

# PERFORMANCE STUDIES OF WIRELESS MULTIHOP NETWORKS

Henri Koskinen

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission for public examination and debate in Auditorium S4 at Helsinki University of Technology (Espoo, Finland) on the 12th of May, 2006, at 12 o'clock noon.

Helsinki University of Technology  
Department of Electrical and Communications Engineering  
Networking Laboratory

Teknillinen korkeakoulu  
Sähkö- ja tietoliikennetekniikan osasto  
Tietoverkkolaboratorio

Distribution:

Helsinki University of Technology

Networking Laboratory

P.O.Box 3000

FIN-02015 TTK

Tel. +358-9-451 2461

Fax. +358-9-451 2474

© Henri Koskinen

ISBN 951-22-8136-8

ISBN 951-22-8137-6 (PDF)

ISSN 1458-0322

Otamedia Oy

Espoo 2006



HELSINKI UNIVERSITY OF TECHNOLOGY P.O.BOX 1000, FIN-02015 TKK <a href="http://www.tkk.fi/">http://www.tkk.fi/</a>		ABSTRACT OF DOCTORAL DISSERTATION	
Author: Henri Koskinen			
Name of the dissertation Performance Studies of Wireless Multihop Networks			
Date of manuscript 5th of December 2005		Date of the dissertation 12th of May 2006	
<input type="checkbox"/> Monograph		<input checked="" type="checkbox"/> Article dissertation (summary + original articles)	
Department	Department of Electrical and Communications Engineering		
Laboratory	Networking Laboratory		
Field of Research	Teletraffic Theory		
Opponent	Professor Patrick Thiran (École Polytechnique Fédérale de Lausanne, Switzerland)		
Supervisor	Professor Jorma Virtamo (Helsinki University of Technology)		
Abstract			
<p>Wireless multihop networks represent a fundamental step in the evolution of wireless communications, a step that has proven challenging. Such networks give rise to a wide range of novel performance and design problems, most of which are of a geometric nature. This dissertation addresses a selection of such problems.</p> <p>The first part of this thesis presents studies in which the network nodes are assumed to receive signals sufficiently clearly only from within some fixed range of operation. Using this simple model, the first two problems addressed are to predict the probabilities that a network with randomly placed nodes is connected or completely covers a given target domain, respectively. These problems are equivalent to determining the probability distribution of the minimal range providing connectivity or coverage. Algorithms for determining these threshold ranges for a given set of network nodes are developed. Because of the complex nature of these problems in finite settings, they are both approached by empirically modeling the convergence of these distributions to their known asymptotic limits. Next, a novel optimization problem is presented, in which the task is to make a given disconnected network into a connected one by adding a minimal number of additional nodes to the network, and heuristic algorithms are proposed for this problem.</p> <p>In the second part, these networks are studied in the context of a more realistic model in which the condition for successful communication between network nodes is expressed as an explicit minimum value for the received signal-to-noise-and-interference ratio. The notion of the threshold range for connectivity is first generalized to this network model. Because connectivity is now affected by medium access control (MAC), two alternative MAC schemes are considered. Finally, an infinite random network employing slotted Aloha is studied under this model. Since the probability of successful reception in a random time slot is a function of the locations of other nodes, this temporal probability is a random variable with its own probability distribution over different node configurations. Numerical approximations for evaluating both the mean and the tail probability of this distribution are developed. The accuracy of these approximations can be improved indefinitely, at the cost of numerical computations.</p>			
Keywords	wireless multihop networks, ad hoc networks, sensor networks, connectivity, coverage, throughput, geometric random graphs, stochastic geometry		
ISBN (printed)	951-22-8136-8	ISSN (printed)	1458-0322
ISBN (pdf)	951-22-8137-6	ISSN (pdf)	
ISBN (others)		Number of pages	84 p. + app. 84 p.
Publisher	Networking Laboratory / Helsinki University of Technology		
Print distribution			
<input checked="" type="checkbox"/> The dissertation can be read at <a href="http://lib.tkk.fi/Diss/">http://lib.tkk.fi/Diss/</a>			





TEKNILLINEN KORKEAKOULU PL 1000, FIN-02015 TKK <a href="http://www.tkk.fi/">http://www.tkk.fi/</a>	VÄITÖSKIRJAN TIIVISTELMÄ
Tekijä: Henri Koskinen	
Väitöskirjan nimi Tutkimuksia langattomien monihyppyisten verkkojen suorituskyvystä	
Käsi kirjoituksen jättämispäivämäärä 5. joulukuuta 2005	Väitöstilaisuuden päivämäärä 12. toukokuuta 2006
<input type="checkbox"/> Monografia	<input checked="" type="checkbox"/> Yhdistelmäväitöskirja (yhteenvedo + erillisartikkelit)
Osasto Sähkö- ja tietoliikennetekniikan osasto Laboratorio Tietoverkkolaboratorio Tutkimusala Teleliikenneteoria Vastaväittäjä Professori Patrick Thiran (École Polytechnique Fédérale de Lausanne, Sveitsi) Työn valvoja Professori Jorma Virtamo (Teknillinen korkeakoulu)	
Tiivistelmä <p>Langattomat monihyppyiset verkot edustavat langattoman viestinnän perustavanlaatuista kehitysaskelta, joka on osoittautunut haasteelliseksi. Tällaisissa verkoissa tulee esiin monenlaisia uusia suorituskyky- ja suunnitteluongelmia, joista useimmat ovat luonteeltaan geometrisia. Tässä väitöskirjassa käsitellään joitakin tällaisia ongelmia.</p> <p>Työn ensimmäinen osa esittelee tutkimuksia, joissa verkon solmujen oletetaan vastaanottavan signaaleja riittävän voimakkaana vain tietyn kiinteän toimintakantaman sisältä. Tämän yksinkertaisen mallin puitteissa kaksi ensimmäistä tarkasteltavaa ongelmaa ovat ennustaa todennäköisyydet sille, että satunnaisesti sijaitsevien solmujen muodostama verkko on yhtenäinen ja toisaalta täysin peittää annetun kohdealueen. Näiden kanssa ekvivalentit ongelmat on määrittää yhtenäisyyden tai täyden peiton tuovan pienimmän kantaman todennäköisyysjakauma. Työssä kehitetään algoritmeja, joilla nämä kynnyskantamat määritetään annetulle solmujoukolle. Koska ongelmat ovat luonteeltaan monimutkaisia äärellisissä tapauksissa, lähestytään kumpaakin mallintamalla empiirisesti näiden jakaumien suppenemista kohti tunnettuja asymptoottisia rajajakaumiaan. Seuraavaksi esitetään uusi optimointiongelma, jossa tehtävänä on tehdä annetusta epäyhtenäisestä verkosta yhtenäinen sijoittamalla verkkoon mahdollisimman vähän lisäsolmuja, ja kehitetään ongelmaan heuristisia algoritmeja.</p> <p>Toisessa osassa näitä verkkoja tutkitaan käyttäen realistisempaa mallia, jossa verkon solmujen välisen viestinnän onnistumisehto ilmaistaan vastaanotetun signaali-häiriösuhteen minimiarvona. Aluksi yleistetään verkon yhtenäisyyden kynnyskantaman käsite tähän verkkomalliin. Koska verkon yhtenäisyyden vaikuttaa nyt solmujen lähetyksuri (medium access control; MAC), tarkastellaan kahta vaihtoehtoista MAC-protokollaa. Lopuksi tutkitaan äärettömästi satunnaista, aikajaettua Aloha-satunnaissiirtymää käyttävää verkkoa tämän mallin puitteissa. Koska lähetyksen vastaanoton onnistumistodennäköisyys satunnaisessa aikavälissä on muiden solmujen sijaintien funktio, tämä ajallinen todennäköisyys on satunnaismuuttuja, jolla on oma todennäköisyysjakaumansa yli erilaisten solmukonfiguraatioiden. Tämän jakauman odotusarvon ja häntätodennäköisyyden arvioimiseksi kehitetään numeerisia approksimaatioita, joiden tarkkuutta voidaan parantaa rajatta, numeerisen laskentatyön kustannuksella.</p>	
Avainsanat langattomat monihyppyiset verkot, spontaanit verkot, anturiverkot, verkon yhtenäisyys, verkon peitto, läpäisy, geometriset satunnaisgraafit, stokastinen geometria	
ISBN (painettu) 951-22-8136-8	ISSN (painettu) 1458-0322
ISBN (pdf) 951-22-8137-6	ISSN (pdf)
ISBN (muut)	Sivumäärä 84 s. + liit. 84 s.
Julkaisija Tietoverkkolaboratorio / Teknillinen korkeakoulu	
Painetun väitöskirjan jakelu	
<input checked="" type="checkbox"/> Luettavissa verkossa osoitteessa <a href="http://lib.tkk.fi/Diss/">http://lib.tkk.fi/Diss/</a>	



## PREFACE

This dissertation is the result of research that began in 2002, when I started to work on my Master's Thesis in the Networking Laboratory of Helsinki University of Technology. The work has been carried out in the AHRAS project, funded by the Finnish Defence Forces Technical Research Centre, as well as in the NAPS project funded by the Academy of Finland. In 2004, I was granted the honor of being admitted to the Graduate School of Electronics, Telecommunications and Automation (GETA), which has since provided the majority of the funding. In addition, personal financial support from the TES and Nokia foundations is gratefully acknowledged.

I wish to express my gratitude to my supervisor, Professor Jorma Virtamo, for the privilege of receiving his guidance throughout my work. His solid expertise in research never ceases to impress me. I also thank Drs. Esa Hyytiä, Jouni Karvo, and Pasi Lassila, as well as M.Sc. Olli Apilo, for their work as co-authors of the publications, and the pre-examiners of this dissertation, Professor Christian Bettstetter and Dr. Bartłomiej Błaszczyszyn, for their time, effort, and constructive comments.

My thanks also go to the rest of the people in the lab for creating such a laid-back working atmosphere, especially coordinator Arja Hänninen and secretaries Raija Halkilahti, Sanna Patana, and Irma Planman, who have all in turn kept things running, and in particular my office mate, Aleksi Penttinen, for sharing both his hysterical sense of humor and those moments of frustration with the beeping Mathematica.

I am also deeply grateful to my dear parents Mauri and Pirjo for their perpetual support, and to my sister and her “own family”, Reetta, Petri, and Pyry, for my delightful role as an uncle. The Ressu gang, my study mates at TKK, and all my other friends also deserve special thanks for being who they are.

Finally, I thank you, Virpi, for the love.

*Henri Koskinen*





# CONTENTS

Preface	i
Contents	iii
List of publications	v
<b>1 Introduction</b>	<b>1</b>
1.1 Wireless multihop networks	1
1.2 Performance problems in wireless multihop networks	1
Connectivity	1
Coverage	3
Throughput and capacity	4
Medium access control	5
Other problems	5
1.3 Common models for wireless multihop networks	6
Homogeneous Poisson point process	6
Boolean models	6
Physical model	7
Random Waypoint (RWP) mobility model	8
1.4 Structure and contribution of this thesis	8
<b>2 Studies under Boolean models</b>	<b>11</b>
2.1 Connectivity of random networks	11
Problem statement	12
Review of existing results	12
Connectivity probability as a learning problem	14
Algorithms for finding $R_k(\mathcal{N})$	14
Empirical models	16
Connectivity probability under the RWP mobility model	21
2.2 Coverage of random networks	23
Problem statement	23
Review of existing results	24
Expected area coverage	25
Probability of complete coverage as a learning problem	26
2.3 Connectivity improvement as an optimization problem	31
Problem statements	31
Review of related problems and existing results	32
Heuristic algorithms	32
2.4 Summary and Conclusions	37
<b>3 Studies under the Physical model</b>	<b>41</b>
3.1 Connectivity	41
Review of existing results	41
Studying graph connectivity under physical models	42
Connectivity boundary in a CDMA network	43

	Connectivity in a slotted-Aloha network . . . . .	44
3.2	Throughput . . . . .	48
	Review of existing results . . . . .	48
	Probability of successful transmission in a random slotted- Aloha network . . . . .	49
	Problem statement . . . . .	50
	Probability of successful reception: expected value . . . . .	51
	Probability of successful reception: distribution . . . . .	56
3.3	Summary and Conclusions . . . . .	60
4	Summaries of publications and author's contributions	63
Appendix A	Supplementary material for publications	65
Appendix B	Erratum	69
	References	71
	Publications	79

## LIST OF PUBLICATIONS

- [1] Henri Koskinen. A simulation-based method for predicting connectivity in wireless multihop networks. *Telecommunication Systems*, 26(2-4): pages 321–338, June 2004.
- [2] Henri Koskinen. Quantile models for the threshold range for  $k$ -connectivity. In *MSWiM '04: Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems*, pages 1–7, October 2004. ACM Press, New York, NY, USA.
- [3] Pasi Lassila, Esa Hyytiä, and Henri Koskinen. Connectivity properties of Random Waypoint mobility model for ad hoc networks. In *Proceedings of the Fourth Annual Mediterranean Workshop on Ad Hoc Networks (Med-Hoc-Net)*, June 2005. 10 pages, printed proceedings to appear.
- [4] Henri Koskinen. On the coverage of a random sensor network in a bounded domain. In *Proceedings of the 16th ITC Specialist Seminar*, pages 11–18, August 2004.
- [5] Henri Koskinen, Jouni Karvo, and Olli Apilo. On improving connectivity of static ad-hoc networks by adding nodes. In *Proceedings of the Fourth Annual Mediterranean Workshop on Ad Hoc Networks (Med-Hoc-Net)*, June 2005. 10 pages, printed proceedings to appear.
- [6] Henri Koskinen. Generalization of critical transmission range for connectivity to wireless multihop network models including interference. In *Proceedings of the Third IASTED International Conference on Communications and Computer Networks (CCN)*, pages 88–93, October 2005.
- [7] Henri Koskinen and Jorma Virtamo. Probability of successful transmission in a random slotted-Aloha wireless multihop network employing constant transmission power. In *MSWiM '05: Proceedings of the 8th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems*, pages 191–199, October 2005. ACM Press, New York, NY, USA.



# 1 INTRODUCTION

## 1.1 Wireless multihop networks

A wireless multihop network refers to a network formed independently by mobile, wireless terminal devices without the aid of any fixed infrastructure. Communication over this kind of network occurs in a decentralized way, with the devices (or, henceforth, network nodes) relaying each other's traffic and connections between node pairs thus being formed over multiple transmission hops. Thus, besides being terminals, the network nodes also function as routers.

The application of wireless multihop networks is generally divided into two scenarios: *sensor networks* consisting of dedicated devices that provide monitoring or measurement data on their surroundings, and *ad hoc* networks formed anywhere and at any time, with communication as the primary purpose. The latter term is often used synonymously with wireless multihop networks in the literature.

Because of their intrinsic properties, wireless multihop networks have remained a challenge for commercial implementation: although, e.g., current Wireless LAN cards feature the ability to operate amongst themselves in an ad hoc mode, they lack the multihop functionality. However, this has far from discouraged the research community: wireless multihop networks have been studied with constantly increasing activity, both before and since the Internet Engineering Task Force (IETF) created the MANET working group (short for Mobile Ad-Hoc NETworks) as a forum for the study of this area in 1999 [MAN05].

## 1.2 Performance problems in wireless multihop networks

The nature of wireless multihop networks poses completely new performance and design problems. In this section, we introduce some characteristic problems, most of which we will be discussing further in this thesis.

