Cambridge University Press
978-1-107-16258-7 — Stochastic Geometry Analysis of Cellular Networks
Bartłomiej Błaszczyszyn , Martin Haenggi , Paul Keeler , Sayandev Mukherjee
Excerpt
More Information

Part I

Stochastic Geometry

Cambridge University Press
978-1-107-16258-7 — Stochastic Geometry Analysis of Cellular Networks
Bartłomiej Błaszczyszyn , Martin Haenggi , Paul Keeler , Sayandev Mukherjee
Excerpt
More Information

Cambridge University Press
978-1-107-16258-7 — Stochastic Geometry Analysis of Cellular Networks
Bartłomiej Błaszczyszyn , Martin Haenggi , Paul Keeler , Sayandev Mukherjee
Excerpt
More Information

1 Introduction

This chapter discusses the reasons for the importance of heterogeneous network (HetNet) architectures in future wireless network deployments. We begin with a brief survey of the state of wireless networks and the chief challenges faced by designers and operators of such networks. We then discuss possible solutions to these problems and single out small-cell networks as probably the best engineering solution. The deployment of small-cell networks then leads directly to the design and configuration of the resulting heterogeneous networks, which is the focus of this book. We make the case for new approaches to complement the traditional simulation-based schemes for design and performance evaluation when the wireless networks in question are heterogeneous, and provide an overview of the principal results that are now known (and that continue to be added to at a rapid pace), mostly as a result of research done since 2010.

1.1 The Demand for Ubiquitous Connectivity

The Information Age and the Computer Age have converged in the form of portable wireless devices that are full-fledged computers and that can, with a connection to the internet, provide instant always-on access to information anywhere. In other words, we are moving toward the "nomadic computing" vision of Kleinrock (Kleinrock 1996), where we can perform information processing "anytime, anywhere," but instead of doing so in Kleinrock's "disconnected" world, we do so by ensuring that the world is fully "connected."

The main source of such ubiquitous connectivity, especially for mobile and outdoor users, is the *cellular network* as it has evolved over three decades and four generations of air interfaces (see Appendix B for details). Indoors, the cellular network is often complemented or even supplanted by a mix of wired and wireless (WiFi) networks, but since the indoor scenario is the simplest one in which to maintain connectedness, we shall focus on outdoor and mobile scenarios in the following discussion of ubiquitous connectivity. At any rate, it is a fait accompli that the *cellular* model, where *base stations* serve as exclusive access points to the internet within defined geographical areas called *cells*, is the foundation for any network that offers such ubiquitous connectivity.[1]

---

[1] We note here that recent trends in the design of future wireless network architectures envision a weakening of the "exclusive" role of base stations within geographic regions, diluting the whole concept of the cell itself. We discuss these topics later in the book.

Cambridge University Press
978-1-107-16258-7 — Stochastic Geometry Analysis of Cellular Networks
Bartłomiej Błaszczyszyn , Martin Haenggi , Paul Keeler , Sayandev Mukherjee
Excerpt
More Information

The value of such ubiquitous connectivity to business and industry is obvious. What was not obvious before, but is now clear from the popularity of social networking applications on these wireless portable devices, is that such ubiquitous connectivity seems to address a fundamental human desire to connect, and stay connected, to other humans. It is not an exaggeration to say that the availability of such “anytime, anywhere” connectivity helps humans live more fulfilled lives. It is therefore imperative to serve more users than ever before, at higher data rates than ever before.

Given the enormous economic and social forces in favor of increasing the availability and quality of such connectivity, the providers of such connectivity (hereafter called *network operators*) are in the happy situation of facing little to no social challenge to the expansion of their business,[2] and it remains only for them to address the technical challenges involved in designing, deploying, and maintaining networks that offer such ubiquitous connectivity at an affordable price. In Section 1.2, we articulate some of these challenges and discuss suggested solutions.

