaranteed constant percentage of *relative gain*, and this percentage does not vanish as $n \to \infty$.

8. The Issue of Connectivity in Wireless Networks

Due to the random network topology, it is not at all guaranteed that any two nodes can send messages to each other. To guarantee that all nodes can reach each other, a minimum requirement is that the network topology (which is usually represented by an undirected graph) is *connected.* Unfortunately, the connectivity requirement is not as innocent as it may look, due to random node positions and limited wireless transmission ranges. It turns out (see, e.g., Gupta and Kumar [50, 61]) that in typical cases, such as placing the nodes in a planar disk independently and uniformly at random, the price of connectivity is very high: the transmission range needs to be set such that it asymptotically results in an infinitely growing number of neighbors.

This phenomenon is a serious *threat to scalability* in these networks. After all, one cannot expect that a small wireless node with limited power and modest capabilities can serve an unbounded number of neighbors.

One might hope at this point that for different modeling assumptions, the situation may perhaps improve. For example, one may try different deployment domains, different probability distributions, different distance metrics, and so forth. Unfortunately, however, it has been proven in very general models that none of these can relieve the scalability bottleneck, see Faragó [8, 10]. It appears that unbounded node degrees are unavoidable whenever full connectivity is required in the limit in a random, geometrically induced topology. This is, of course, bad news for hoping a scalable implementation.

It is therefore of keen importance whether better scalability can be achieved if we are willing to give up full connectivity and substitute it with the milder requirement of *partial connectivity*. This means that, as a price for keeping the node degrees bounded, we accept that only most, but not all, nodes are in a connected component. The motivation is that in many potential applications, such as a network of randomly placed sensors, it is acceptable to have only a majority (say, 99%) of nodes in a connected component and the rest are possibly disconnected.

Next, we show through two case study examples that our abstract geometric random graph model and the related results can serve as powerful tools to analyze the situation, even when that would be very hard to do via the traditional methods.

9. Case Study 1: A Sensor Network Problem

9.1. Problem Setting. Let us consider a large sensor network. Due to the limited processing capabilities of the small sensor nodes, each one is capable of maintaining connections only

two at most three other nodes in our example. The existence of wireless links depends on distance, but the actual form of the dependence is unknown. Moreover, random obstacles to radio waves are also present, and two nodes can only communicate if no such obstacle separates them.

The sensor nodes are distributed in space independently, according to a common, but unknown probability distribution. The locations of the random obstacles are also independent of each other and of the node locations, but otherwise the position, size, and shape of each obstacle can have an arbitrary probability distribution, which is again unknown. We only assume that the events that links are blocked by an obstacle can be considered independent.

Without further information about this sensor network, is it possible to provide a nontrivial lower bound on the number of sensors that will be necessarily pushed to the "periphery"? By periphery, we mean those nodes that are not part of the largest connected component of the network topology.

The traditional approach to answer this question would be to specify the probability distributions and other parts of the model (such as how link existence depends on distance, etc.), and then do (tedious) calculations under the specific conditions. If, however, anything changes in the conditions, the results may not carry over. Our general approach will make it possible to avoid this, and provide a nontrivial bound that is valid for all practically relevant cases, as explained next.

9.2. Solution via Abstract Geometric Random Graphs. In the solution, we use the terminology and results introduced in Section 5. In the above-described sensor network example, we can observe that the model is described by a local and name invariant abstract geometric graph model, no matter what the unknown probability distributions are. The reason for locality is that once we choose the positions of two sensors, the probability that a link exists between them does not depend on the locations of other sensors. Although it does depend on the obstacles, they are independent of the sensor positions and block the links independently. Name invariance also holds in this example, as the sensor positions are i.i.d., which is a special case of an exchangeable system of random variables. The node degree bound of 3 yields $D_{\mathcal{M}} \leq 3$. By Theorem 8, the threshold function for β-connectivity in our case satisfies

$$f(D_{\mathcal{M}}) = 1 - e^{-D_{\mathcal{M}}} \leq 1 - e^{-3} \approx 0.95. \tag{50}$$

Thus, we can conclude that despite the very vague information about the system, we are still able to calculate that at least about 5% of the nodes cannot belong to the largest connected component.

Thus, our general result was able to easily come to a conclusion that would otherwise be rather hard to obtain without having further information.