### Connectivity

Due to the nature of wireless links and the unconstrained locations of network nodes, the topology of these networks is dynamic. This leads to perhaps the most fundamental problem, namely, the requirement that the nodes form a single connected network that allows them to communicate with each other in a multihop fashion.

The natural analytical framework for studying connectivity is graph theory. However, applying graph theory to these networks requires a definition of when a single link is connected. This boils down to issues on the physical layer: the quality of reception of a radio transmission depends on the signal-to-noise-and-interference ratio (SINR) at the receiver. From an information theory point of view, any positive SINR makes successful communication possible; only the achievable rate of communication depends on the SINR

[Sha49]. From this viewpoint, any link and hence any given network can always be said to be connected.

To say that some link is not connected therefore requires us to set a minimum value for the SINR corresponding to a required minimum rate of communication. Such a minimum value may also be dictated by technical design choices, such as the existence of proper modulation and coding schemes within a given communication framework. We will return to this issue in the next section, where we introduce two network models, both based on this idea but one more detailed than the other.

Another choice to be made when harnessing graph theory is whether to consider unidirectional links, which may often exist in practice: transmissions are well received in one direction but not in the other. As in most studies, we choose not to consider unidirectional links, the reason being, e.g., that they render acknowledged forms of communication difficult. This allows us to concentrate on undirected graphs.

Finally, connectivity itself is also subject to definition. Throughout this thesis, we study connectivity as defined in graph theory, namely, by the requirement that all node pairs are connected by the network. We also study the generalized property of  $k$ -connectivity, which means that the network remains connected after the removal of any  $k - 1$  nodes. In other words, by  $k$ -connectivity we mean  $k$ -node-connectivity, as opposed to link connectivity, which characterizes resilience against the removal of links. For comparison, connectivity has also been identified with *percolation*, i.e., the existence of an unbounded connected component in infinite random networks; we will also review existing results on percolation in the coming chapters.

The assumptions outlined above lead to the modeling of the topology of wireless multihop networks by geometric random graphs. In the simplest case, such a graph contains an undirected edge between all node pairs – and those node pairs only – that are less than some predefined distance apart. As we will see in the next section, this case results from the Boolean network model. However, the earliest existing analytical results concern pure random graphs, or Erdős-Rényi random graphs [ER60], where every pair among a total of  $n$  vertices is connected by an undirected edge independently with some common probability,  $p$ . It has been shown that the probability that such a graph is connected tends to one asymptotically if and only if  $p(n)$  is such that  $n \cdot p(n) - \log n$  tends to infinity with  $n$ , i.e., if the expected degree of a vertex  $n \cdot p(n)$  increases faster than the logarithm of  $n$  (see, e.g., [Bol85]). This condition was later shown also to hold true for simple geometric random graphs on vertices distributed uniformly at random [GK98], a case whose asymptotic  $k$ -connectivity properties have since been derived fairly exhaustively [WY04]. Recently, it has been shown that in the general case where the existence of an edge is dictated by any probability function of the pair of vertex locations, the logarithmic growth of the expected degree is still a necessary condition for asymptotic connectivity [Far05].

Beside the simplest geometric random graph, the connectivity problem has recently been studied under increasingly diverse modeling assumptions. A log-normal radio model – which also falls under the above gen-

eralized random graphs – is used in [HM04], where it is found that, as a result of the randomness in the radio conditions, the connectivity behavior resembles more closely that of pure random graphs. The benefit of using randomly directed antennae for connectivity is investigated in [BHM05].

### Coverage

One application of wireless multihop networks is sensor networks, where the main purpose of the devices is to measure or monitor some property in their surroundings, and the multihop communication ability serves the secondary purpose of delivering the sensor data to some central entity; see [ASSC02] for a survey on sensor networks. In this context, the coverage of a sensor network measures how well it is able to monitor a given target area (as, e.g., in the case of motion sensors).

Because of the large variety of possible sensor applications, coverage, like connectivity, is subject to a choice of definition. In [MKPS01], it is generally characterized as the measure of the quality of service (or quality of surveillance) of a sensor network. Generalizing the Boolean network model to sensor coverage leads to the modeling of the coverage region of a sensor by a circular disk whose radius equals the sensing range of the sensor, which indeed is the underlying assumption in the majority of existing studies.

Research topics related to sensor coverage can roughly be divided into algorithmic problems and analytical work on coverage processes [Hal88]. Examples of the former include finding optimal paths through a given bounded sensor network; the *best-coverage* path minimizes the distance of all its points to the nearest sensor, while the *worst-coverage* path maximizes this distance for all points. Centralized algorithms for both problems utilizing the Voronoi diagram and its dual structure, the Delaunay triangulation, are presented in [MKPS01], whereas localized algorithms are given in [LWF02]. Polynomial-time distributed algorithms for the sensor nodes to decide whether the target domain of the network is  $k$ -covered, i.e., whether every point is covered by at least  $k$  sensors, have also been developed [HT03].

Given that the sensors have limited energy reserves, *density control* aims at maximizing the operational lifetime of the network by keeping a minimal subset of sensors active at each moment and letting the remaining sensors stand by in a low-power mode. One of the first studies addressing this problem under the constraint of preserving coverage is [TG02], and a recent, evolved algorithm for preserving both connectivity and coverage, along with an extensive survey of algorithmic sensor coverage problems, is given in [ZH05]. As in the latter study, many algorithms for maintaining connected coverage rely on the elementary relation that if a given set of sensors provides  $k$ -coverage of their target domain with a given common sensing range, then they form a  $k$ -connected network with a transmission range twice as great as the sensing range. Fundamental limits for the achievable lifetime in large random sensor networks are explored in [ZH04]. More results on the coverage of random networks will be reviewed in the next chapter.

Recent analytical results suggest that sensor mobility improves the coverage of sensor networks [LBD<sup>+</sup>05]. Indeed, several algorithmic studies

consider directing mobile sensors so as to maximize coverage (see, e.g., [WCP04] and the references therein).

### Throughput and capacity

The fact that the wireless medium must be shared among the network nodes, together with the dual role of the nodes as both routers and terminals in multihop networks, gives rise to a limitation as fundamental in nature as the problem of connectivity. In the seminal paper [GK00], it was shown that as the density of nodes increases in relation to the typical distance between communicating source-destination pairs (as is also the case with larger and larger networks with constant node density and random destinations for each node), the burden from relaying other nodes' traffic grows for each node, with the result that the throughput obtained by each node for its own traffic diminishes. Thus, if every node is assumed to add its own contribution to the total traffic demand, then, because of the restrictions imposed by the spatial aspect, wireless multihop networks do not enjoy the virtually unlimited scalability of other networks.

Stated more precisely, one of the main results in [GK00], further refined in [AK04], states that under a so-called Physical model of communication, where the bit rate extracted from receiving a transmission is some step function of the prevailing SINR, a wireless network of  $n$  nodes spanning a domain of area  $A$  is capable of transferring at most  $\Theta(\sqrt{An})$  bit-meters per second.<sup>1</sup> (We will introduce the Physical model in detail shortly; however, this result holds also for the Generalized Physical model, where the bit rate is the Shannon-capacity logarithmic function of the SINR ([AK04, Gup00]).) Thus, if all nodes generate traffic with some common constant rate, with destined receivers at average distance  $\Theta(\sqrt{A})$ , there are  $n$  flows requiring  $\Theta(\sqrt{A})$  bit-meters per second each, which yields a throughput of  $\Theta(1/\sqrt{n})$  bit/s for each flow. Naturally, if the domain area increases with  $n$ , i.e.,  $A = \Theta(n)$ , but communication remains local so that each flow requires  $\Theta(1)$  bit-meters per second, the per-flow throughput also remains at a constant level.

It is important to note that these results are not ultimate information-theoretic capacity limits. The assumptions behind the Physical model are based on the paradigms that dictate how current communication technology operates, but they are unnecessarily restrictive for ultimate limits for information transfer. For example, the assumption that all interference is essentially regarded as noise rules out the diverse possibilities of co-operative communication, such as active interference cancellation by some nodes in the network to improve the quality of reception of others. With this motivation, the aim in [XK04a] is to connect information theory to the world of networking, and find ultimate limits for how much information wireless networks can transport without making preconceived assumptions. Among the results in [XK04a], it is found that whenever the wireless medium is absorptive – which is generally the case – the transport capacity is upper bounded by a multiple of the total transmission power of all the nodes, which means that there is a lower bound on the energy price in joules

---

<sup>1</sup> $f(n) = \Theta(g(n))$  denotes that  $f(n) = O(g(n))$  and  $g(n) = O(f(n))$ .



per bit-meter of information transport. In addition, it is shown that at least in certain basic scenarios, the “multihop strategy” of coding and decoding packets successively hop by hop and letting concurrent transmissions be useless noise – which is a commonly agreed-upon assumption in current protocol development – is almost or completely order-optimal, meaning that it does not lead to drastically suboptimal scaling of transport capacity. This strategy is shown to be appropriate also for fading environments in [XXK05].

### Medium access control

One important design problem is medium access control (MAC). Although it is not so great an issue in wired networks, the solution used in wireless networks should allow the spatial reuse of the shared medium. Moreover, the time-varying network topology and the lack of centralized control in multihop networks render the use of coordinated MAC schemes difficult, making random access seem the preferred choice.

For later reference, probably the simplest random access protocol is Aloha [Abr70], in which network nodes transmit whenever they desire, and conflicts resulting from simultaneous transmissions destructively interfering are deduced from missing acknowledgements. Retransmissions are randomly delayed so as to avoid repeated collisions. The efficiency of this scheme is improved if transmissions are only allowed to occupy synchronized time slots; this is referred to as slotted Aloha [Abr73].

Another approach to accessing the medium, known as Carrier Sense Multiple Access (CSMA), is that network nodes determine whether or not to transmit by “listening” to any possible ongoing transmissions [KT75]. Although this is perfectly viable as such in wired networks, its implementation in wireless networks requires additional procedures around the receiver in order to overcome problems concomitant with the spatial aspect, such as so-called hidden and exposed terminals. CSMA is the basis for the medium access protocols used in WLANs [IEE99], while slotted Aloha is used, e.g., in the Random Access CHannel of GSM/GPRS [3GP05].

It has recently been shown that slotted Aloha, while making decentralized implementation possible, reaches the above upper bound for the scaling of transport capacity in wireless multihop networks [BBM04] under the Physical model. We will return to this matter in Chapter 3.

### Other problems

Among the performance and design problems that we will not discuss further in this thesis, one worth mentioning is *routing*. The dynamic topology of these networks makes the task of maintaining routing information challenging. A proactive approach may result in overwhelming signaling traffic as the information in a highly mobile network is updated, whereas relying only on reactive routing can lead to long connection set-up delays.

Finally, the fact that wireless devices are bound to have limited energy supplies calls for efficient *power management*: as we already mentioned, maximizing the operational lifetime of a sensor network is a widely-studied problem.

### 1.3 Common models for wireless multihop networks

This thesis presents the results of studies where the above performance problems are approached using diverse mathematical analysis techniques. As always when applying mathematics to complex real-life phenomena, this requires the object of interest to be represented with a model, stripped of all the complexity that is not essential – and perhaps, for the sake of tractability, even of some that may be essential. In this section, we introduce for later reference some popular models used in analyzing wireless multihop networks that our studies also rely on.

#### Homogeneous Poisson point process

As already made apparent by the problem of connectivity, the geographical locations of network nodes are an important factor affecting the performance of wireless multihop networks, which underlines the need to model these locations. To take into account the possibility of practically any configuration of nodes, the locations are usually treated as random. Furthermore, unless more specific information is given, it is reasonable to assume that, a priori, the locations are uniformly distributed.

To this end, let us assume that  $n$  nodes are randomly and independently located according to the uniform distribution over some bounded domain  $\mathcal{A} \subset \mathbb{R}^2$  with an area  $|\mathcal{A}| = A$  (generalization to a higher number of dimensions is straightforward). Then the number of nodes in any subdomain  $\mathcal{D} \subset \mathcal{A}$  is random, with the distribution  $\text{Bin}(n, |\mathcal{D}|/|\mathcal{A}|)$ . However, given the number of nodes  $\tilde{n}$  in any other non-intersecting subdomain  $\tilde{\mathcal{D}}$ , the conditional distribution is different; hence, the two numbers are not independent.

Keeping our attention on the arbitrarily selected domains  $\mathcal{D}$  and  $\tilde{\mathcal{D}}$ , let us consider the effect of letting the domain  $\mathcal{A}$  become larger and larger, while keeping the average node density  $n/A \stackrel{\text{def}}{=} \lambda$  constant. Since this makes  $|\mathcal{D}|/|\mathcal{A}|$ , the probability that an arbitrary node is in  $\mathcal{D}$ , diminish but keeps the expected number of nodes therein  $n|\mathcal{D}|/|\mathcal{A}|$  constant, in the limit  $n, A \rightarrow \infty$  the above binomial distribution tends to a Poisson distribution with the parameter  $\lambda|\mathcal{D}|$ . In the same limit, the number of nodes in  $\tilde{\mathcal{D}}$  depends less and less on  $\tilde{n}$  (because this leaves  $n - \tilde{n}$  nodes outside  $\tilde{\mathcal{D}}$ , which also tends to infinity).

The point process that results in the limit is the homogeneous Poisson point process, often denoted by  $\Phi$ . It can be interpreted as points “uniformly distributed” over the whole plane with average density  $\lambda$ . It is completely characterized by the following two properties:

1. The number of points of  $\Phi$  in a bounded set  $\mathcal{D}$  has a Poisson distribution of mean  $\lambda|\mathcal{D}|$  for some constant  $\lambda$ .
2. The numbers of points of  $\Phi$  in  $k$  disjoint sets form  $k$  independent random variables, for arbitrary  $k$ .

#### Boolean models

As mentioned earlier, applying graph theory to studying the connectivity of wireless multihop networks requires defining when a single link is con-

ected by setting a minimum value for the signal-to-noise-and-interference ratio (SINR).

Let us consider the case where we neglect interference altogether and instead only require some minimum value  $T > 0$  for the ratio of the received signal power to that of a constant-level ambient noise, i.e., the signal-to-noise ratio (SNR). Then any node  $j$  located at point  $x_j$  in Euclidean space is assumed to successfully decode the signal transmitted by another node  $i$  at  $x_i$  with power  $P_i$  if and only if

$$\frac{P_i l(\|x_i - x_j\|)}{N_0} \geq T,$$

where  $N_0$  is the power of the background noise on the frequency channel utilized by the network and  $l(\cdot)$  is some strictly decreasing attenuation function of propagated distance, i.e.  $l(\|x_i - x_j\|)$  gives the path loss in power for a signal propagated from point  $x_i$  to  $x_j$ .

The above condition is equivalent to

$$\|x_i - x_j\| \leq l^{-1} \left( \frac{N_0 T}{P_i} \right),$$

i.e., we may translate it into a maximum distance from the transmitter  $i$  within which the signal is successfully received, referred to as the *transmission range* of node  $i$ . This simple model, where the assumed condition for any two nodes being directly connected is that they are within each other's transmission ranges and hence only depends on these nodes, is known as the *Boolean model*.

A similar Boolean model can be used to study coverage in sensor networks, by assuming that each sensor covers a disk around it with a certain radius. This *sensing range* models the range within which a sensor must be from a given point in order for this point to be considered reliably monitored – or covered – by that sensor, perhaps to a predefined level of confidence. For an example of where such a notion of coverage is valid, one may think of motion sensors.