1.2 Technical Challenges for a Network Operator

A network operator designing or deploying a wireless communication network needs to consider, at a minimum, the following three criteria:

1. *Affordability*: When we talk of affordability from the point of view of the network operator, we mean the costs of deploying (capital expenses, or CapEx) and maintaining (operating expenses) such a network. Note that CapEx includes not only the cost of the equipment on the network, but also the cost of acquiring the spectrum license(s), which can often be substantial for the low (<2 GHz) frequency bands. These bands are in higher demand because of the reduced propagation losses compared to higher-frequency bands, but clearly if higher-frequency bands could be used efficiently by the network operator, they could save on the acquisition costs of this spectrum. Further, large cellular base station towers are expensive to deploy and operate, and efficiencies therein directly translate to the bottom line of the network operator.
2. *Availability*: The challenge of mere access to the internet has been largely addressed nearly worldwide – even the all but obsolete 2G (second generation; see Appendix B) cellular networks offer low-speed (around 300 kbps) access to the internet, and such networks are available worldwide and cover most locations with a settled population. The main technical challenge is the next criterion, namely the quality of the data connections to the users of this operator’s network.
3. *Quality*: In Section 7.3 we will provide a more detailed and technical definition of the quality of a data connection, or a collection of data connections (say, one per user of the network), but for the moment, let us say that from the point of view of the network operator, the quality of the network is indicated by the average total

---

[2] Except from some people who do not want unsightly cellular towers in their towns; as we shall see, even their concerns can be addressed by technology.

throughput per unit area in the network (that is, from all base stations to all users per unit area).[3] This average total throughput per unit area, which we loosely refer to as the *capacity* of the network,[4] is given by the product of the average number of active (actually transmitting and/or receiving data) connections per unit area with the average rate on such a data connection, which we know is in turn the product of the *bandwidth* and the average *spectral efficiency* on this data connection, where the spectral efficiency of a link is closely related to the signal-to-interference-plus-noise ratio (SINR) on that link due to Shannon's formula. In other words, the network capacity is the product of the area-density of active connections, the average spectral efficiency on each such connection, and the bandwidth available to such a connection.

Note that the goal of serving more users at higher data rates than before translates into increasing both the average number of active connections per unit area and the average data rate per active connection. Of the three quantities that together determine the network capacity (the "quality" of the network), the bandwidth is the least easy to increase. Typically, new spectrum is freed up for auction by a national regulator at infrequent intervals, often years apart, and the winners of such spectrum auctions pay large sums for the bandwidth they get to use. Moreover, there is simply never enough spectrum released in a single tranche to increase the bandwidth by a factor of 10, say. Thus, the average data rate cannot increase by much due to increased bandwidth, but relies on improvements in spectral efficiency.

Multiple-antenna transmission technologies (collectively called *multiple-input multiple-output or MIMO*) do yield improvements in spectral efficiency, but the physical sizes of wireless devices preclude the use of more than, say, four antennas (at least for <5 GHz bands), and this in turn limits the achievable spectral efficiency gains. Further, the transmission schemes that seek to exploit these gains are complex and require sophisticated signaling between base stations and user devices.

This leaves us with the final factor determining the network capacity, namely the average activity (number of active data connections per unit area that can be supported with acceptable data rate by the network). This number could be increased by, for example, dividing the available total bandwidth into two parts (bands), and distributing the transmissions over these two bands, provided that each of the two bands is adequate to support an acceptable data rate. In light of the scarcity of available spectrum, it is clear that the above division of spectrum cannot occur too many times. An alternative to spectrum division is spectrum reuse, where we attempt to support all transmissions on the same band while mitigating their mutual interference. It has been estimated that between 1955 and 2010, there was an approximately million-fold increase in network capacity, of which only about a five-fold increase can be attributed to spectral efficiency gains, but a 25,000-fold increase is attributable to spectrum reuse (Andrews 2011).

---

[3] As we shall see, distribution of the total throughput per unit area is more descriptive of network performance than the average, but for now let us focus on the average.

[4] There are several definitions of *capacity* but for the purposes of this section, this is the definition we shall use.

Cambridge University Press
978-1-107-16258-7 — Stochastic Geometry Analysis of Cellular Networks
Bartłomiej Błaszczyszyn , Martin Haenggi , Paul Keeler , Sayandev Mukherjee
Excerpt
More Information

Highly sophisticated means of interference avoidance and/or cancellation exist, but for a robust and scalable network architecture, it would be most desirable to have interference between cells be naturally small. This can indeed be obtained by moving to *small cells*, as we shall see in Section 1.3.