10. Case Study 2: Mobile Ad Hoc Network

10.1. Problem Description. In this example, we model a mobile wireless ad hoc network. The initial position of each node is chosen in the following way. Let P be a probability measure over a planar domain D. First we choose k pivot points independently at random, using P. Then the actual node positions are generated such that each potential node is chosen independently at random from P, but it is kept only if it is within a given distance d_0 to at least one of the random pivot points, otherwise it is discarded. Note that this way of generating the nodes makes them dependent, as the nondiscarded ones cluster around the random pivot points, thus modeling a clustered, nonindependent node distribution.

Remark 15. The fact that the above node generation may indeed result in *dependent* node locations can be seen the easiest in the special case when $k = 1$, the domain D is a circular disk, and P is the uniform distribution over D. Since the pivot point is then chosen uniformly at random over D, therefore, any location in D will have a positive probability of falling within d_0 distance of the pivot. Therefore, the individual probability density of any given nondiscarded node is strictly positive everywhere in D. On the other hand, if we condition on the event that a nondiscarded node falls within some distance d of the center of the disk, then it implies that the pivot must have fallen within $d+d_0$ distance of the center. Then, since any other nondiscarded node must fall within distance d_0 of the pivot, therefore, any nondiscarded node will be at most at distance $d + 2d_0$ of the center. If $d + 2d_0$ is less than the radius of the disk, then the conditioning on the position of one nondiscarded node changed the distribution of others, since under the condition they cannot fall *anywhere* with positive probability density in D. Thus, they are not independent, since if they were independent, then conditioning on the location of one of them could not influence the distribution of others.

Let us model the mobility of the nodes in this example in the following way. Over some time horizon T_n, that may depend on n, the number of nodes, each node moves along a random curve from its initial position with a constant speed v_0. The curve is chosen from a set $\mathcal{C}$ of available potential trajectories in D. For simplicity, it is assumed that each curve can be identified by a real parameter. This parameter is chosen using a probability distribution $Q_{x,y}$ that depends on the initial position (x, y) of the node. Then the randomly obtained curve is shifted so that its start point coincides with the random initial position of the node and then the node will move along this random trajectory. It is assumed that $\mathcal{C}$ and D are such that the shifted curves still remain in the domain.

Let $d(x, y)$ be a nonnegative real-valued function over $D \times D$, with the only restriction that $d(x, x) = 0$ holds for any x. This function is intended to measure "radio distance" in D. The assumption is that whenever $d(x, y)$ is small enough, then two nodes positioned at x and y can receive each others' transmissions. The function $d(x, y)$, however, does not have to satisfy the usual distance axioms, it may reflect complex radio propagation characteristics, such as expected attenuation and fading, it may account for the heterogeneity of the terrain, for propagation obstacles, and so forth.

We may also include random effects, making $d(x, y)$ a random variable, reflecting special conditions of interest, such as the random presence of eavesdroppers that can trigger the inhibition of certain links. We assume, however, that if there is randomness in $d(x, y)$, then it is independent of the other random variables in the model.

We now define the links of the network, as follows. Consider two nodes with initial position vectors $X_1(0)$ and $X_2(0)$, respectively. As they move along their random trajectories, their positions at time t is denoted by $X_1(t)$ and $X_2(t)$, respectively. The two nodes are considered connected by a link, if there is a closed subinterval of length at least t_n within the time horizon $[0, T_n]$, such that $d(X_1(t), X_2(t)) \leq r_n$ holds for every time t within the subinterval, with the possibly complicated radio distance. Here, t_n and r_n are parameters that may also depend on the number n of nodes. The motivation for this link definition is that the nodes should be within range at least for the time of sending a packet.

Now the question is for given P, D, $\mathcal{C}$, $Q_{x,y}$, and $d(x, y)$ and for the described way of dependent node generation, can we somehow choose the model parameters k, d_0, v_0, T_n, t_n, and r_n, such that the expected node degrees remain bounded by, say, 4 and still at least 99% of nodes belong to a connected component?

We believe that it would be rather hard to answer this question with a direct analysis for arbitrary complex choices of P, D, $\mathcal{C}$, $Q_{x,y}$, and $d(x, y)$. On the other hand, in view of our general results that we build up in the subsequent sections, it becomes quite straightforward.