### Physical model

The Boolean model, by neglecting all interference, makes successful communication depend on the SNR rather than the SINR. The more accurate Physical model, first introduced in [GK00], explicitly takes into account interference from concurrent transmissions and replaces the above condition with

$$\frac{P_i l(\|x_i - x_j\|)}{N_0 + \sum_{k \neq i, j} P_k l(\|x_k - x_j\|)} \geq T.$$

Note that the inclusion of the interference term makes things considerably more complicated than is the case with the Boolean model: whether or not two nodes are able to communicate directly no longer depends only on their transmission ranges and the distance between them, but also on the locations of all other nodes and their *instantaneous* transmission powers  $P_k$ .

It was assumed in [GK00] that whenever the above condition holds, the rate of communication from node  $i$  to  $j$  is always the same, no matter how far beyond  $T$  the achieved SINR is; when referring to the Physical model, we also incorporate this assumption. (In contrast, the relaxed case where the bit rate is the Shannon logarithmic function of the SINR is named the Generalized Physical model in [AK04].)

### Random Waypoint mobility model

The mobility of nodes is integral to these networks and therefore also needs to be modeled somehow. One of the most popular mobility models used for wireless multihop networks, originally proposed for studying the performance of routing protocols for ad hoc networks in [JM96], is the Random Waypoint (RWP) model. In this model, a mobile node is assumed to move in a convex domain  $\mathcal{A}$  between successive waypoints drawn independently and randomly from  $\mathcal{A}$ ; again, the lack of more specific information makes uniform distribution a reasonable assumption. The leg between any two waypoints is traversed directly along a straight line segment, with a constant velocity that is also assumed to be an independent and identically distributed random variable for each leg. Furthermore, the node may be assumed to spend a random i.i.d. pause time at each waypoint. Finally, all the nodes in the network are assumed to move independently, governed by the same probability distributions for the waypoints, velocities, and pause times.

## 1.4 Structure and contribution of this thesis

This thesis gathers observations from various performance studies of wireless multihop networks. The first part comprises studies where these networks are modeled using Boolean models, and the second part treats them using the Physical model.

### Studies under Boolean models

We first discuss the connectivity of random networks. Determining the probability that a random network is  $k$ -connected is equivalent to knowing the distribution of the threshold range for  $k$ -connectivity. These distributions are known only asymptotically, as the number of nodes in the network tends to infinity. The joint contribution of Publications [1] and [2] is an approach for predicting these distributions for finite configurations, based on empirical models that describe the convergence of observed (simulated) distributions to the known asymptotic ones. In Publication [1], we present algorithms to determine the threshold range for a given set of nodes. These algorithms facilitate simulations, on the basis of which we also present initial, purely empirical models that do not yet take into account asymptotic distributions. The prior information regarding limit distributions is then taken as the basis for these models in Publication [2], resulting in good predictive power. Finally, in Publication [3] we use the qualitative implications of asymptotic distributions to approximate the probability of connectivity in a mobile network where the nodes move according to the Random Waypoint mobility model.

We then move on to the coverage problem of random networks. In Publication [4], we first note that the covered fraction of a bounded domain is a random variable and determine the expected coverage in a simple circular domain. We then point out that the problem of determining the probability of complete coverage, like that of connectivity, is also equivalent to knowing the distribution of a well-defined threshold range; we show how this threshold range can be determined for a given set of nodes. Interpreting previous analytical results as an asymptotic distribution of this threshold range for complete coverage, we generalize the approach of predicting this distribution for finite configurations by using empirical models. In a case where the limit distribution is not known as a result of a complex border effect, we derive an approximation for the asymptotic distribution.

Finally, we study connectivity as an optimization problem. Publication [5] presents a novel problem where a given disconnected network is to be made into a connected one by adding additional nodes to the network; the objective is to minimize the number of nodes added. We point out some connections of this problem to existing problems that are NP-hard and present gradually better-performing heuristic algorithms, along with their complexity analysis.

### **Studies under the Physical model**

So far the percolation properties of infinite random networks have been studied under the Physical model [DFM<sup>+</sup>06, DT04], but little has been done to address connectivity as defined in graph theory when taking interferences into account. In Publication [6], we generalize the notion of the threshold range for connectivity to networks under the Physical model. Connectivity is now affected by the medium access scheme used in the network, through the time-varying interference; we consider two scenarios from existing studies. Because there is now more than one free parameter in the network, the threshold range generalizes into a boundary in the space of these parameters that implies tradeoffs between different performance quantities.

It has recently been shown that the optimal throughput scaling under the Physical model can be achieved when the medium access control is handled using slotted Aloha [BBM04] (see also the journal version [BBM06]). However, the quantitative results in this study are based on the assumption that the transmission powers in the network are exponentially distributed and hence unbounded. As our contribution in Publication [7], we extend the analysis of the proposed scenario: assuming that all nodes use some common constant transmission power, we develop numerical approximations for determining the probability of successful transmission in an infinite random network. We point out that this probability is a function of the locations of all surrounding nodes and therefore a random variable; we address both the expected value and the tail probability of its distribution.



## 2 STUDIES UNDER BOOLEAN MODELS

*This chapter describes studies that all rely on Boolean models. The first two sections focus on the connectivity and coverage, respectively, of random networks. The final section presents an optimization problem of connectivity.*

### 2.1 Connectivity of random networks

We begin by discussing the problem of connectivity of random wireless multihop networks when their topology is modeled using the Boolean model. Before proceeding to the connectivity problem that we focus on, we briefly review some related problem variations and previous work on them.

One problem studied under the Boolean model, motivated by the need for distributed topology control, is the following. Assume that every node in the network, by adjusting its transmission power, sets its transmission range equal to the distance to its  $m$ -th nearest neighbor, so that, taking only bidirectional links into account, any two nodes are directly connected if and only if they are both one of each other's  $m$  nearest neighbors. The problem is to find such  $m$  that a network of  $n$  nodes uniformly and independently distributed in a domain with a simple shape is connected with high probability. Ending a long series of studies proposing different constants, or “magic numbers”, it was shown in [XK04b] that  $m$  must grow like the logarithm of the number of nodes, and explicit numerical, asymptotically almost sure lower and upper bounds for the multiplying constant involved were derived. In [WY04], the upper bound was improved and was furthermore shown to hold for  $k$ -connectivity in general.

The above problem statement deviates from the mainstream of the existing literature in that the majority of studies are based on the assumption that all nodes have the same transmission range. This assumption can be motivated by thinking of the common transmission range as resulting from a common maximum transmission power that the nodes can achieve, which can be deemed reasonable in many cases. Thus, such a range models the distance over which other nodes can be reached if need be, allowing us to address ultimate limits for connectivity.

In this spirit, assuming that all nodes have some common transmission range  $r$ , the connectivity of infinite networks has also been studied, in the setting where the nodes of the network are located at the points of a homogeneous Poisson point process with some intensity  $\lambda$  in the infinite plane. This rules out studying graph-theoretical connectivity, since for any finite range  $r$ , there always exist isolated nodes with no other nodes within this range. Instead, the connectivity problem is then related to percolation theory. A fundamental theorem from continuum percolation [MR96] states that there exists a finite critical value of the relative intensity  $\lambda r^2$  (which is scale-independent) below which all connected components in the network are almost surely bounded, whereas a unique unbounded connected component almost surely exists above this critical value. The probability that an

arbitrary node belongs to the infinite component is referred to as the percolation probability. The exact value of the critical relative intensity and the explicit expression for the percolation probability are still open problems. Percolation in such infinite networks was first studied in [Gil61] and later on, e.g., in [PPT89, DTH02].

In [PPT89], Philips et al. studied the graph connectivity of finite networks with a common transmission range. Assuming the Poisson process marking the node locations is restricted to a square with area  $A$ , it was shown that the expected number of direct neighbors of a node,  $\lambda\pi r^2$ , must grow logarithmically with the network area to ensure, in the limit, a moderate probability of network connectivity. This setting is in fact very close to the problem that we study and will present next; the only difference is that the number of nodes in our problem statement is given, not random.

### Problem statement

The problem that we examine closer is also based on the assumption that all nodes in the network have some common transmission range  $r$ , i.e., any two nodes are directly and bidirectionally connected if and only if they are within distance  $r$  from each other.

We assume that the network consists of  $n$  nodes located independently and randomly in some bounded, connected domain  $\mathcal{D}$  in  $d$ -dimensional Euclidean space with  $d > 1$ , and that the locations are identically distributed according to some probability density function  $f_{\mathcal{D}}(\cdot)$  over  $\mathcal{D}$ . If  $\mathcal{N} = \{X_i \in \mathcal{D} \mid i = 1, 2, \dots, n\}$  denotes the set of random node locations, then by the assumption of one common transmission range  $r$  for the nodes, the network topology can be represented by an undirected geometric graph  $\mathcal{G}(\mathcal{N}, \mathcal{E}(\mathcal{N}, r)) = \mathcal{G}(\mathcal{N}, r)$  with vertex set  $\mathcal{N}$  and edge set  $\mathcal{E}(\mathcal{N}, r) = \{(X_i, X_j) \mid X_i, X_j \in \mathcal{N}, i \neq j, \|X_i - X_j\| \leq r\}$ .

The problem is then stated as follows.

*Given  $n$ ,  $f_{\mathcal{D}}(\cdot)$  and  $r$ , what is the probability that the network, i.e., the random geometric graph  $\mathcal{G}(\mathcal{N}, r)$ , is  $k$ -connected?*

This problem can also be stated in an alternative but equivalent form. To this end, let  $R_k(\mathcal{N})$  denote the smallest transmission range  $r$  with which the graph  $\mathcal{G}(\mathcal{N}, r)$  with given node locations  $\mathcal{N}$  is  $k$ -connected; we refer to  $R_k(\mathcal{N})$  as the *threshold range for  $k$ -connectivity* (also often called the *critical range* in literature). Then the event  $\{\mathcal{G}(\mathcal{N}, r) \text{ is } k\text{-connected}\}$  is equivalent to  $\{r \geq R_k(\mathcal{N})\}$ , whence the probability of interest to us is equal to the cumulative distribution function of  $R_k(\mathcal{N})$  (with given  $n$  and  $f_{\mathcal{D}}(\cdot)$ ), evaluated at  $r$ . Therefore, being able to answer the above question with any  $r$  reduces to knowing the probability distribution of  $R_k(\mathcal{N})$  with given  $n$  and  $f_{\mathcal{D}}(\cdot)$ .

### Review of existing results

For a finite number of nodes  $n$ , the distribution of  $R_k(\mathcal{N})$  is not known even in the simplest cases such as uniform  $f_{\mathcal{D}}(\cdot)$  on a domain  $\mathcal{D}$  with a simple shape; all the existing precise analytical results are asymptotic in nature. Consequently, the tools used to address the problem in the finite case can be divided into analytical approximations and empirical methods.



In [Pen97], Penrose proved the following theorem for  $R_1(\mathcal{N})$  which, as pointed out in [SMH99], is equal to the greatest edge-length in the Euclidean minimum spanning tree of  $\mathcal{N}$ . In analogy with  $R_k(\mathcal{N})$ , let  $M_k(\mathcal{N})$  denote the threshold range for minimum degree  $k$ , i.e., the smallest transmission range  $r$  with which every vertex in the graph  $\mathcal{G}(\mathcal{N}, r)$  has degree at least  $k$ .

**Theorem 2.1** [Pen97] For uniform  $f_{\mathcal{D}}(\cdot)$  on  $\mathcal{D} = [0, 1]^2$ ,

$$\lim_{n \rightarrow \infty} \Pr[R_1(\mathcal{N}) = M_1(\mathcal{N})] = 1.$$

This implies that  $R_1(\mathcal{N})$  has the same asymptotic distribution as  $M_1(\mathcal{N})$ : in [DH89], for the same  $f_{\mathcal{D}}(\cdot)$  and  $\mathcal{D}$ , the distribution of  $n\pi M_1(\mathcal{N})^2 - \log n$  has been shown to converge weakly to the Gumbel distribution. A similar but weaker result has been derived by Gupta and Kumar in [GK98], stating that for uniform  $f_{\mathcal{D}}(\cdot)$  when  $\mathcal{D}$  is a unit-area disk, the probability  $\Pr[r(n) \geq R_1(\mathcal{N})]$  tends to one if and only if  $r(n)$  is such that  $n\pi r(n)^2 - \log n \rightarrow \infty$ .

As explained in [Pen97], Theorem 2.1 means that the longest edge is likely to be the same for the Euclidean minimum spanning tree as for the nearest-neighbor graph, and the qualitative meaning of the Gumbel distribution for  $n\pi M_1(\mathcal{N})^2 - \log n$  is that the asymptotics for  $M_1(\mathcal{N})$  are as if the nearest-neighbor distances of the points  $\mathcal{N}$  were independent. Penrose conjectures these two properties to hold for more general distributions  $f_{\mathcal{D}}(\cdot)$ : in [Pen98], he shows them to hold when  $f_{\mathcal{D}}(\cdot)$  is the standard  $d$ -dimensional normal distribution.

In [Pen99], Penrose generalized Theorem 2.1 to hold for  $R_k(\mathcal{N})$  and  $M_k(\mathcal{N})$  with any  $k > 1$ , in the unit cube in  $d$  dimensions with any  $d > 1$ . However, the exact asymptotic distribution of  $R_k(\mathcal{N})$ ,  $k > 1$ , remained undetermined because of a dominating *border effect*: the complicated effect of nodes near the boundary of  $\mathcal{D}$  having fewer direct neighbors, discussed already in [DH90a].

The boundary was successfully analyzed and the asymptotic distribution of  $R_k(\mathcal{N})$  for all  $k$  thus derived only recently in [WY04], for uniform  $f_{\mathcal{D}}(\cdot)$  when  $\mathcal{D}$  is both the unit-area square and the unit-area disk: it turns out that when  $k = 1$  and the border effect does not dominate, the distribution is the same in both domains, whereas it is different in the two domains when  $k > 1$ . Thus, for uniform  $f_{\mathcal{D}}(\cdot)$ , the asymptotic distribution of  $R_1(\mathcal{N})$  is at least to some extent independent of the shape of the domain  $\mathcal{D}$ , which is no longer the case for  $k > 1$ .

Extensive work on the subproblem of determining such transmission range  $r$  that results in a  $k$ -connected network with a predefined, high probability, has been done by Bettstetter. The theoretical basis for his approach is Theorem 2.1 and its generalization to  $k > 1$ , along with the above-mentioned asymptotic independent-like statistics of nearest-neighbor distances and the assumption of this holding in general for  $k$ -nearest-neighbor distances. These assumptions make way to approximating the probability of a  $k$ -connected network simply by the probability of a random node in the network having at least  $k$  other nodes within range, raised to the power  $n$ . In [Bet02], this was computed without consideration of the border effect, a shortcoming remedied with Zangl in [BZ02]. Further applications of the

approach when  $f_{\mathcal{D}}(\cdot)$  is the stationary node location distribution of the Random Waypoint mobility model or the normal distribution are demonstrated in [Bet04]. In short, the approach gives reasonably tight lower bounds for the required range when the target probability is at least 99% and the number of nodes is at least on the order of 100.