1.3 The Case for Small-Cell Architectures

In the wireless industry, the term *small cell* refers to any cell that is smaller in area than the typical cell of a cellular deployment today, called a *macrocell*, which has an area of about 0.5–2 km^2. There is a loose classification of small cells by size, with the names *microcell* and *picocell* frequently used in the literature. The analysis in this book applies to all small cells of any size, and indeed to macrocells too.

At any rate, even if a base station transmits with power comparable with that of a macrocell base station, its area of coverage will be a small cell if the transmission occurs in a high-frequency band, because propagation losses increase with frequency. In other words, a simple and elegant way to get small cells, and thereby obtain the benefits of spectrum reuse, is to employ high-frequency spectrum. Of course, even without switching to a different spectrum from macrocells, the transmission powers of small-cell base stations can be scaled down compared with macrocells, to reduce the size of the cell.

In short, a smallcell (as opposed to a macrocell) architecture brings the benefits of spectrum reuse while also offering the possibility of increased spectral efficiency (simply because the users are now nearer to the base stations than they would be in macrocells) and reduced equipment costs.[5] These are the primary reasons why small-cell architectures play such an important role in future wireless network deployments.

Consider the following rough calculation (Andrews 2010): take an urban area with a medium-to-high user density of $10,000\,\text{km}^2$, and suppose each of these users is to be served at an average rate of 1 Mbps. Then the required network capacity is $10\,\text{Gbps}/\text{km}^2$. If we tried to support this with $1\,\text{km}^2$-area macrocells, then, even with the unrealistic assumption of a spectral efficiency of 2 bps/Hz, we would require a bandwidth of 5 GHz! On the other hand, if we used small cells with an average radius of 100 m (instead of the roughly 600-m radius of a macrocell with area $1\,\text{km}^2$), then the density of base stations rises by a factor of approximately 30, and the bandwidth requirement may be satisfied with as little as 500 MHz.

1.4 Future Wireless Networks Will Be Heterogeneous

With all the promise of small-cell architectures, it should not be forgotten that nearly all present-day wireless cellular networks use macrocells. Network operators have spent

---

[5] Although each small-cell base station may cost less than a macrocell base station, the density of small-cell base stations is greater than that of macrocell base stations, so whether or not CapEx is reduced depends on the actual prices and densities.

enormous amounts of money and time in deploying and optimizing them, and they expect to amortize these investments over several decades. In other words, the existing macrocellular networks are not going to go away. It follows that the wireless networks of the future will be *heterogeneous* in nature, comprising a macrocell *tier* overlaid with one or more tiers of small cells. Thus the design of wireless networks that meet the desired criteria of affordability, availability, and quality must take into account the fact that such networks are likely to be HetNets.

Notwithstanding the benefits from operating small-cell networks in a different (higher) frequency band from the existing macrocells, it is crucial to study scenarios where the small-cell and macrocell tiers use the same frequency band. This is because nearly all deployments today are macrocellular, so wireless devices support only the bands used in macrocells, hence upcoming small-cell deployments are also more likely than not to use the same bands. However, there are some interesting proposals for future wireless networks that envision entirely different classes of devices that operate solely on dedicated frequency bands in small cells. In this book, we analyze both classes of architectures (with overlapping and nonoverlapping spectrum between the tiers).

1.5 Approaches to the Design of Future Wireless Networks

A key part of the standards development process for future wireless networks is the evaluation of network performance for proposed network architectures and transmission schemes. This evaluation is mostly done according to the following two criteria (Novlan et al. 2015):

1. The plot of block error rate (BLER) vs. SINR on an arbitrarily selected link for the proposed modulation and coding scheme.
2. The cumulative distribution function of the SINR at an arbitrarily selected user in the network, with the details of the modulation and coding scheme abstracted for the sake of simplicity.

The symbol error rate (SER) (or BLER) vs. SINR plot on a single link has traditionally always been analyzed via detailed (down to the bits going into transmitted packets) simulation, because there are no exact analytical results for almost any modulation and coding schemes (although bounds on the SER or BLER can often be derived). These simulators are called link-level simulators.