10.2. Solution. In the solution, we use again the terminology and results introduced in Section 5. As we have already noted, it would be hard to solve the problem for arbitrary choices of P, D, $\mathcal{C}$, $Q_{x,y}$, and $d(x, y)$ with directly analyzing the stochastic geometry of the model. On the other hand, we can easily check that it satisfies our general conditions, as shown below.

Let us choose the model domain S as a 3-dimensional phase space, in which each node is represented by a point such that the first two coordinates describe the initial position of the node and the last coordinate encodes which random trajectory was chosen from $\mathcal{C}$ for the node. Let X_1, $X_2, \ldots$ be the representations of the nodes in this phase space.

We can now check that, for any n, the joint distribution of $X_1, \ldots, X_n$ is invariant to reindexing them. The reason is that both the initial positions and the trajectory choices are generated by processes in which the indices do not play any role. Therefore, the model is *name invariant*. Interestingly, this remains true despite having a lot of dependencies among the nodes: the initial positions of different nodes are not independent (due to clustering), and the trajectory of a given node is also not independent of its initial position, as it is drawn from a probability distribution that may depend on the location. Through this, the trajectories and initial positions of different nodes also become dependent, making their whole movement dependent. Yet, the model is still name invariant.

Let us now consider the links. As defined in Section 9, two nodes are considered connected if during their movement over the time horizon $[0, T_n]$ there is a subinterval of time, of length at least t_n, such that they remain within "radio distance" $\leq r_n$ during the entire subinterval. The radio distance, however, may be very different from the Euclidean distance, it may be described by an arbitrary function that may account for complex propagation characteristics, attenuation, obstacles, and it may also contain independent randomness.

Given some possibly complicated radio distance $d(x, y)$ and the node generation and movement process with possibly complex trajectories, it may not be easy to compute whether a link actually exists between two nodes according to the above definition. On the other hand, for us, it is enough to note that once the phase space representations X_i, X_j of any two nodes are given, plus the realization of the independent randomness of the distance, they together determine whether a link exists between the two nodes or not. The reason is that the initial positions and the trajectories, given in the phase space representation, fully determine the movement of the nodes. Once this is known, it determines, along with the realization of the independent randomness of the distance function, whether the link definition is satisfied, that is, if there is a subinterval of length $\geq t_n$ in $[0, T_n]$, such that the nodes stay within radio distance $\leq r_n$ during the entire subinterval. To actually compute, it may not be easy for a sophisticated case, but for our purposes it is enough to know that it is *determined* by the listed factors, without knowing anything about the other nodes. This implies that the model is *local*.

Thus, we have established that the problem can be described by a local and name invariant abstract geometric graph model, for any choice of the parameters. Then, by Theorem 12, the threshold function for β-connectivity is

$$f(D_{\mathcal{M}}) = 1 - e^{-D_{\mathcal{M}}}. \tag{51}$$

If we require that node degrees are bounded by, say, 4, then we have $D_{\mathcal{M}} \leq 4$, implying

$$f(D_{\mathcal{M}}) = 1 - e^{-D_{\mathcal{M}}} \leq 1 - e^{-4} < 0.9817. \tag{52}$$

Thus, the threshold in this case falls below 0.99, so by Theorem 12, it is impossible to achieve that asymptotically 99% of the nodes belong to a connected component, no matter how the other parameters are chosen.

Note that the direct application of our general results was able to cut through a lot of complexity that would otherwise arise if we wanted to reach the same conclusions by directly analyzing the stochastic geometry of such a model.

11. Conclusion

We have reviewed a number of models and results that can describe and analyze large, random network topologies, primarily (but not exclusively) focusing on wireless networks. We have also demonstrated that it is possible to find common generalizations of various sets of models. These generalizations can help in solving hard problems that

would be much more difficult to solve via the traditional approaches.

We hope that the introduction of numerous models and results could convince the interested reader that the modeling of very large networks with various random mathematical structures is a fruitful, vibrant, and promising research area.

Acknowledgment

The author gratefully acknowledges the support of NSF Grant no. CNS-1018760.

References

[1] Y. Hirate, S. Kato, and H. Yamana, "Web structure in 2005," in *Proceedings of the Algorithms and Models for the Web Graph (WAW '06)*, pp. 36–46, Banff, Canada, December 2006.