The approach taken by D’Souza et al. to this problem in [DRL03] is in the same spirit as the one we will be discussing next. Studying the distribution of  $R_1(\mathcal{N})$  by simulation when the nodes are uniformly distributed in a square region, their aim was to see whether determining the distribution with various  $n$  had predictive power, by modeling the behavior of the mean of the distribution as a function of  $n$ . The estimated parameter characterizing the model for the mean was the asymptotic value of  $R_1(\mathcal{N})$  as  $n$  tends to infinity while the node density remains fixed — but this is in contradiction with the result by Philips et al. mentioned in the beginning of this section which implies that  $R_1(\mathcal{N})$  has no such finite limit.

### Connectivity probability as a learning problem

Exact analytical determination of the  $k$ -connectivity probability or, equivalently, the distribution of the threshold range for  $k$ -connectivity  $R_k(\mathcal{N})$ , in finite networks is complicated: unlike in the asymptotic limit, the nearest-neighbor distances cannot be treated as independent, and minimum degree  $k$  does not imply  $k$ -connectivity with high probability.

Because analytical treatment of these complicated phenomena is daunting, we opt to encapsulate them in an empirical model. The purpose of this model is to describe how the distribution of  $R_k(\mathcal{N})$  changes with  $n$ , with the aim of allowing the prediction of this distribution over as wide a range of different  $n$  as possible. Thus, we approach our problem as that of *learning*, by which we mean improving our knowledge with the aid of observed data.

In our case, this observed data consists of samples of  $R_k(\mathcal{N})$  determined from a large number of simulated random realizations of  $\mathcal{N}$ , with different numbers of nodes  $n$ . Our data acquisition therefore requires algorithms that determine  $R_k(\mathcal{N})$  for given input  $\mathcal{N}$ ; in Publication [1], we present such algorithms. Detailed algorithms are described for  $k = 1, 2, 3$ , but the general principle is applicable for any  $k$ .

The next section will briefly describe these algorithms, and the empirical models for the distribution of  $R_k(\mathcal{N})$  are discussed in the section that follows.

### Algorithms for finding $R_k(\mathcal{N})$

All our algorithms treat the input  $\mathcal{N}$  as a fully connected Euclidean graph, in which the number of edges is quadratic in  $n$ . Since, for example, the Euclidean minimum spanning tree for  $n$  points in the plane can be computed in  $O(n \log n)$  time by utilizing the Delaunay triangulation [Aur91], the strength of our algorithms lies rather in their very simple implementation and effective operation with small  $n$  than in good scaling to large  $n$ .

The idea in the algorithms is to find the threshold range incrementally. For a given  $\mathcal{N}$ , the initial range  $r_0$  is chosen so that the geometric graph

$\mathcal{G}(\mathcal{N}, r_0)$  satisfies the necessary conditions for  $k$ -connectivity. Finally, the range is increased if required, to satisfy the sufficient condition as well. The necessary conditions are that  $\mathcal{G}(\mathcal{N}, r_0)$  is  $(k - 1)$ -connected and has minimum degree  $k$ ; the sufficient condition is that there is no  $(k - 1)$ -tuple of nodes  $\mathcal{T}$  whose removal would disconnect the graph. The smallest range needed to eliminate a given such  $\mathcal{T}$  equals  $R_1(\mathcal{N} \setminus \mathcal{T})$ .

For  $k = 1$ , the above descriptions of  $(k - 1)$ -connectivity and  $(k - 1)$ -tuples of nodes are naturally not defined, and the steps of the algorithm can be described very concisely:

1. Set the range to  $r = M_1(\mathcal{N})$  and find the connected components of  $\mathcal{G}(\mathcal{N}, r)$ .
2. If there is only one connected component,  $R_1(\mathcal{N}) = r$ . Else, treating the components as single elements making up the set  $\mathcal{N}$  and defining the distance between two components as the shortest node distance between them, go to step 1.

Figure 2.1 illustrates an example network after two rounds of the above steps. This algorithm can also be seen as a simplified variation of Boruvka's algorithm for finding the minimum spanning tree [NMN01], where we only keep record of the greatest edge length in the tree formed on each round, instead of the tree itself.

The algorithm generalized to  $k > 1$  can be summarized as follows:

1. Set the initial range to  $r_0 = \max\{M_k(\mathcal{N}), R_{k-1}(\mathcal{N})\}$ .
2. Find all the  $(k - 1)$ -tuples of nodes  $\mathcal{T}_i$  whose removal would disconnect  $\mathcal{G}(\mathcal{N}, r_0)$ .
3. If no such  $\mathcal{T}_i$  were found,  $R_k(\mathcal{N}) = r_0$ , else  $R_k(\mathcal{N}) = \max_i\{R_1(\mathcal{N} \setminus \mathcal{T}_i)\}$ .

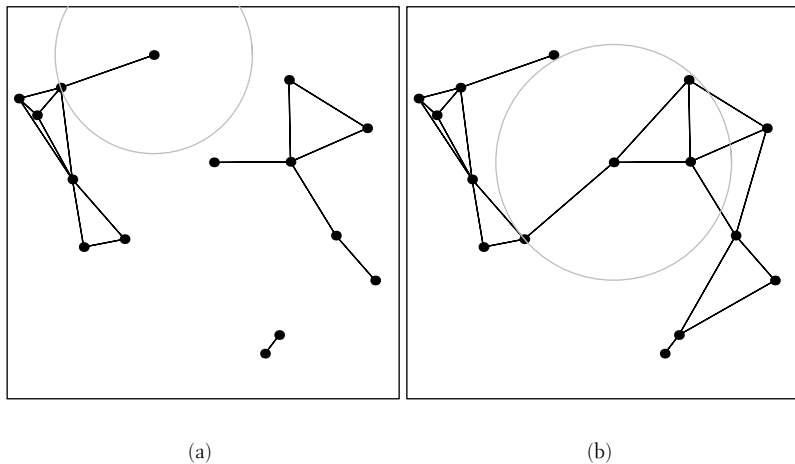


Figure 2.1: The network formed by a sample set  $\mathcal{N}$  of 15 nodes with the range (a):  $M_1(\mathcal{N})$  and (b):  $R_1(\mathcal{N})$ .

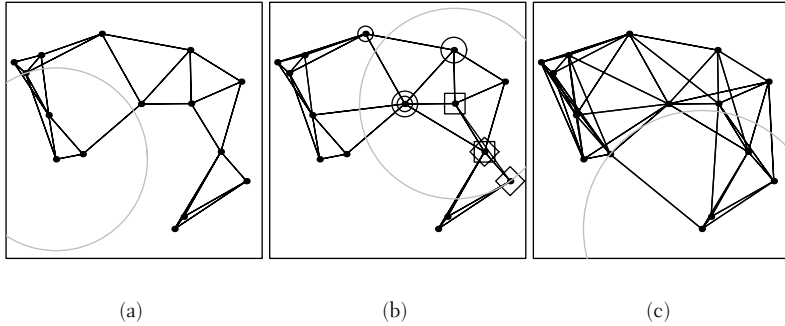


Figure 2.2: The network formed by the 15-node sample set  $\mathcal{N}$  with the range (a):  $M_3(\mathcal{N})$ , (b):  $r_0 = R_2(\mathcal{N}) > M_3(\mathcal{N})$  (each separation pair  $\mathcal{T}_i$  is marked with a distinct symbol), and (c):  $R_3(\mathcal{N})$ .

(See Figure 2.2 for an example with  $k = 3$ .) Despite this seemingly general formulation, the task of finding the disconnecting  $(k - 1)$ -tuples  $\mathcal{T}_i$  (also done in step 1 with  $k - 2$  to check whether  $\mathcal{G}(\mathcal{N}, M_k(\mathcal{N}))$  is  $(k - 1)$ -connected) becomes increasingly complex very rapidly as  $k$  increases. When  $k = 2$ , the disconnecting nodes, or *cutvertices*, are found in linear time with respect to the size of the graph by using depth-first-search (DFS), a basic text-book graph traversal algorithm. With  $k = 3$ , finding the *separation pairs* is already notably more difficult, but a data structure called the SPQR-tree makes it possible to do this in linear time as well.

The SPQR-tree was introduced by Di Battista and Tamassia in [DT96] and only recently correctly implemented by Gutwenger and Mutzel in [GM00]. Due to the apparent difficulty of implementing the SPQR-tree and the fact that no implementation is publicly available, a simpler – and suboptimal – algorithm for finding the separation pairs is developed in Publication [1]. In short, this algorithm is based on storing a DFS tree of the biconnected network (obtained with the initial range  $r_0$ ) with enough information to justify each node not being a cutvertex. The effects of single node removals on the DFS tree are then examined to find out whether a removal creates cutvertices in the network: if so, the removed node comprises separation pairs with the cutvertices. The emergence of cutvertices is determined by preserving those parts of the DFS tree that are known to be unaffected by the removal and rebuilding the rest of the tree. To find all the potential separation pairs,  $n - 1$  node removals have to be considered in this way. For further details, the reader is referred to Publication [1].

Note that these algorithms are also motivated by Theorem 2.1, its generalization to  $k > 1$ , and the conjectured generalizations to other spatial distributions of nodes: as  $n$  increases, the initial range  $r_0 = M_k(\mathcal{N})$  is the sought range with increasing probability.

### Empirical models

We apply the use of empirical models to our connectivity problem with uniform  $f_{\mathcal{D}}(\cdot)$  on  $\mathcal{D} = [0, 1]^2$ . The data used as the basis for these models

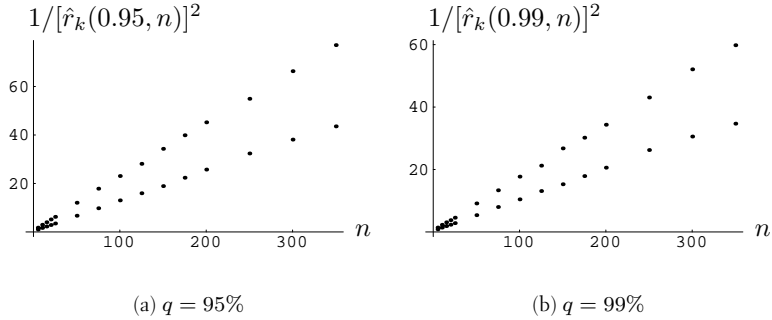


Figure 2.3: Squared inverses of estimated  $q$ -quantiles of  $R_k$ ,  $k = 1$  (upper points) and  $k = 3$  (lower points).

consists of samples of  $R_1(\mathcal{N})$ ,  $R_2(\mathcal{N})$ , and  $R_3(\mathcal{N})$  determined using the above algorithms from 5000 random realizations of  $\mathcal{N}$  with every fixed  $n$ ; the different values of  $n$  ranged from 5 to 350.

Our first models are presented in Publication [1]. There we observe, purely by visual inspection, that the squared inverse of any fixed quantile of the simulated distribution of  $R_k(\mathcal{N})$  seems to grow linearly with  $n$  (see Figure 2.3). This suggests that the  $q$ -quantile  $r_k(q, n)$  behaves like  $r_k(q, n) = 1/\sqrt{a(k, q) \cdot n + b(k, q)}$  with some parameters  $a(k, q)$  and  $b(k, q)$ . (Note that the  $q$ -quantile of the distribution of  $R_k(\mathcal{N})$  is the transmission range that provides a connected network with probability  $q$ .)

This model implies that when  $n$  increases while the transmission range equals  $r_k(q, n)$ , thus maintaining the probability  $q$  for  $k$ -connectivity, the expected degree of a node  $n\pi[r_k(q, n)]^2$  (ignoring the border effect) has the limit  $\lim_{n \rightarrow \infty} n\pi[r_k(q, n)]^2 = \pi/a(k, q)$ . On the other hand, the asymptotic Gumbel distribution of  $n\pi R_1(\mathcal{N})^2 - \log n$  implies that in this limit,  $n\pi[r_1(q, n)]^2$  equals the sum of  $\log n$  and  $-\log(-\log q)$ , the  $q$ -quantile of the Gumbel distribution, and therefore increases indefinitely. (Naturally, it then follows that  $n\pi[r_k(q, n)]^2$  must increase indefinitely for all  $k$ .) Thus, this model exhibits the same contradiction with asymptotic results as the one used in [DRL03]. We point out this contradiction already in Publication [1].

In Publication [2], we correct this deficiency. Taking the asymptotic distributions as the bases of the models, we let a model for a quantile encompass only the deviation from the asymptotic distribution, caused by the various complicated phenomena in the non-asymptotic regime.

More precisely, in the case  $k = 1$  we let the model describe the deviation of  $n\pi[r_1(q, n)]^2 - \log n$  from its asymptotic limit  $-\log(-\log q)$ . As an example, Figure 2.4(a) shows these deviations as estimated from the simulation data for  $q = 0.5$ , together with the fitted four-parameter regression model which assumes that

$$n\pi[r_1(q, n)]^2 - \log n + \log(-\log q) = a \cdot n^{-b} - c \cdot e^{-d \cdot n}, \quad a, b, c, d > 0, \quad (2.1)$$

where the power-law part is sufficient to describe the tail of the model.

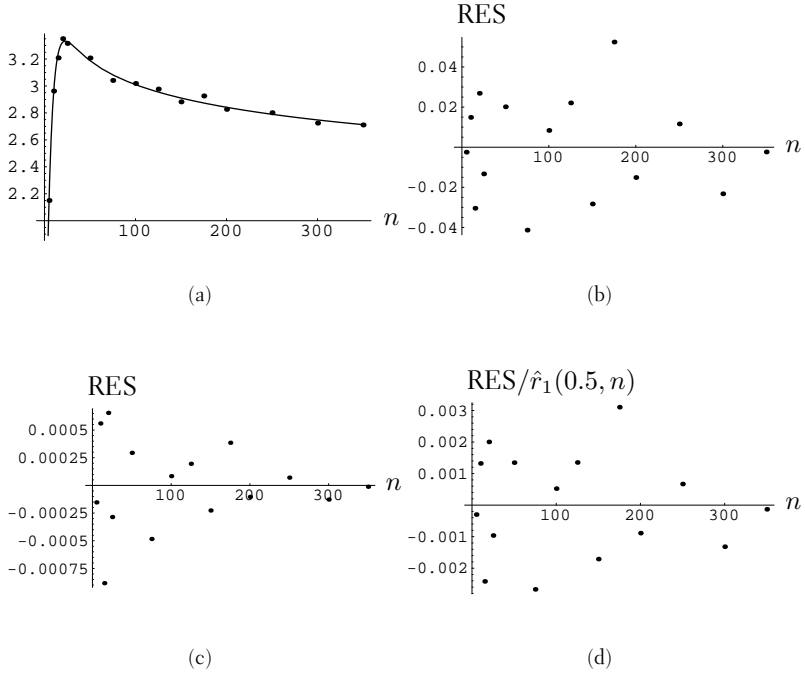


Figure 2.4: The model of the form (2.1) (a) and its residuals (b) obtained for  $n\pi[r_1(0.5, n)]^2 - \log n + \log(-\log 0.5)$ , and the overall absolute (c) and relative (d) residuals for  $r_1(0.5, n)$ .

The residuals, i.e. the differences between the data points and the fitted model, plotted in Figure 2.4(b), seem to be evenly scattered around zero level, showing no trend as an indication of the convergence parting with the model. Furthermore, their variance around the zero level seems to remain constant as  $n$  increases, implying that all the data points yield equally accurate information for fitting the model. Note that this is not the case with the residuals thus obtained for  $r_1(0.5, n)$  itself, plotted in Figure 2.4(c), which demonstrate how the variance of  $R_1(\mathcal{N})$  and hence that of the quantile estimate  $\hat{r}_1(q, n)$  decreases with  $n$ . Finally, Figure 2.4(d) shows the relative residuals for  $r_1(0.5, n)$ : the near-identical pattern with Figure 2.4(b) implies that the fitting of the model can be considered almost equivalent to minimizing the sum of squared relative residuals of  $\hat{r}_1(q, n)$ .