In this book, we are interested in the aggregate network performance involving multiple users in multiple cells. As described in detail in Section 1.8 until about 2010 there were no analytical results on SINR distribution at an arbitrary user location in such a network. Thus, the SINR distribution at an arbitrary location had to be found via simulation. For the sake of computational feasibility, such multicell simulations usually abstracted the individual link-level packets and their transmission and reception, replacing the link-level behavior with look-up tables modeling what was separately obtained via link-level simulation. These multicell simulations with abstracted physical layer coding/decoding are called *system-level simulations*. In this chapter, when we talk of simulations to evaluate network performance, we mean these system-level simulations.

As stated in the preceding paragraph, until about 2010 the simulation approach was the only one available to evaluate the performance of proposed network architectures. To generate results in a consistent format to compare standards proposals by different companies, the standards bodies defined detailed scenarios for simulation together with detailed lists of results that had to be obtained from such simulation. For example, there were several different possible base station layouts, for each of which several different antenna models had to be simulated, and for each of which in turn several different propagation models had to be simulated. Even though nearly all simulation scenarios are for single-tier macrocell layouts, the number of total scenarios to evaluate via simulation is often very large, requiring several person-weeks of effort to code, debug, and run in each company before every major standards meeting. For example, a typical exhaustive simulation-based investigation for a macrocellular network might include all combinations from a set of two layouts, three or four antenna heights, three cell densities, four or five base station transmit powers, and ten user densities (Novlan et al. 2015).

1.6 The Case against Pure Simulation-Based Investigation

The above set of scenarios to simulate is daunting enough, but even if they are all simulated in full, the total set of all possible scenarios is too large to be explored comprehensively with only a handful of selections each for antenna heights, cell density, base station transmit power, and user density. In other words, the comparison of simulation results across companies for these numerous, yet still limited, set of scenarios serves mainly to calibrate the simulators developed independently by the different companies. It is not even obvious that a significant fraction of practically useful usage situations are captured and/or modeled by the handful of scenarios defined for simulation. It follows that the results for the simulated scenarios may not necessarily yield useful insights as to network performance in other scenarios that have not been simulated.

These comments apply to a single-tier macrocell network. Note that the number of scenarios to be simulated will be squared if there is a second, small-cell, tier overlaid on the macrocell tier, and even if this resulting enormous number of simulations were done, they would cover an even smaller fraction of the space of all possible scenarios than was true for a single-tier macrocell network.[6]

To summarize, exclusive reliance on simulation to evaluate multitier HetNet performance may not only be computationally infeasible, it may even be misleading if used as a basis for HetNet design.

---

[6] This is a consequence of the *curse of dimensionality*, which in this case is simply the fact that if $f < 1$ is the fraction of scenarios simulated for the single-tier network, then even if we simulated the same set of scenarios for each tier of a two-tier network, the resulting fraction of scenarios covered is $f^2 < f$.

Cambridge University Press
978-1-107-16258-7 — Stochastic Geometry Analysis of Cellular Networks
Bartłomiej Błaszczyszyn , Martin Haenggi , Paul Keeler , Sayandev Mukherjee
Excerpt
More Information

1.7 The Case for an Analytical Approach to HetNet Design

Suppose that, by some means, we could sift through the space of scenarios to definitively eliminate those that are *not* going to yield the performance we seek (as measured by the distribution of the SINR at an arbitrarily located user, say). Then we could attempt careful, thorough, and detailed simulation of the remaining scenarios of practical interest with the assurance that we did not overlook any potentially useful or important scenario.

We would have a quick way to eliminate network architecture proposals that do not result in the desired performance if, for example, there existed analytical or semi-analytical results relating the user SINR distribution to the HetNet deployment parameters such as base station transmit powers and densities in the tiers. These analytical results would still be useful even if they were obtained for a slightly simplified (but still realistic) HetNet model, so long as we had reason to believe that the relative performances of two sets of deployment parameters, as obtained via analysis, would mirror what we would obtain via simulation.

The message of this book is that such analytical and semianalytical results do exist, covering a wide range of HetNet models. Nearly all of these results were derived since 2010. In subsequent chapters, we develop the analytical techniques whereby such results can be derived, but an overview of the principal results obtained from this approach is given in Section 1.8.