[2] S. Basagni, A. Faragó, M. A. Nanni, and D. T. Tran, "Increased connectivity at lower cost: the case for multi-radio nodes in multi-hop wireless networks," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '09)*, Honolulu, Hawai, USA, December 2009.

[3] S. Basagni, F. Nidito, and A. Faragó, "The multi-radio advantage," in *Proceedings of IEEE Radio and Wireless Symposium (RWS '09)*, pp. 478–481, San Diego, Calif, USA, January 2008.

[4] A. Faragó, "Analysis and modeling of very large network topologies," *Infocommunications Journal*, vol. 4, pp. 1–10, 2012.

[5] A. Faragó, "On the structure of classes of random graphs," *Chicago Journal of Theoretical Computer Science*, 2010.

[6] A. Faragó, "Scalability of node degrees in random wireless network topologies," *IEEE Journal on Selected Areas in Communications*, vol. 27, pp. 1238–1244, 2009.

[7] A. Faragó, "Quantifying the trade-off between the level of connectivity and local complexity in random wireless network topologies," in *Proceedings of the 6th International Conference on the Quantitative Evaluation of Systems (QEST '09)*, pp. 237–246, Budapest, Hungary, September 2009.

[8] A. Faragó, "On the fundamental limits of topology control in ad hoc networks," *Algorithmica*, vol. 49, no. 4, pp. 337–356, 2007.

[9] A. Faragó, "Analysis of fundamental limits for partial connectivity in wireless networks," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '09)*, Honolulu, Hawai, USA, December 2009.

[10] A. Faragó, "Asymptotically optimal trade-off between local and global connectivity in wireless networks," *Performance Evaluation*, vol. 68, no. 2, pp. 142–156, 2011.

[11] A. Faragó, "Random graph models and the limits of scalability in Ad Hoc and sensor networks," in *Proceedings of the International Workshop on Theoretical and Algorithmic Aspects of Sensor and Ad-hoc Networks (WTASA '07)*, pp. 1–12, Miami, Fla, USA, June 2007.

[12] A. Faragó and S. Basagni, "The effect of multi-radio nodes on network connectivity—a graph theoretic analysis," in *Proceedings of the IEEE International Workshop on Wireless Distributed Networks (WDN '08)*, Cannes, France, September 2008.

[13] A. Faragó, "Algorithmic challenges in Ad Hoc networks," in *Mobile Ad Hoc Networking*, S. Basagni, M. Conti, S. Giordano, and I. Stojmenovic, Eds., pp. 427–445, IEEE Press and Wiley-Interscience, 2004.

[14] A. Faragó, "Efficient blocking probability computation of complex traffic flows for network dimensioning," *Computers and Operations Research*, vol. 35, no. 12, pp. 3834–3847, 2008.

[15] A. Faragó, "Algorithmic challenges in learning path metrics from observed choices," *Applied Artificial Intelligence*, vol. 22, no. 7-8, pp. 749–760, 2008.

[16] A. Faragó, "A general tractable density concept for graphs," *Mathematics in Computer Science*, vol. 1, no. 4, pp. 689–699, 2008.

[17] A. Faragó, "Random graph models and the limits of scalability in Ad Hoc and sensor networks," in *Sensor and ad Hoc Networks, Theoretical and Algorithmic Aspects*, S. K. Makki, X. Y. Li, and N. Pissinou, Eds., pp. 19–42, Springer, 2008.

[18] A. Faragó, "On the typical case complexity of graph optimization," *Discrete Applied Mathematics*, vol. 153, no. 1–3, pp. 73–88, 2005.

[19] A. Faragó and V. R. Syrotiuk, "MERIT: a scalable approach for protocol assessment," *Mobile Networks and Applications*, vol. 8, no. 5, pp. 567–577, 2003.

[20] A. Faragó and S. Basagni, "Connecting two worlds: physical models and graph models of wireless network topologies," in *Proceedings of the 13th IEEE International Workshop on Future Trends of Distributed Computing Systems (FTDCS '11)*, pp. 739–744, Banff, Canada, September 2011.

[21] A. Faragó and D. T. Tran, "The multigraph advantage and its application in network analysis," in *Proceedings of the EURO 24th Conference on Operations Research*, Session 6.2.52, (Telecommunications), Paper 1, Lisbon, Portugal, July 2010.