The median was chosen as the first quantile for building this model to maximize the accuracy of the used quantile estimates, as these estimates are obtained from simulation data with limited sample sizes. However, the same model is well able to describe the convergence of any quantile: Figure 2.5 shows the equivalent of Figure 2.4(d) for two more extreme quantiles. One can observe that due to the increasing inaccuracy inherent in estimating extreme quantiles from limited-sized data, the model error is roughly within 1% for  $q = 0.95$  and within 1.5% for  $q = 0.99$ , as opposed to only 0.3% for  $q = 0.5$ .

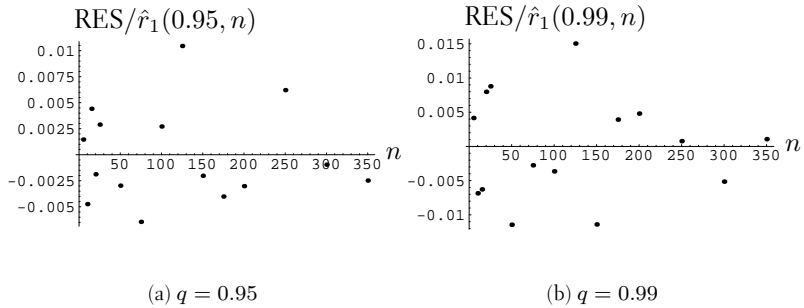


Figure 2.5: Relative residuals for other quantiles of  $R_1$ , obtained by using the model (2.1).

We now demonstrate the qualitative meaning of the asymptotic Gumbel distribution of  $n\pi M_1(\mathcal{N})^2 - \log n$  with uniform  $f_{\mathcal{D}}(\cdot)$  on  $\mathcal{D} = [0, 1]^2$ . Let us first write the probability that a random node of  $\mathcal{N}$  is isolated, i.e., is out of range  $r$  from all other nodes. If we neglect the possibility that this node is within range from the border, the probability that a given other node is out of range is  $1 - \pi r^2$ . Because the node locations are independent, the probability that the node is isolated then equals  $(1 - \pi r^2)^{n-1}$ .

We are interested in what happens when  $n$  tends to infinity; naturally, if the range  $r$  is fixed, the above probability goes to zero. In the non-trivial case where  $r$  decreases as  $n$  increases, the probability satisfies

$$\begin{aligned} \Pr[\text{Random node isolated}] &= \frac{(1 - \pi r^2)^n}{1 - \pi r^2} \xrightarrow[n \rightarrow \infty]{} (1 - \pi r^2)^n \\ &= \left(1 - \frac{n\pi r^2}{n}\right)^n \xrightarrow[n \rightarrow \infty]{} \exp(-n\pi r^2), \end{aligned}$$

where the last limit is equal to the probability that there are no points of a Poisson process with intensity  $n$  in a circle with radius  $r$ . Note that the complement probability, as a function of  $r$ , is the cumulative distribution function of a randomly selected node's nearest-neighbor distance. On the other hand, the probability that *none* of the nodes is isolated is the cumulative distribution function of  $M_1(\mathcal{N})$ .

Now, let us write the latter cumulative distribution as that of the *maximum* of  $n$  independent and identically distributed nearest-neighbor distances, and take the logarithm of both sides:

$$\begin{aligned} \Pr[M_1(\mathcal{N}) \leq r] &= [1 - \exp(-n\pi r^2)]^n \quad (2.2) \\ \Leftrightarrow \log \Pr[M_1(\mathcal{N}) \leq r] &= n \log [1 - \exp(-n\pi r^2)]. \end{aligned}$$

Again, in the non-trivial case where the probability above does not tend to zero with increasing  $n$ , the probability that one random node is

isolated,  $\exp(-n\pi r^2)$ , diminishes with increasing  $n$ . Making use of the limit  $\log(1+x) \xrightarrow{|x| \rightarrow 0} x$ , we then have

$$\begin{aligned} \log \Pr[M_1(\mathcal{N}) \leq r] &\xrightarrow{n \rightarrow \infty} -n \exp(-n\pi r^2) \\ \Leftrightarrow -\log(-\log \Pr[M_1(\mathcal{N}) \leq r]) &\xrightarrow{n \rightarrow \infty} n\pi r^2 - \log n \stackrel{\text{def}}{=} \alpha. \end{aligned}$$

Expressing the event  $M_1(\mathcal{N}) \leq r$  equivalently using the definition of  $\alpha$ , we arrive at

$$\Pr[n\pi M_1(\mathcal{N})^2 - \log n \leq \alpha] \xrightarrow{n \rightarrow \infty} \exp(-e^{-\alpha}),$$

i.e., the asymptotic Gumbel distribution of  $n\pi M_1(\mathcal{N})^2 - \log n$ .

The above derivation shows that the border effect does not dominate in the asymptotic distribution of  $M_1(\mathcal{N})$  (and hence that of  $R_1(\mathcal{N})$ ). In Publication [2], which was submitted before becoming aware of the exact asymptotic distributions recently derived in [WY04], we derive approximate asymptotic distributions for  $R_k(\mathcal{N})$ ,  $k > 1$ , as the bases of the corresponding quantile models. We do this as above, by neglecting border effects and writing the distribution of  $M_k(\mathcal{N})$  as that of the maximum of  $n$  i.i.d.  $k$ -nearest-neighbor distances; we refer the reader to the publication for the details.

Finally, as the most important argument in favor of using these models, in Publication [2] we demonstrate their ability to predict the independent simulation data presented in [Bet02]. Figure 2.6, excerpted from [Bet02], shows the simulation results: here,  $r$  has been fixed while  $n$  has been varied. (The analytical curve represents the asymptotic relation (2.2).) The

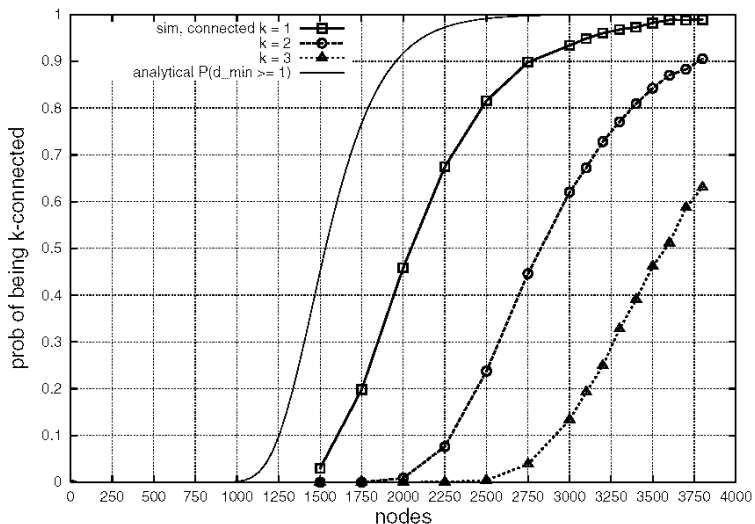


Figure 2.6: “Simulation results for  $n$  nodes with  $r_0 = 20\text{m}$  uniformly distributed on  $A = 500 \times 500\text{m}^2$  {...}, 3000 random topologies” [Bet02].



Table 2.1: The number of nodes  $n$  required to achieve  $k$ -connectivity with probability  $q$  when  $r/\sqrt{A}$  is as in Figure 2.6, as predicted by our quantile models for  $R_1(\mathcal{N})$ ,  $R_2(\mathcal{N})$ , and  $R_3(\mathcal{N})$

	$q$				
	50%	75%	90%	95%	99%
$k = 1$	2057	2387	2790	3144	3871
$k = 2$	2805	3262	3807	4208	
$k = 3$	3533	4065			

predictions of our quantile models to this example scenario are given in Table 2.1. Comparison of the two shows that although the models were fitted to simulation data involving no more than  $n = 350$  nodes, their predictions turn out to be quite accurate even with up to ten times as many nodes. Furthermore, the models for  $R_2(\mathcal{N})$  and  $R_3(\mathcal{N})$  seem to perform as well as those for  $R_1(\mathcal{N})$ , implying that the derived asymptotic distributions serve as reasonable approximations when  $k > 1$ .

### Connectivity probability under the Random Waypoint mobility model

Recall from the literary review of this section that Theorem 2.1 generalizes to  $R_k(\mathcal{N})$  and  $M_k(\mathcal{N})$  when  $k > 1$ , and that the asymptotic distribution of  $M_1(\mathcal{N})$  as that of the maximum of  $n$  independent nearest-neighbor distances also holds for normally distributed points, which gives reason to assume that these properties should hold for more general spatial distributions.

The Random Waypoint (RWP) mobility model has been treated as a special case of a more general class of models using Palm calculus in [LV05], where it is rigorously proven that the RWP model reaches a stationary state distribution if and only if the inverse of the random velocity drawn for each leg and the pause time drawn after each leg have finite expectations, and that this stationary distribution is unique.

Assuming such a velocity distribution, consider the steady-state distribution  $f_{\mathcal{D}}(\cdot)$  for the location of a node moving according to the RWP model with uniform waypoint distribution over the convex domain of movement  $\mathcal{D}$  and no pause times (apart from the condition for reaching stationarity, this distribution is independent of the velocity distribution [Le 05]). This distribution is not uniform: the probability mass is concentrated around the center of  $\mathcal{D}$  whereas the probability density reaches zero at the boundary of  $\mathcal{D}$ . While this stationary distribution is expressed in a complicated form for rectangular  $\mathcal{D}$  in [NC04], approximated for various shapes of  $\mathcal{D}$  in [BRS03], and given a formal yet high-level representation in [LV05], an explicit expression for any convex  $\mathcal{D}$  has been derived recently [HV05, HLV06]. With the theoretical motivation that started this subsection, in Publication [3] we utilize this exact distribution in estimating the probability that a network of  $n$  nodes moving according to the RWP model is  $k$ -connected at a random time instant; in other words, we address the problem statement of this sec-

tion when  $f_{\mathcal{D}}(\cdot)$  is the stationary node location distribution in the RWP model.

The approximation method is the same as in the work by Bettstetter, i.e., we approximate the probability of  $k$ -connectivity by the probability that a random node has at least  $k$  other nodes within range, raised to the power  $n$ . This is equivalent to approximating the distribution of  $R_k(\mathcal{N})$  by that of the maximum of  $n$  i.i.d.  $k$ -nearest-neighbor distances. However, whereas Bettstetter uses an approximation in deriving the latter probability from the spatial distribution, we derive it exactly. The other difference is that we use the exact node location distribution, unlike the approximation from [BRS03] used by Bettstetter in [Bet04].

Figure 2.7, demonstrating the accuracy of our approximation, shows that the approximation is quite poor with small  $n$  but indeed improves as  $n$  increases, which supports our assumption that the approximation is in fact asymptotically accurate. Furthermore, the approximation turns out to

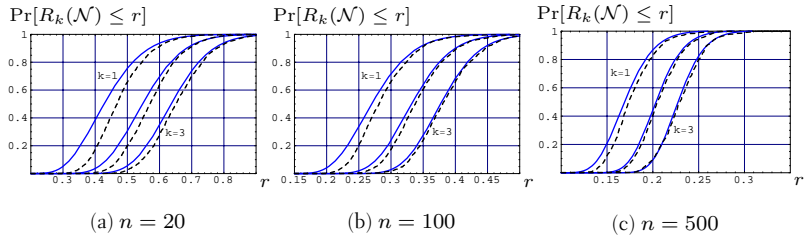


Figure 2.7: Probability that  $n$  nodes with range  $r$  moving in a unit disk make a  $k$ -connected network at a random point in time, as determined using the approximation in Publication [3] (solid lines) and by simulation (dashed lines).

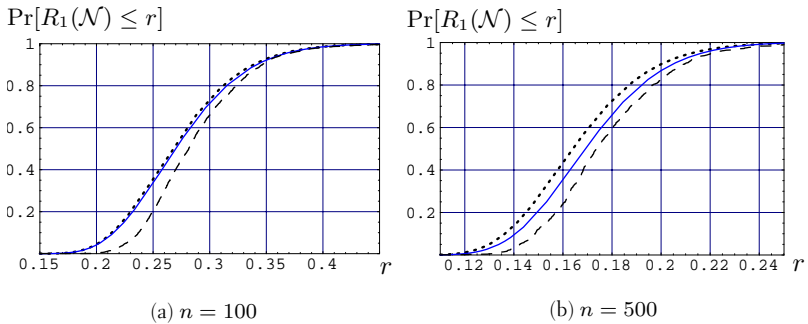


Figure 2.8: Probability that a network of  $n$  nodes with range  $r$  moving in a unit disk is connected, as determined using the approximations presented in [Bet04] (dotted lines) and in Publication [3] (solid lines), and by simulation (dashed lines).

improve with increasing  $k$ . The improvement brought by our approximation to that used in [Bet04] is depicted in Figure 2.8.

The general conclusion drawn in [Bet04] was that with any given  $n$  and  $r$ , the connectivity probability under the spatial node distribution caused by RWP mobility is always lower than when the nodes are uniformly distributed. One additional finding in Publication [3] is that with small  $n$ , the case is in fact the opposite. This can be observed from simulation data but is also correctly predicted by our approximation. The reason why this was not discovered in [Bet04] was that only values of  $n$  greater than 200 were observed: the reverse situation can be observed roughly when  $n < 100$ . The intuitive explanation for this phenomenon is that when only few points are drawn from the centralized RWP spatial distribution, they are all likely to lie in the center of the domain, close to each other in comparison to the uniform case. As the number of points increases, it becomes more likely that there are individual outlying points located far – relative to the uniform case – from the rest of the nodes.

## 2.2 Coverage of random networks

In this section, we study the coverage of random sensor networks under a Boolean coverage-disk model. The random locations of the sensors are motivated by the vision of large numbers of small sensors, often referred to as “smart dust”, being scattered over some terrain from, say, an aircraft. On the other hand, networks of mobile sensors have also been studied; at a random time instant, the locations of such sensors are random.

Existing work on the coverage of mobile sensors under this coverage model has addressed, e.g., the time until a point is covered by sensors in Brownian motion [KKP03] and various coverage dynamics of sensors under a random-direction mobility model [LBD<sup>+</sup>05]. Taking into account the limited operational lifetime of sensors, the temporal aspect of coverage also with stationary sensors has been studied in [ZH04], where the authors derive an upper bound for the  $\alpha$ -lifetime of large random networks, i.e., the maximum time for which at least the fraction  $\alpha$  of some target domain is covered.

### Problem statement

We preserve much of the notation and assumptions used in the context of our connectivity problem in the previous section. We now assume throughout that the locations  $\mathcal{N}$  of the  $n$  sensors are independent and *uniformly* distributed in some domain  $\mathcal{S} \supseteq \mathcal{D}$  where  $\mathcal{D}$  is the bounded domain to be covered, and that all the sensors have a common sensing range  $r$ : see Figure 2.9 for an illustration. We take  $\mathcal{S}$  and  $\mathcal{D}$  to be subsets of the Euclidean plane  $\mathbb{R}^2$ , although the generalization to a higher number of dimensions is straightforward.