1.8 The Stochastic Geometric Approach to HetNet Analysis

The set of analytical techniques used to study HetNets that is described in this book comes from a field of mathematics called *stochastic geometry*. The key idea in the application of stochastic geometry to wireless cellular network analysis is to model the locations of base stations and user terminals as realizations of a class of random mathematical objects called *point process*.

1.8.1 A preview of the main results in the book

The remainder of the book is devoted to developing such models and deriving analytical expressions for relevant network performance criteria such as user SINR distribution, beginning with a self-contained overview of the basic theory of point processes in the next chapter. Some of the major results that have been derived are listed below.

1. Consider an arbitrary layout of a tier of base stations in the plane with independent and identically distributed (iid) log-normal[7] fading on all links to the user. Then, under mild conditions on the density of the base stations, asymptotically in the standard deviation of the log-normal fading, the point process of link losses at the user terminal converges in distribution to the point process of link losses from a tier of base stations whose locations are the points of a Poisson point process,

---

[7] See Appendix C for more details on the probability distributions that appear in this book.

CAMBRIDGE

Cambridge University Press
978-1-107-16258-7 — Stochastic Geometry Analysis of Cellular Networks
Bartłomiej Błaszczyszyn , Martin Haenggi , Paul Keeler , Sayandev Mukherjee
Excerpt
More Information

which is often abbreviated PPP or called just the Poisson process (Błaszczyszyn, Karray, & Keeler 2013).

2. This result holds irrespective of the propagation loss model.
3. This result also holds if the iid log-normal fading is replaced by iid Suzuki fading, which combines log-normal *shadow fading* or *shadowing*, with Rayleigh fading (Błaszczyszyn, Karray, & Keeler 2015).
4. For an arbitrarily located user in a tier of base stations whose locations are the points of a homogeneous (in other words, constant-density) Poisson process, with arbitrary iid fading to the user, the point process of link losses at this user terminal from all base stations in the tier has the same intensity function as the path-loss (with no fading) point process from a different homogeneous Poisson-process-located tier of base stations whose density is that of the original tier of base stations times a moment of the fading coefficient. In other words, a network with iid fading can be replaced for analysis purposes by an equivalent network with no fading (Błaszczyszyn, Karray, & Klepper 2010; Haenggi 2008; Madhusudhanan et al. 2014).
5. The signal-to-signal-plus-interference ratio, called *signal-to-total-interference ratio* or STIR, at the typical location from all base stations in a given tier (modeled as a homogeneous Poisson process) is a type of Poisson-Dirichlet point process (Błaszczyszyn & Keeler 2014), and results on this process also hold for the process formed by the STIR and, in some instances, the *signal-to-total-interference-plus-noise ratio* values.
6. A consequence of the this is that the joint distribution of the n strongest STIRs from each tier is available in exact analytic form (Handa 2009).
7. Similarly, the joint distribution of the SINR at the typical location from the *nearest* base stations (one per tier) in the accessible tiers of the HetNet (with the locations of the base stations in the tiers modeled by independent homogeneous Poisson processes) can be obtained in exact analytic form (Mukherjee 2012).

The above selection of results concerns a HetNet where the locations of base stations in the tiers are modeled as independent homogeneous Poisson processes. Although the Poisson model has great advantages for analytical tractability, and is also the asymptotic limit of an arbitrary base station layout, it is not necessarily the best model for a finite layout of base stations.

Fortunately, we note that although the model of independent Poisson point process tiers for the HetNet does not fit real-world SINR distributions exactly, in several practically important scenarios it is possible to “shift” the Poisson process curves by a known amount to obtain excellent agreement with real-world measurements. This shift depends only on a quantity called the *mean interference-to-signal ratio* (MISR). The resulting method, called *approximate signal-to-interference ratio analysis based on the Poisson point process*, provides a quick and computationally efficient scheme that is eminently suitable for engineering applications (Guo & Haenggi 2015). As one example, we also study the use of the MISR in approximating the ergodic spectral efficiency of links under advanced transmission techniques such as interference alignment.