[22] A. Faragó, "Structural properties of random graph models," in *Proceedings of the Computing: The Australasian Theory Symposium (CATS '09)*, Wellington, New Zealand, January 2009.

[23] A. Faragó and S. Basagni, "The effect of multi-radio nodes on network connectivity—a graph theoretic analysis," in *Proceedings of the IEEE International Workshop on Wireless Distributed Networks (WDM '08)*, Cannes, France, September 2008.

[24] A. Faragó, "Graph theoretic analysis of Ad Hoc network vulnerability," in *Proceedings of the Workshop on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt '03)*, pp. 171–180, INRIA Sophia-Antipolis, France, March 2003.

[25] N. Meghanathan and A. Faragó, "On the stability of paths, Steiner trees and connected dominating sets in mobile ad hoc networks," *Ad Hoc Networks*, vol. 6, no. 5, pp. 744–769, 2008.

[26] C. Wang, M. A. Park, J. Willson, Y. Cheng, A. Faragó, and W. Wu, "On approximate optimal dual power assignment for biconnectivity and edge-biconnectivity," *Theoretical Computer Science*, vol. 396, no. 1–3, pp. 180–190, 2008.

[27] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins, "Trawling the web for emerging cyber communities," in *Proceedings of the 8th World Wide Web Conference*, Toronto, Canada, 1999.

[28] A. L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[29] A. L. Barabási, R. Albert, and H. Jeong, "Scale-free characteristics of random networks: the topology of the world-wide web," *Physica A*, vol. 281, no. 1, pp. 69–77, 2000.

[30] A. Broder, R. Kumar, F. Maghoul et al., "Graph structure in the web," in *Proceedings of the 9th World Wide Web Conference*, Amsterdam, The Netherland, May 2000.

[31] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication (SIGCOMM '99)*, pp. 251–262, New York, NY, USA, 1999.

[32] A. Bonato, *A Course on the Web Graph*, American Mathematical Society, 2008.

[33] S. Bornholdt and H. G. Schuster, *Handbook of Graphs and Networks*, Wiley-VCH, 2003.

[34] G. Caldarelli, *Scale-Free Networks*, Oxford University Press, 2007.

[35] F. Chung and L. Lu, *Complex Graphs and Networks*, American Mathematical Society, 2006.

[36] M. Newman, A. L. Barabási, and D. J. Watts, *The Structure and Dynamics of Networks*, Princeton University Press, 2006.

[37] L. Li, D. Alderson, J. C. Doyle, and W. Willinger, "Towards a theory of scale-free graphs: definition, properties, and implications," *Internet Mathematics*, vol. 2, pp. 431–523, 2005.

[38] D. Alderson, L. Li, W. Willinger, and J. C. Doyle, "Understanding internet topology: principles, models, and validation," *IEEE/ACM Transactions on Networking*, vol. 13, no. 6, pp. 1205–1218, 2005.

[39] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády, "The degree sequence of a scale-free random graph process," *Random Structures and Algorithms*, vol. 18, no. 3, pp. 279–290, 2001.

[40] W. Aiello, F. Chung, and L. Lu, "Random evolution in massive graphs," in *Handbook on Massive Data Sets*, J. Abello, P. M. Pardalos, and M. G. C. Resende, Eds., pp. 97–122, Kluwer Academic Publishers, 2002.

[41] J. Kleinberg, R. Kumar, P. Raghavan et al., "The web as a graph: measurements, models and methods," in *Proceedings of the International Conference on Combinatorics and Computing*, 1999.

[42] F. Chung and L. Lu, "Coupling off-line and on-line analysis of random power law graphs," *Internet Mathematics*, vol. 1, pp. 409–461, 2004.

[43] J. Leskovec, D. Chakrabarti, J. Kleinberg, and C. Faloutsos, "Realistic, mathematically tractable graph generation and evolution, using Kronecker multiplication," in *Proceedings of the Knowledge Discovery in Databases (PKDD '05)*, vol. 3721 of *Lecture Notes in Computer Science*, pp. 133–145, Springer, 2005.

[44] F. Chierichetti, R. Kumar, S. Lattanzi, A. Panconesi, and P. Raghavan, "Models for the compressible web," in *Proceedings of the 50th Annual Symposium on Foundations of Computer Science (FOCS '09)*, pp. 331–340, October 2009.