For given  $\mathcal{N}$ ,  $\mathcal{D}$  and  $r$ , let  $C(\mathcal{N}, \mathcal{D}, r)$  denote the *area coverage* as defined in [LT03], i.e., the fraction of the area of  $\mathcal{D}$  that is covered by at least one sensor. As a function of the random sensor locations  $\mathcal{N}$ ,  $C(\mathcal{N}, \mathcal{D}, r)$  is a random variable. Our problem is the following.

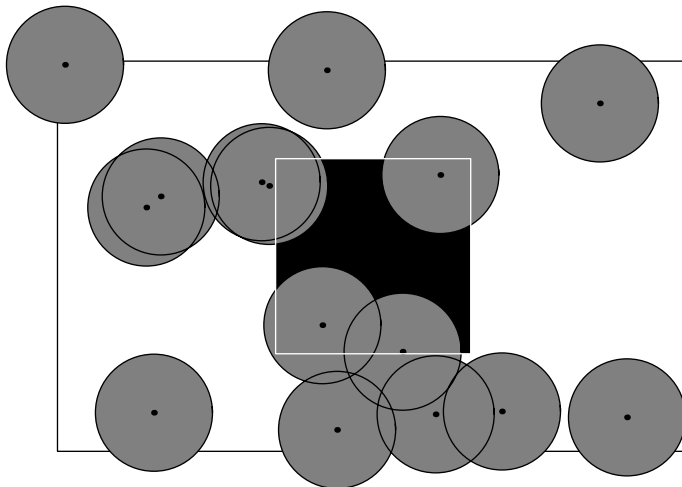


Figure 2.9: Illustration of the coverage problem definitions. The locations  $\mathcal{N}$  (black points) of  $n = 15$  sensors have been drawn uniformly at random in the rectangular region  $\mathcal{S}$ . Each sensor has the same sensing range  $r$  and thus covers a disk drawn in gray. With this particular realization, the fraction  $C(\mathcal{N}, \mathcal{D}, r)$  of the area of the square-shaped target domain  $\mathcal{D}$  covered by the sensors is around 50%.

*Given  $n, \mathcal{S}, \mathcal{D}$  and  $r$ , what is the probability distribution of  $C(\mathcal{N}, \mathcal{D}, r)$ ?*

Note that in the unbounded limit where  $\mathcal{S}$  is infinitely enlarged to be  $\mathbb{R}^2$  while keeping the average sensor density  $n/|\mathcal{S}|$  fixed, the point process marking the sensor locations becomes a Poisson process with intensity  $\lambda = n/|\mathcal{S}|$ .

### Review of existing results

To our knowledge, this problem has not been addressed as such, in the general form stated above. Furthermore, many existing studies on different subproblems assume instead that the sensors are located on the points of a Poisson point process on some bounded set to be covered. Overall, as with the problem of connectivity of random networks, exact analytical results in finite settings do not exist.

When  $\mathcal{D} = \mathcal{S}$  in the above unbounded limit, the area coverage has a deterministic value: by the properties of the Poisson process, this value is  $1 - e^{-\lambda\pi r^2}$ , as pointed out in [LT03].

In the case of sensors on the points of a Poisson process on a unit-area disk  $\mathcal{S} = \mathcal{D}$ , with some intensity  $\lambda$ , it has been shown in [Hal88] that the probability that  $\mathcal{D}$  is completely covered has the bounds

$$\begin{aligned} \max\{0, 1 - 3(1 + \pi r^2 \lambda^2)e^{-\lambda\pi r^2}\} &< \Pr[C(\mathcal{N}, \mathcal{D}, r) = 1] \\ &< \frac{1}{20} \max\{0, 1 - (1 + \pi r^2 \lambda^2)e^{-\lambda\pi r^2}\}. \end{aligned}$$

As noted in [GK98], if the number of points  $n$  is fixed to  $n = \lambda$  and  $r(n)$  is such that  $n\pi r(n)^2 - (\log n + \log \log n) \xrightarrow[n \rightarrow \infty]{} \infty$ , then this result implies that  $\Pr[C(\mathcal{N}, \mathcal{D}, r(n)) = 1] \rightarrow 1$ , and if  $n\pi r(n)^2 - (\log n + \log \log n) \rightarrow -\infty$ , then  $\lim_{n \rightarrow \infty} \Pr[C(\mathcal{N}, \mathcal{D}, r(n)) = 1] \leq 19/20$ .

The following asymptotic result by Janson, also about complete coverage, is actually a special case, adapted to the context of our problem, of a much more general result presented in [Jan86]. Let  $|\mathcal{S}|$  denote the Lebesgue measure of  $\mathcal{S}$ .

**Theorem 2.2** *Suppose that  $\mathcal{D} = [0, 1]^2$ , that  $\text{Closure}(\mathcal{D}) \subset \text{Interior}(\mathcal{S})$  and that  $|\mathcal{S}| < \infty$ . Suppose further that  $A$  is the disk with unit radius.*

*For  $r > 0$ , consider the set of disks  $rA + X$ , where  $X$  is a set of random points uniformly distributed on  $\mathcal{S}$ , and let  $N_r$  be the number of disks required to cover  $\mathcal{D}$  completely.*

*Let  $U$  have the extreme value distribution  $\Pr(U \leq u) = \exp(-e^{-u})$ . Then, as  $r \rightarrow 0$ ,*

$$\frac{\pi r^2}{|\mathcal{S}|} N_r + \log \pi r^2 - 2 \log(-\log \pi r^2) \xrightarrow{d} U.$$

The original theorem is stated for covering a more general  $\mathcal{D}$  any given number of times with random sets  $A$ , in any number of dimensions, but in this general case the above asymptotic expression is more complicated. Note that just as with the threshold range for connectivity, we have here again a quantity with an asymptotic Gumbel distribution.

Complementing the above results, sufficient conditions for asymptotically almost surely covering  $\mathcal{D} = [0, 1]^2$  any given number of times, and also for not covering  $\mathcal{D}$ , have been derived in [KLB04], in the cases of  $n$  sensors located on the points of a regular grid or points drawn uniformly at random, and sensors at the points of a Poisson point process.

Finally, as a curiosity, we remark that the sufficient conditions regarding asymptotic probability of complete coverage derived in [ZH04] can be obtained as direct corollaries from [Jan86].

In Publication [4], we address the expected value of  $C(\mathcal{N}, \mathcal{D}, r)$  over  $\mathcal{N}$ , as well as the probability of complete coverage in the non-asymptotic regime. We devote the next two sections to discussing these quantities.

### Expected value of $C(\mathcal{N}, \mathcal{D}, r)$ over $\mathcal{N}$

In line with our problem statement, assume that  $n$ ,  $\mathcal{S}$ ,  $\mathcal{D}$  and  $r$  are given. Then the conditional expected value of the area coverage over the different sensor configurations,  $\mathbb{E}_{\mathcal{N}}[C(\mathcal{N}, \mathcal{D}, r) | n, \mathcal{S}, \mathcal{D}, r]$ , is simply the integral of  $C(\mathcal{N}, \mathcal{D}, r)$  over the joint probability distribution of  $\mathcal{N}$ , in this case, a uniform distribution over  $\mathcal{S}^n$ . On the other hand, when also  $\mathcal{N}$  is fixed,  $C(\mathcal{N}, \mathcal{D}, r)$  is equal to the conditional probability that a random point in  $\mathcal{D}$  is covered by a sensor. Thus, it is easy to see that  $\mathbb{E}_{\mathcal{N}}[C(\mathcal{N}, \mathcal{D}, r) | n, \mathcal{S}, \mathcal{D}, r] = \Pr_{\mathcal{N}}[\text{Random point in } \mathcal{D} \text{ is covered} | n, \mathcal{S}, \mathcal{D}, r]$ .

In Publication [4], we examine the expected area coverage when  $\mathcal{D}$  is a disk with radius  $R$  by solving this probability. We consider two alternative cases. When  $\mathcal{S} = \mathcal{D}$ , we must account for the border effect, i.e., the fact

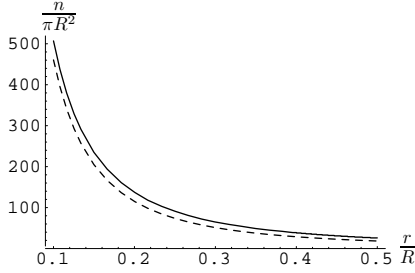


Figure 2.10: The number of sensors per area  $\pi R^2$  required to ensure  $E_{\mathcal{N}}[C(\mathcal{N}, \mathcal{D}, r) | n, \mathcal{S}, \mathcal{D}, r] = 0.99$  if  $\mathcal{D}$  is a disk with radius  $R$ , when  $\mathcal{S} = \mathcal{D}$  causing a border effect (solid line) and when  $\mathcal{S} \rightarrow \mathbb{R}^2$ , eliminating the border effect (dashed line).

that points near the boundary are less likely to be covered. In this case, we obtain by straightforward calculation

$$E_{\mathcal{N}}[C(\mathcal{N}, \mathcal{D}, r) | n, \mathcal{S}, \mathcal{D}, r] = 1 - \left[ \frac{\pi(R-r)^2}{\pi R^2} \left( 1 - \frac{\min\{\pi r^2, \pi R^2\}}{\pi R^2} \right)^n + \int_{|R-r|}^R \frac{2\rho}{R^2} \left( 1 - \frac{A(\rho, R, r)}{\pi R^2} \right)^n d\rho \right],$$

where we allow also the case  $R > r$ . Here,  $A(\rho, R, r)$  denotes the area of the intersection of two disks with radii  $r$  and  $R$  when their centers are separated by  $\rho > |R - r|$  (else this area equals  $\min\{\pi r^2, \pi R^2\}$ ).

In the second case, we consider the limit  $\mathcal{S} \rightarrow \mathbb{R}^2$ , which eliminates the border effect; by the properties of the resulting Poisson process, we know that in this case the expected area coverage equals  $1 - e^{-\lambda\pi r^2}$ .

Figure 2.10 demonstrates the border effect by comparing the sensor densities required to ensure a 99% expected area coverage on this particular  $\mathcal{D}$  in the two cases, as implied by these results.

### Probability of complete coverage as a learning problem

We now turn to the subproblem of determining the probability that  $\mathcal{D}$  is completely covered, i.e.  $\Pr[C(\mathcal{N}, \mathcal{D}, r) = 1]$ . In Publication [4], we point out that this problem, like the connectivity problem discussed earlier, also reduces to knowing the distribution of a well-defined threshold range: given  $\mathcal{D}$ , let  $R_c(\mathcal{N})$  denote the smallest sensing range  $r$  with which  $\mathcal{D}$  is completely covered by sensors at given locations  $\mathcal{N}$ , i.e., with which  $C(\mathcal{N}, \mathcal{D}, r) = 1$ . We refer to  $R_c(\mathcal{N})$  as the *threshold range for complete coverage*. Now, the event  $\{C(\mathcal{N}, \mathcal{D}, r) = 1\}$  is equivalent to  $\{r \geq R_c(\mathcal{N})\}$ , whence this problem reduces to knowing the probability distribution of  $R_c(\mathcal{N})$  with given  $n$ ,  $\mathcal{S}$  and  $\mathcal{D}$ .

For given  $\mathcal{N}$ ,  $R_c(\mathcal{N})$  is equal to the distance from a point in  $\mathcal{D}$  to the nearest point in  $\mathcal{N}$ , maximized over all points in  $\mathcal{D}$ . It follows that  $R_c(\mathcal{N})$  can be easily determined using the sensors' Voronoi diagram: it can be deduced to be the longest distance from a sensor to an edge of its Voronoi cell

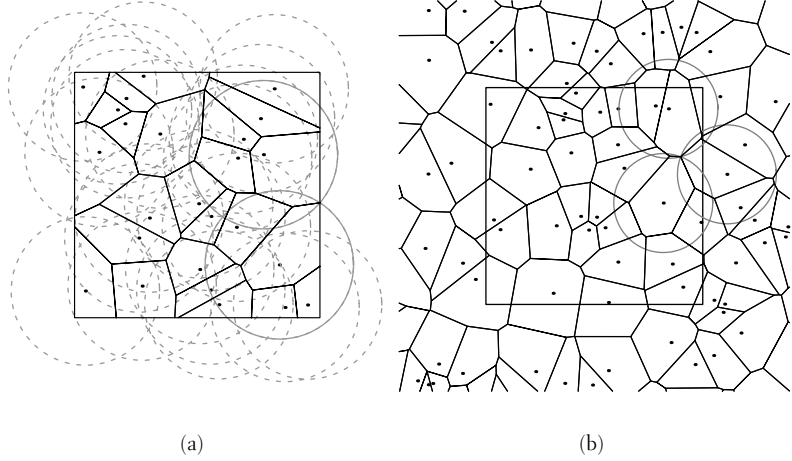


Figure 2.11: Example of the threshold range for complete coverage when a) 25 sensors are placed randomly inside the square-shaped domain  $\mathcal{S} = \mathcal{D}$ , and b) sensors are scattered uniformly over the whole plane with average density  $25/|\mathcal{D}|$ . The critical coverage ranges are shown with solid circles.

in  $\mathcal{D}$ . Moreover, when the boundary of  $\mathcal{D}$  is piecewise linear, it is sufficient to concentrate on cell corners. Like the expected area coverage, the probability of complete coverage is strongly influenced by the border effect: Figure 2.11 shows examples of the threshold range determined for a square-shaped  $\mathcal{D}$ , both when  $\mathcal{S} = \mathcal{D}$  and when the border effect is eliminated.

Recall that Theorem 2.2 gives the asymptotic distribution of the number of sensors required on the set  $\mathcal{S}$  to cover  $\mathcal{D}$  completely, as their sensing range tends to zero. It also allows us to solve the inverse cumulative distribution function of this number: to have  $\mathcal{D}$  completely covered with probability at least  $q$ , the number of sensors with range  $r$  must be at least  $n(r) = \lceil -\log(-\log q) - \log \pi r^2 + 2 \log(-\log \pi r^2) \rceil \cdot |\mathcal{S}|/\pi r^2$ .

On the other hand, we may also want to know the inverse cumulative distribution function of the threshold range for complete coverage, i.e., ask how large the sensing range should be with a given number of sensors  $n$  to guarantee a coverage probability of at least  $q$ : of course, the range should then not be less than the smallest  $r$  with which the given  $n$  is at least  $n(r)$ . By the continuity and monotonicity of  $n(r)$ , it follows that  $r$  should be at least  $r^{-1}(n)$ , the inverse function of  $n(r)$ . We may thus conclude that Theorem 2.2 also holds with the given number of sensors  $n$  and the threshold range for complete coverage  $R_c(\mathcal{N})$  substituted in the place of  $N_r$  and  $r$ , respectively. Because the resulting expression with the asymptotic Gumbel distribution is a monotonically increasing function of  $R_c(\mathcal{N})$  with given  $n$ , it follows that this result gives the asymptotic distribution of the threshold range for complete coverage when the number of sensors  $n$  tends to infinity.

The above observations provide the ingredients for treating the probability of complete coverage as a learning problem, exactly as done and discussed earlier with the connectivity probability. In what follows, we model

the convergence of the distribution of the threshold range for complete coverage using empirical models when  $\mathcal{D} = [0, 1]^2$ .

The purpose of requiring that the set  $\mathcal{S}$  completely encompass  $\mathcal{D}$  in Theorem 2.2 is to eliminate the border effect, i.e., to avoid complications resulting from the otherwise likely event that the boundary of  $\mathcal{D}$  is the last part to be covered. In other words, Theorem 2.2 does not apply to the case  $\mathcal{S} = \mathcal{D}$  demonstrated in Figure 2.11(a). We will next discuss the empirical models in the setting without the border effect but will return to the case  $\mathcal{S} = \mathcal{D}$  shortly.

To eliminate the border effect, the set  $\mathcal{S}$  where the  $n$  sensors are randomly placed for each sample of the threshold range  $R_c(\mathcal{N})$  is chosen as a larger square centered around the unit square  $\mathcal{D}$ . This other square is made large enough so that the resulting threshold range is never greater than the shortest distance from the boundary of  $\mathcal{D}$  to that of  $\mathcal{S}$ ; this way, sensors outside  $\mathcal{S}$  would not affect the threshold range. (This requirement is checked for each sample of  $R_c(\mathcal{N})$ ; if it is not met, the set  $\mathcal{S}$  is enlarged further and all the simulations for the particular sensor density are repeated.) As when simulating the threshold range for connectivity,  $R_c(\mathcal{N})$  was determined from 5000 random realizations of  $\mathcal{N}$  with every fixed sensor density  $n/|\mathcal{S}|$ , and the value of  $n/|\mathcal{S}|$  was varied up to 350.

Now, let  $r_c(q, n)$  denote the  $q$ -quantile of  $R_c(\mathcal{N})$ . In the same spirit as with the models in Publication [2], we let the model describe the deviation of  $\frac{n}{|\mathcal{S}|} \pi r_c(q, n)^2 + \log \pi r_c(q, n)^2 - 2 \log(-\log \pi r_c(q, n)^2)$  from its asymptotic limit  $-\log(-\log q)$ . Figure 2.12 is the equivalent of Figure 2.4, now showing as an example the corresponding information for these deviations for  $q = 0.95$  along with a fitted four-parameter regression model of the form

$$\begin{aligned} \frac{n}{|\mathcal{S}|} \pi r_c(q, n)^2 + \log \pi r_c(q, n)^2 - 2 \log(-\log \pi r_c(q, n)^2) + \log(-\log q) \\ = a \cdot n^{-b} + c \cdot n^{-d}, \quad a, b, c, d > 0, \end{aligned} \quad (2.3)$$

i.e., a combined power-law where a slower-decaying component characterizes the tail of the model. In short, this figure allows for the same conclusions as Figure 2.4.

We now turn to the case  $\mathcal{S} = \mathcal{D}$ . In [Jan86], Janson is able to derive the equivalent of Theorem 2.2 when  $\mathcal{S} = \mathcal{D} = [0, 1]^2$  and the set  $A$  is the square  $[-1/2, 1/2]^2$  instead of the unit disk. In this case, covering  $\mathcal{D}$  can be decomposed into the asymptotically seemingly independent events of covering the interior  $[r/2, 1 - r/2]^2$ , i.e., the region where the effect of the randomly placed scaled squares  $rA$  is not affected by the boundary of  $\mathcal{D}$ , and covering its normal projections to the boundary such as  $\{0\} \times [r/2, 1 - r/2]$  with intersections of the overlapping squares, with length  $r$ . This is because when  $A$  is chosen as the square, covering the boundary of  $\mathcal{D}$  implies covering the strip between the boundary and the interior. Thus, the problem with border effects can be decomposed into subproblems in one and two dimensions, both free of border effects.

When  $A$  is chosen as the disk with unit radius as in Theorem 2.2 adapted to our problem, covering the boundary of  $\mathcal{D}$  no longer implies covering the strip between the boundary and the interior, and this decomposition is no



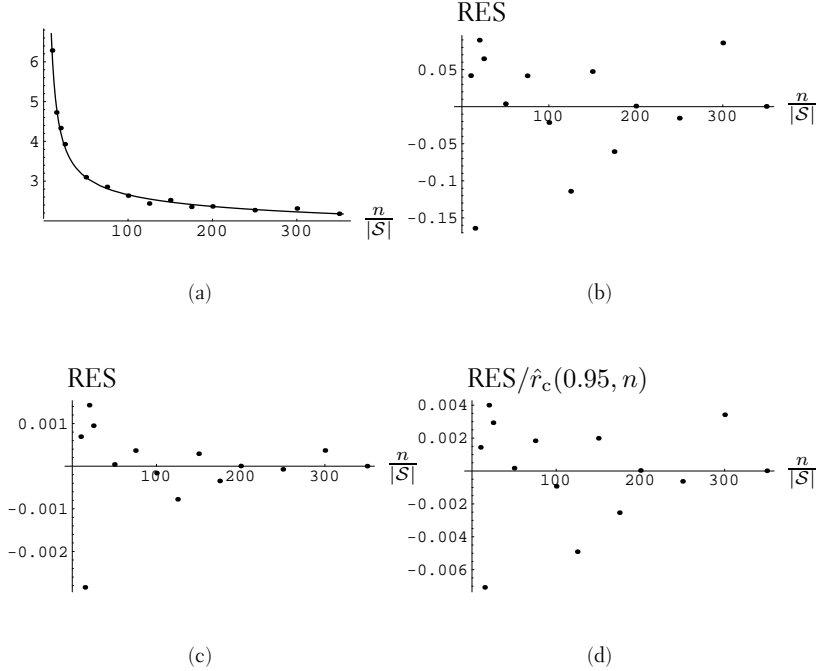


Figure 2.12: The model of the form (2.3) (a) and its residuals (b) obtained for  $\frac{n}{|\mathcal{S}|} \pi r_c(0.95, n)^2 + \log \pi r_c(0.95, n)^2 - 2 \log(-\log \pi r_c(0.95, n)^2) + \log(-\log 0.95)$ , and the overall absolute (c) and relative (d) residuals for  $r_c(0.95, n)$ .

longer equivalent to the original problem. However, we may still derive the asymptotic probability that the interior  $[r, 1-r]^2$  and its normal projections to the boundary of  $\mathcal{D}$  are covered by the scaled disks  $rA$ , by repeating the steps shown in [Jan86] when  $A$  is  $[-1/2, 1/2]^2$ ; the reader is referred to the appendix of Publication [4] for the details. Given these coverings, we assume that the remaining parts are covered asymptotically with high probability and thus use the result as an *approximation for the asymptotic distribution for the threshold range for complete coverage of  $\mathcal{S} = \mathcal{D} = [0, 1]^2$* :

$$\begin{aligned} \lim_{n \rightarrow \infty} \Pr [n\pi R_c(\mathcal{N})^2 + \log \pi R_c(\mathcal{N})^2 - 2 \log(-\log \pi R_c(\mathcal{N})^2) \leq u] \\ = \exp \left( -\frac{4}{\sqrt{\pi}} e^{-u/2} - e^{-u} \right). \end{aligned} \quad (2.4)$$

Note that here  $|\mathcal{S}| = 1$ , so the quantity with the converging distribution is in effect the same as in the case without the border effect; only the limiting distribution is now different.

We use again the model (2.3) to describe the convergence as in the previous case, with the obvious difference that the limiting  $q$ -quantile is now  $-2 \log(\sqrt{4/\pi - \log q} - 2/\sqrt{\pi})$  instead of that of the Gumbel distribution,  $-\log(-\log q)$ . We show here as examples only the relative residuals ob-

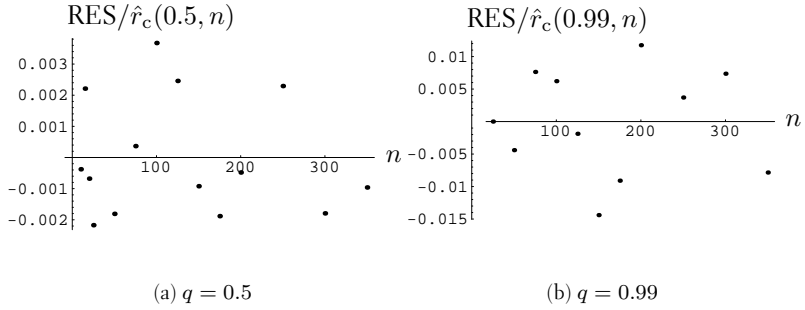


Figure 2.13: Relative residuals for two quantiles of  $R_c(\mathcal{N})$ , obtained by applying the model (2.3) to the approximate asymptotic distribution (2.4) when  $\mathcal{S} = \mathcal{D} = [0, 1]^2$ .

tained after fitting this model to 50% and 99% quantiles estimated from simulation data, again with 5000 samples generated for each  $n$  and  $n$  ranging up to 350: see Figure 2.13.

We opt to validate these models in Publication [4] only in the latter case with  $\mathcal{S} = \mathcal{D} = [0, 1]^2$ , i.e. when the models are based on the approximate asymptotic distribution (2.4), since it is fair to assume that the models based on the exact distribution in the case without border effects should perform at least as well. To this end, we test how well these models are able to predict the distribution of additional samples of  $R_c(\mathcal{N})$  determined from simulated random realizations with  $n = 1000$  sensors; because of the resulting extensive computation times, only 1000 samples were generated. The result is shown in Figure 2.14: slight differences are noticeable, which may be due to the approximation (2.4), invalidity of the model (2.3), inaccuracy in the model parameter estimates, or the small number of samples

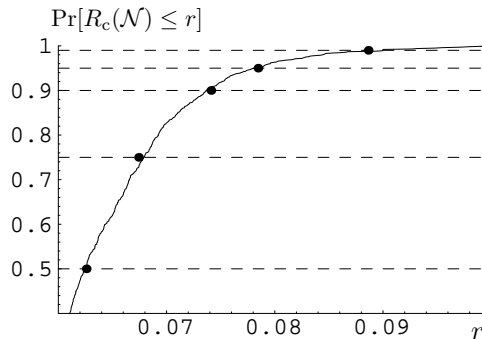


Figure 2.14: The empirical cumulative distribution of  $R_c(\mathcal{N})$  determined from 1000 random realizations with  $n = 1000$  sensors in  $\mathcal{S} = \mathcal{D} = [0, 1]^2$  (solid line), and the predictions (drawn as points) of empirical models fitted to five different quantiles (indicated with the dashed lines).

in the validation data. However, the fact that the prediction errors are in both directions makes a systematic error seem less likely.

### 2.3 Connectivity improvement as an optimization problem

In Publication [5], we view the connectivity problem from a new perspective: we are concerned with what can be done when a wireless ad hoc network needs to be formed but the network nodes are too far apart to form a network with a desired level of connectivity. More precisely, we study the option of improving the connectivity of a static ad hoc network by placing additional network nodes at optimized locations. Networks where adding extraneous nodes is feasible are some sensor networks and such ad hoc networks that are used in a controlled situation where some central entity can organize the deployment of the nodes.

To our knowledge, the connectivity problem in ad hoc networks has not been addressed so far from this practical viewpoint: probably the closest related work has been done in [DTH02], where the authors study how the existence of base stations attached to fixed, wired network infrastructure improves the connectivity of random networks.

#### Problem statements

Reusing the notation introduced in Section 2.1, we assume that the existing network is disconnected and consists of  $n$  nodes with given locations  $\mathcal{N}$  in a bounded and convex deployment region  $\mathcal{D} \subset \mathbb{R}^d$ ,  $d > 1$ , and that we may place additional nodes in  $\mathcal{D}$  to be used as relays in the network. To distinguish between the two, we will refer to the existing nodes as *terminal nodes* and the additional nodes as *relay nodes*. We assume that the terminal nodes and the relay nodes have equal transmission and reception capabilities and hence have a common transmission range  $r$ . In this setting, the basic form of our problem is as follows:

*Given  $\mathcal{N}$  and  $r$ , find a set of locations  $\mathcal{N}_r$  with minimum cardinality  $n_r$  that makes the geometric graph  $\mathcal{G}(\mathcal{N} \cup \mathcal{N}_r, r)$  connected.*

Obviously,  $\mathcal{N}_r$  is the set of locations to place relay nodes.

We also study the following problem variant which comes into question when too few relay nodes are available to make the network connected. In this case, we need to define the *utility* of the network.

*Given  $\mathcal{N}$  and  $r$ , find  $\mathcal{N}_r$  that maximizes the chosen utility metric  $U = U(\mathcal{G}(\mathcal{N} \cup \mathcal{N}_r, r))$ , subject to  $n_r \leq n_r^{\max} \in \mathbb{Z}_+$ .*

The choice of the utility metric depends on the target application. Throughout Publication [5], we use the greatest number of terminal nodes that are all in the same connected component of the graph  $\mathcal{G}(\mathcal{N} \cup \mathcal{N}_r, r)$ . In light of the above problem, this is equivalent to using the greatest number of terminal node pairs, i.e. the number of possible connections that can be formed within a connected component. In addition to this latter problem, we use the utility metric also in defining greedy approaches to the first problem.

The key assumptions behind these problem statements are that the locations of the terminal nodes are known and that the location information of these nodes can be collected even though the network is not connected. The motivation behind the latter assumption is that depending on the solutions on the physical layer, it can be possible to be able to sustain a low rate of communication over much further distance than to provide quality of service. In this case, the network is able to convey control information even if efficient communications are not possible. In other words, in this problem the transmission range  $r$  models the longest distance that allows direct communication at the rate required by the application to be utilized over the network.

### Review of related problems and existing results

The closest related problem with existing literature is that of the Euclidean *shortest-connection network*, where the task is to form a connected network between a given set of points with minimum total edge length. This problem has two variants, depending on whether or not the addition of new points is allowed before forming the edges: when permitted, the problem is known as the *minimum Steiner tree* (or Steiner minimal tree) problem, and when not, it is the minimum spanning tree problem. Whereas finding the minimum spanning tree is a straightforward task, the Steiner tree problem is NP-hard [GGJ77]. A fundamental relation between the two problems is the greatest possible ratio between their optimal solutions, known as the Steiner ratio [DH90b]: for any set of points in the plane, the total edge length of their Euclidean minimum spanning tree is at most  $2/\sqrt{3} \approx 1.15$  times the total edge length of their Euclidean Steiner minimal tree.

Our problem in its basic form can be seen as a special variation of the shortest-connection network problem, where the geometric graph provides the edges implicitly and we are therefore restricted solely to adding new points. Accordingly, the objective function is then the number of added points.

Our problem could as well be stated with the generalized goal of making the network  $k$ -connected. The corresponding generalization of the shortest-connection network problem without the possibility to add points leads to the graph augmentation problem: in commonly used terms, the task in the *minimum augmentation* problem is to add to a given graph the set of edges with minimum total weight so that the resulting graph is  $k$ -connected. Thus, the case  $k = 1$  corresponds to the minimum spanning tree problem, but for any  $k > 1$  the problem is known to be NP-hard: see, e.g., [Hsu93] and the references therein. In this context, a somewhat related problem has been studied in [BR04], where an ad hoc network of mobile robot nodes is already assumed to be connected, and the task is to move the robots to make the network biconnected so that the total distance travelled by the robots is minimized.

### Heuristic algorithms

Our problem differs from the original shortest-connection network problems, e.g., in that it is solely the addition of new points that it is all about; hence, unlike in the original counterparts, the objective function is integer-

valued. However, as we point out in Publication [5], our problem reduces to the shortest-connection network problem in the limiting case as  $r \rightarrow 0$ : the optimal solution is then to place the relay nodes along the edges of the Euclidean minimum Steiner tree for  $\mathcal{N}$ . In the general case, our problem poses the additional complication that we are not connecting only single points to each other, but connected components in the graph  $\mathcal{G}(\mathcal{N}, r)$ , where the best point for connecting to other components must be chosen.