[45] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graphs over time: densification laws, shrinking diameters and possible explanations," in *Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 177–187, Chicago, Ill, USA, August 2005.

[46] A. D. Flaxman, A. M. Frieze, and J. Vera, "A geometric preferential attachment model of networks," *International Workshop on Algorithms and Models for the Web-Graph*, vol. 3243, pp. 44–55, 2004.

[47] W. Aiello, A. Bonato, C. Cooper, J. Janssen, and P. Pralat, "A spatial web graph model with local influence," *Internet Mathematics*, vol. 5, pp. 175–196, 2009.

[48] A. D. Flaxman, "Expansion and lack thereof in randomly perturbed graphs," Microsoft Research Technical Report TR-2006-118, Web Algorithms Workshop, 2006.

[49] F. Chung and L. Lu, "The average distance in a random graph with given expected degrees," *Internet Mathematics*, vol. 1, pp. 91–114, 2003.

[50] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[51] S. Toumpis and A. Goldsmith, "Ad hoc network capacity," in *Proceedings of the 34th Asilomar Conference on Signals, Systems and Computers*, pp. 1265–1269, November 2000.

[52] P. Gupta and P. Kumar, "Internets in the sky: the capacity of three dimensional wireless networks," *Communications in Information and Systems*, vol. 1, pp. 39–49, 2001.

[53] S. Yi, Y. Pei, and S. Kalyanamaran, "On the capacity improvement of Ad Hoc wireless networks using directional antennas," in *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC '03)*, pp. 108–116, Annapolis, Md, USA, June 2003.

[54] C. Peraki and S. D. Servetto, "On the maximum satble throughput problem in random networks with directional antennas," in *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC '03)*, pp. 76–87, Annapolis, Md, USA, June 2003.

[55] U. C. Kozat and L. Tassiulas, "Throughput capacity of random Ad Hoc networks with infrastructure support," in *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom '03)*, pp. 55–65, San Diego, Calif, USA, September 2003.

[56] B. Liu, Z. Liu, and D. Towsley, "On the capacity of hybrid wireless networks," in *Proceedings of the 22nd Annual Joint Conference on the IEEE Computer and Communications Societies (INFOCOM '03)*, pp. 1543–1552, San Francisco, Calif, USA, April 2003.

[57] S. Toumpis, "Capacity bounds for three classes of wireless networks: asymetirc, cluster and hybrid," in *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC '04)*, pp. 133–144, Roppongi, Japan, May 2004.

[58] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 477–486, 2002.

[59] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay trade-off in wireless networks," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM '04)*, pp. 464–475, Hong Kong, March 2004.

[60] A. Özgür, O. Lévêque, and D. Tse, "Hierarchical cooperation achieves optimal capacity scaling in Ad Hoc networks," *IEEE Transactions on Information Theory archive*, vol. 53, no. 10, pp. 3549–3572, 2007.

[61] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," in *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*, W. M. McEneany, G. Yin, and Q. Zhang, Eds., pp. 547–566, Birkhauser, Boston, Mass, USA, 1998.

[62] J. Li, C. Blake, D. S. J. De Couto, H. I. Lee, and R. Morris, "Capacity of Ad Hoc wireless networks," in *Proceedings of the ACM 7th Annual International Conference on Mobile Computing and Networking (MOBICOM '01)*, pp. 61–69, Rome, Italy, 2001.

[63] O. Dousse, P. Thiran, and M. Franceschetti, "On the throughput scaling of wireless relay networks," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2756–2761, 2006.

[64] M. Franceschetti and R. Meester, *Random Networks for Communication*, Cambridge University Press, 2007.

[65] M. Penrose, *Random Geometric Graphs*, Oxford University Press, 2003.

[66] B. Bollobás and O. Riordan, "Mathematical results on scale-free random graphs," in *Handbook of Graphs and Networks*, S. Bornholdt and H. G. Schuster, Eds., pp. 1–34, Wiley-VCH, 2003.

[67] B. N. Clark, C. J. Colbourn, and D. S. Johnson, "Unit disk graphs," *Discrete Mathematics*, vol. 86, no. 1–3, pp. 165–177, 1990.

[68] S. Schmid and R. Wattenhofer, "Modeling sensor networks," in *Algorithms and Protocols for Wireless Sensor Networks*, A. Boukerche, Ed., John Wiley & Sons, 2008.

[69] F. Kuhn, R. Wattenhofer, and A. Zollinger, "Ad hoc networks beyond unit disk graphs," *Wireless Networks*, vol. 14, no. 5, pp. 715–729, 2008.

[70] I. F. Akyildiz, W. Y. Lee, and K. R. Chowdhury, "Spectrum management in cognitive radio ad hoc networks," *IEEE Network*, vol. 23, no. 4, pp. 6–12, 2009.

[71] F. Kuhn, T. Moscibroda, and R. Wattenhofer, "On the locality of bounded growth," in *Proceedings of the 24th Annual ACM Symposium on Principles of Distributed Computing (PODC '05)*, pp. 60–68, New York, NY, USA, July 2005.

[72] T. K. Philips, S. S. Panwar, and A. N. Tantawi, "Connectivity properties of a packet radio network model," *IEEE Transactions on Information Theory*, vol. 35, no. 5, pp. 1044–1047, 1989.

[73] P. J. Wan and C. W. Yi, "Asymptotic critical transmission radius and critical neighbor number for k-connectivity in wireless ad hoc networks," in *Proceedings of the 5th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MoBiHoc '04)*, pp. 1–8, Roppongi, Japan, May 2004.

[74] M. D. Penrose, "The longest edge of the random minimal spanning tree," *Annals of Applied Probability*, vol. 7, no. 2, pp. 340–361, 1997.

[75] M. D. Penrose, "A strong law for the longest edge of the minimal spanning tree," *Annals of Probability*, vol. 27, no. 1, pp. 246–260, 1999.

[76] F. Xue and P. R. Kumar, "The number of neighbors needed for connectivity of wireless networks," *Wireless Networks*, vol. 10, no. 2, pp. 169–181, 2004.

[77] M. D. Penrose, "On k-connectivity for a geometric random graph," *Random Structures and Algorithms*, vol. 15, no. 2, pp. 145–164, 1999.

[78] M. J. B. Appel and R. P. Russo, "The minimum vertex degree of a graph on uniform points in $[0, 1]^d$," *Advances in Applied Probability*, vol. 29, no. 3, pp. 582–594, 1997.

[79] H. Dette and N. Henze, "The limit distribution of the largest nearest neighbour link in the Unit d-Cube," *Journal of Applied Probability*, vol. 26, pp. 67–80, 1989.

[80] B. Bollobás, *Random Graphs*, Cambridge University Press, 2001.

[81] S. Janson, T. Luczak, and A. Rucinski, *Random Graphs*, John Wiley & Sons, 2000.

[82] K. Pahlavan and P. Krishnamurthy, *Principles of Wireless Networks*, Prentice Hall, 2002.

[83] J. E. Yukich, *Probability Theory of Classical Euclidean Optimization Problems*, vol. 1675 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 1998.

[84] C. Avin, Y. Emek, E. Kantor, Z. Lotker, D. Peleg, and L. Roditty, "SINR diagrams: towards algorithmically usable SINR models of wireless networks," in *Proceedings of the 28th Symposium on Principles of Distributed Computing (PODC '09)*, pp. 200–209, Calgary, Alberta, Canada, August 2009.

[85] C. Yan, Y. Xiang, and F. F. Dragan, "Compact and low delay routing labeling scheme for unit disk graphs," in *Proceedings of the Workshop on Algorithms and Data Structures (WADS '09)*, vol. 5664 of *Lecture Notes in Computer Science*, pp. 566–577, 2009.

[86] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks*, vol. 47, no. 4, pp. 445–487, 2005.

[87] R. Bruno, M. Conti, and E. Gregori, "Mesh networks: commodity multihop ad hoc networks," *IEEE Communications Magazine*, vol. 43, no. 3, pp. 123–131, 2005.

[88] B. Awerbuch, D. Holmer, and H. Rubens, "High throughput route selection in multi-rate ad hoc wireless networks," in *Proceedings of the 1st Working Conference on Wireless On-demand Networks Systems (WONS '04)*, R. Battiti, M. Conti, and E. Lo Cigno, Eds., vol. 2998 of *Lecture Notes in Computer Science*, pp. 253–270, Madonna di Campiglio, TN, Italy, January 2004.

[89] J. Tang, G. Xue, and W. Zhang, "Interference-aware topology control and QoS routing in multi-channel wireless mesh networks," in *Proceedings of 6th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MOBIHOC '05)*, pp. 68–77, Urbana-Champaign, Ill, USA, May 2005.

[90] R. Chandra, P. Bahl, and P. Bahl, "Multinet: connecting to multiple IEEE 802.11 networks using a single wireless card," in *Proceedings of IEEE Conference on Computer Communications (INFOCOM '04)*, pp. 882–893, Hong Kong, March 2004.

[91] P. Kyasanur and N. H. Vaidya, "Capacity of multi-channel wireless networks: impact of number of channels and interfaces," in *Proceedings of the 11th Annual International Conference on Mobile Computing and Networking (MobiCom '05)*, pp. 43–57, Cologne, Germany, September 2005.

[92] M. Kodialam and T. Nandagopal, "Characterizing achievable rates in multi-hop wireless mesh networks with orthogonal channels," *IEEE/ACM Transactions on Networking*, vol. 13, no. 4, pp. 868–880, 2005.

[93] M. Kodialam and T. Nandagopal, "Characterizing achievable rates in wireless multi-hop networks: the joint routing and scheduling problem," in *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom '03)*, pp. 42–54, San Diego, Calif, USA, September 2003.

[94] "Wireless community network," http://en.wikipedia.org/wiki/Wireless_community_network.

[95] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," in *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking (MobiCom '04)*, pp. 114–128, Philadelphia, Pa, USA, October 2004.

[96] P. Bahl, A. Adya, J. Padhye, and A. Wolman, "Reconsidering wireless systems with multiple radios," *ACM SIGCOMM Computer Communications Review*, vol. 34, no. 5, pp. 39–46, 2004.

[97] A. Miu, H. Balakrishnan, and C. E. Koksal, "Improving loss resilience with multi-radio diversity in wireless networks," in *Proceedings of the 11th Annual International Conference on Mobile Computing and Networking (MobiCom '05)*, pp. 16–30, Cologne, Germany, September 2005.

[98] J. Shin, H. Choi, P. G. Traynor, and T. F. La Porta, "Network formation schemes for dynamic multi-radio, multi-hop wireless celluar systems," Tech. Rep. NAS-TR-0005-2005, The Pennsylvania State University, University Park, Pa, USA, 2005.

[99] J. Shin, P. Kishen, and T. F. La Porta, "Dynamic multi-frequency, multi-hop wireless cellular networks," in *Proceedings of the IFIP Mediterranean Ad Hoc Networking Workshop*, Lipari, Italy, June 2006.

[100] M. Alicherry, R. Bhatia, and L. Li, "Joint channel assignment and routing for throughput optimization in multi-radio wireless mesh networks," in *Proceedings of the 11th Annual International Conference on Mobile Computing and Networking (MobiCom '05)*, pp. 58–72, Cologne, Germany, September 2005.

[101] M. Shin, S. Lee, and Y. A. Kim, "Distributed channel assignment for multi-radio wireless networks," in *Proceedings of IEEE International Conference on Mobile Ad Hoc and Sensor Sysetems (MASS '06)*, pp. 417–426, October 2006.

[102] K. N. Ramachandran, E. M. Belding, K. C. Almeroth, and M. M. Buddhikot, "Interference-aware channel assignment in multi-radio wireless mesh networks," in *Proceedings of the 25th IEEE International Conference on Computer Communications (INFOCOM '06)*, pp. 1–12, Barcelona, Spain, April 2006.

[103] A. Adya, P. Bahl, J. Padhye, A. Wolman, and L. Zhou, "A multi-radio unification protocol for IEEE 802.11 wireless networks," in *Proceedings of the 1st International Conference on Broadband Networks (BroadNets '04)*, pp. 344–354, San Jose, Calif, USA, October 2004.

[104] R. Hekmat and P. Van Mieghem, "Connectivity in wireless ad-hoc networks with a log-normal radio model," *Mobile Networks and Applications*, vol. 11, no. 3, pp. 351–360, 2006.

[105] R. Gowaikar, B. Hochwald, and B. Hassibi, "Communication over a wireless network with random connections," *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 2857–2871, 2006.