Therefore, given the complexity of our problem, in Publication [5] we propose heuristic algorithms that are applicable to both of its forms stated earlier. In what follows, we present these algorithms and conclude with a brief analysis of their performance. Our complexity analysis is to a large part based on results gathered in [Aur91]; see the publication for more details.

### *Minimum Spanning Tree algorithm*

Our first algorithm utilizes the minimum spanning tree: if we only require that each relay node or contiguous chain of relay nodes connect exactly two connected components of the graph  $\mathcal{G}(\mathcal{N}, r)$ , the optimal solution is to place the relay nodes along the edges of the Euclidean minimum spanning tree (MST) calculated for the components, when the distance between two components is defined as the shortest distance between two terminal nodes in these distinct components. See Figure 2.15 for an example.

In fact, it is not difficult to show that this MST consists of exactly those edges of the MST for  $\mathcal{N}$  that are longer than  $r$ . To this end, consider, e.g., Kruskal's algorithm [Kru56] for finding the MST applied to the set of points  $\mathcal{N}$ : form the MST by drawing edges between points in the order of increasing length, drawing each edge only if it does not create a cycle with the edges drawn earlier. It is easy to see that when we halt this algorithm at edge length  $r$ , the terminal nodes in any given subtree formed so far are exactly those in a connected component of the graph  $\mathcal{G}(\mathcal{N}, r)$ , and hence the rest of the algorithm also finds the MST for these components.

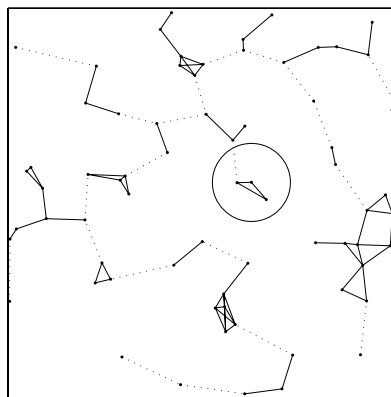


Figure 2.15: Minimum Spanning Tree algorithm. The initially connected components in this example realization of 70 terminal nodes are connected with solid edges, and the edges to place relay nodes are dotted. The transmission range is 10% of the side of the domain, as illustrated by the circle.

If the number of relay nodes available limits the solution, the optimal selection of edges of the MST to place nodes generally requires going through all possibilities. In this case, we propose the greedy method of selecting edges in the order of added utility (with respect to the initial graph  $\mathcal{G}(\mathcal{N}, r)$ ) per used relay node. With this approach, the running time of the algorithm is in any case determined by that of finding the MST, which is  $O(n \log n)$  in the plane.

Note that with a non-negligible transmission range, the Steiner ratio no longer gives a valid approximation ratio for utilizing the MST in our problem: as a simple example, consider a regular pentagon whose vertices are on a circle with radius equal to the transmission range, and assume one terminal node at each of these vertices. These initially disconnected terminal nodes can be connected with a single relay node placed at the center of the circle, whereas the MST suggests placing as many as four relay nodes.

### *Greedy Tessellation algorithm*

The stricter requirement that a single relay node should, when possible, join more than two connected components of the network suggests points that are equally distant from several components as potential points of placement: our second algorithm is based on the observation that the points equally distant from three components are a subset of the vertices, i.e. the coinciding corners of the convex sets also called cells, of the Voronoi diagram of the existing nodes. Note that in practice, points equally distant to more than three nodes do not exist, but placing a relay node at a vertex close to other vertices may well result in connecting more than three components.

This is the motivation for our *Greedy Tessellation* algorithm. In every iteration of the algorithm, we examine the Voronoi diagram of  $\mathcal{N} \cup \mathcal{N}_r$  and

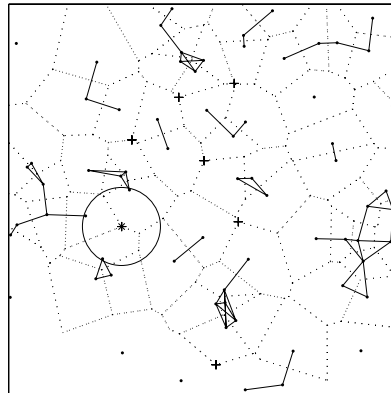


Figure 2.16: An example of the Greedy Tessellation algorithm, when applied to the same realization as in Figure 2.15. The edges of the Voronoi tessellation are shown with dotted lines, the candidate points for relay node insertion with '+'-signs. The first location to add a relay node is marked with an asterisk.

regard as candidate points for node insertion the coinciding corners of such Voronoi cells that contain nodes all in different connected components; the candidate point that yields the maximal increase in the chosen utility metric is added to  $\mathcal{N}_r$ . Finally, if the candidate points are all further than  $r$  from the existing components, the remaining network is connected using the Minimum Spanning Tree algorithm. The final result is an algorithm that takes  $O(n^2)$  time in  $\mathbb{R}^2$  and  $O(n^3)$  in  $\mathbb{R}^3$  to connect the network. Figure 2.16 illustrates the algorithm.

### *Greedy Triangle algorithm*

With a closer look, we see that points equally distant from different components are not always optimal for relay nodes: the point equally distant from three given terminal nodes may fall outside the triangular convex hull of their locations, in which case it cannot be the optimal place for a relay node to connect the three nodes (optimal in the sense that the range required from the relay node to connect the terminal nodes is minimized). For example, the point marked in Figure 2.16 as the place for the first relay node is such a point. Thus, if looking only at the Voronoi diagram, one may not find all the places where connecting three components with a single relay node is possible.

However, for given locations of three disconnected terminal nodes, the optimal place for a relay node can be exhaustively determined: the point equally distant from the terminal nodes is optimal only if it is inside the triangle spanned by the terminal nodes; if the point is outside the triangle, the midpoint of the longest side of the triangle is the optimal place. Similarly,

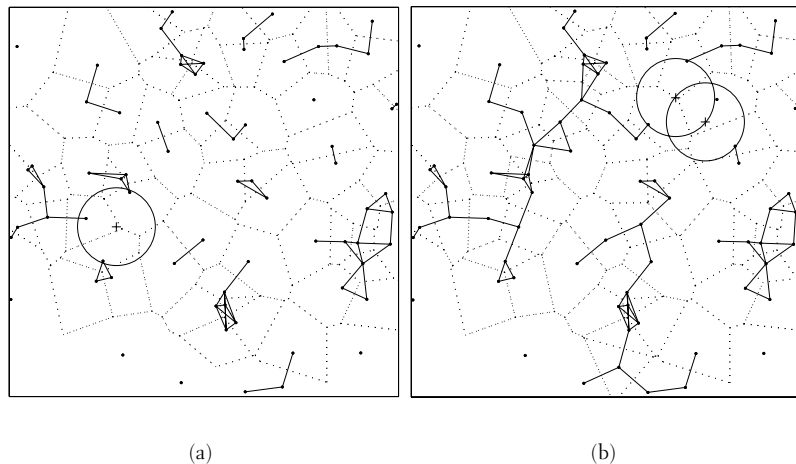


Figure 2.17: Applying the Greedy Triangle algorithm to the realization of Figure 2.15. (a): The first point to place a relay node in the first phase, as indicated by the '+'-sign. Note the difference from Figure 2.16 in the placement. (b): The first pair of points to place relay nodes, as determined in the second phase, after several relay nodes have been added in the first phase. Note that in this case, four components are connected.

one can find jointly optimal places for two relay nodes to connect three terminal nodes. Our analysis of this case, which is already more complicated, is presented in Appendix 4.

The analyses of the above two cases are the basis for our last algorithm. The *Greedy Triangle* algorithm proceeds through two rounds of iterations. The first round consists of repeatedly finding candidate triplets of terminal nodes that are all in different components and pairwise at most  $2r$  apart, and connecting the triplet that yields the maximal increase in the chosen utility metric with a single relay node. This is then repeated in the second round by connecting triplets at most  $4r$  apart with two relay nodes in each iteration. Finally, as in the Greedy Tessellation algorithm, the Minimum Spanning Tree algorithm is invoked to connect the network partitions that remain. The result of one iteration in both the two rounds is illustrated in Figure 2.17. The running time of this algorithm is of the same order as in the previous algorithm in two and three dimensions.

#### *Comparison of the algorithms*

Finally, we present results from applying our three algorithms to simulated realizations of randomly and uniformly distributed terminal nodes in a square-shaped domain in the plane. The purpose is partly to compare the performance of the algorithms relative to each other, and in part to gain some idea on how close to optimal their solutions are. The latter is a problematic task, as the optimal solution for a general realization is usually unknown. We used as a benchmark the method of placing the relay nodes on those edges of the Euclidean minimum Steiner tree for  $\mathcal{N}$  that connect different components of the graph  $\mathcal{G}(\mathcal{N}, r)$ . This method should be close to optimal with sparse networks, i.e. when the transmission range is small compared to the typical distance between neighboring terminal nodes.

Figure 2.18(a) shows the average number of relay nodes needed to connect random configurations with varying number of terminal nodes using each of the different algorithms. The transmission range was set to 10% of the side of the square domain, in order to demonstrate a “feasible” scenario where the number of relay nodes needed is still a fraction of the number of terminal nodes, making the addition of relay nodes sensible. As expected, our three algorithms produce gradually better solutions. The two greedy algorithms also outperform utilizing the Steiner tree with these parameters, as the minimum Steiner tree simply optimizes the wrong measure from our problem’s viewpoint.

The gain from utilizing the Steiner tree is captured in Figure 2.18(b) which shows corresponding results with the transmission range set to 5% of the side of the domain. In a very sparse initial network, the existence of suitable candidate triangles is unlikely, and the Greedy Triangle algorithm practically reduces to the Minimum Spanning Tree algorithm, while the Steiner tree yields the best results. As the density of the initial network increases, the Greedy Triangle algorithm surpasses the Steiner tree method in performance.

The quantity that best describes what we referred to as the density of the network is the average number of other terminal nodes directly connected to a random terminal node in the initial configuration. Ignoring the border effect, this quantity is given by  $n/A \cdot \pi r^2$  where  $A$  is the area of the domain.



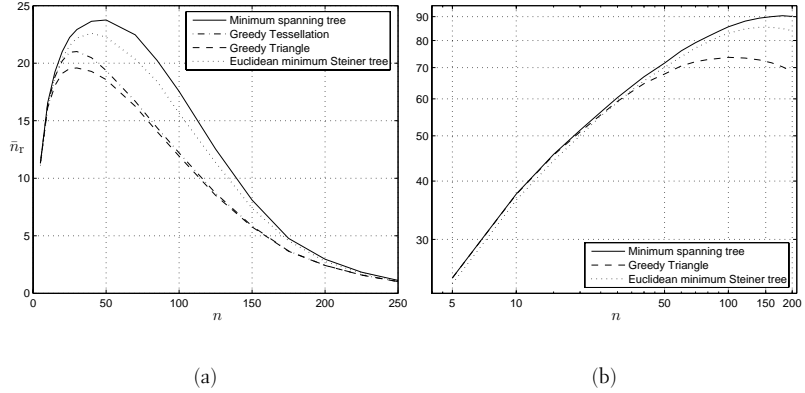


Figure 2.18: Average number of relay nodes needed to connect the network, as a function of the number of terminal nodes initially in the network, taken over 1000 random realizations. (a): The transmission range is 10% of the side of the square-shaped domain. (b): The transmission range is 5% of the side of the domain; the Greedy Tessellation algorithm has been omitted for clarity.

In essence, this quantity determines which method yields the best results, and for example the two greedy algorithms bring significant advantage over the MST algorithm at proper intermediate values of this quantity, when suitable candidate triangles are likely to exist.

## 2.4 Summary and Conclusions

This chapter focused on studies where wireless multihop networks are modeled with Boolean models. When used to describe the network topology, these models implicitly account for some constant-level background noise but not for interference from concurrent transmissions in the network. Thus, such a model is adequate for studying ultimate limits for connectivity, or connectivity in cases where interference is not a limiting factor for communications. The latter is the case in networks with low transmission activity or transmission powers, or when nodes are sparsely located and hence separated by long distances.

We first addressed the problem of determining the  $k$ -connectivity probability of random, finite networks. Scenarios where this problem arises include network dimensioning: for example, how densely should network nodes with given communication capabilities be scattered in order to have a  $k$ -connected network with high probability? As this problem is equivalent to knowing the distribution of the threshold range for  $k$ -connectivity, our first approach to this problem was to model empirically the convergence of these distributions, studied by simulation, to their previously analytically derived asymptotic limits, allowing the prediction of the  $k$ -connectivity probability for finite configurations. To this end, we developed algorithms to

determine the threshold range for 1-, 2-, and 3-connectivity for a given set of network nodes.

We also studied the connectivity of mobile networks, where  $n$  nodes are assumed to move according to the Random Waypoint (RWP) mobility model. In this context, we applied an alternative approach, namely, an analytical approximation where the threshold range for  $k$ -connectivity is treated as the maximum of  $n$  i.i.d.  $k$ -nearest-neighbor distances, which reflects the properties of the known asymptotic statistics.

Overall, because our empirical method is devoid of the systematic error present when assuming that the asymptotic properties of connectivity statistics also hold in the non-asymptotic regime, it provides the most accurate prediction results thus far. However, the requirements for using this approach with good results are that obtaining ample observations from finite configurations is not an obstacle and that the asymptotic distribution is known in analytical form; the case of RWP mobility is an example of where the latter does not hold true.

Next, we discussed the coverage of random networks under the Boolean model, which is of importance when dimensioning a sensor network with randomly scattered nodes. In this context, we formulated the problem as that of determining the distribution of the area coverage caused by randomly placed sensors in a bounded target domain. We showed that the expected value of this distribution is equal to the probability that a random point in the domain is covered, making possible the analytical calculation of this quantity in a simple setting. We also addressed the problem of determining the probability of full coverage and pointed out that this problem, like that of connectivity, is also equivalent to knowing the distribution of a well-defined threshold range. We showed how to determine this range for a given set of nodes and target domain, and interpreted existing analytical results as an asymptotic distribution of this threshold range. We also derived an approximate limit distribution of the threshold range in a case with border effects. With all the same ingredients as with the connectivity problem at hand, we also treated the problem of full coverage as a learning problem, with the result of predicting the distribution of the threshold range for full coverage from independent simulations with good accuracy. To our knowledge, no analytical approximations aiming at accurately predicting the probability of full coverage have been presented to date, which makes our empirical approach all the more important.

Finally, we presented a novel problem representing the algorithmic side of connectivity research. The problem is motivated by a practical disaster recovery scenario: given a disconnected network of terminal nodes deployed at known locations and the possibility of adding relay nodes with the same properties to the network, the task is to connect the network by adding a minimal number of relay nodes. We interpreted this problem as a special variation of the shortest-connection network problem and, in particular, pointed out its connection to the Euclidean minimum Steiner tree problem, known to be NP-hard. We presented three increasingly advanced, centralized heuristic algorithms along with their complexity analysis, and compared their performance by simulation. The more evolved algorithms are most useful at proper intermediate densities of the initial network.

In the assumptions of the last problem, we precluded the mobility of terminal nodes. The approach of adding relay nodes in optimized locations has little application if all the terminal nodes tend to move all over the network region. However, by keeping track of the locations of terminal nodes over time, it should be possible to recognize those nodes that are nearly stationary and place relay nodes to connect these nodes. This question has been taken under further study.



### 3 STUDIES UNDER THE PHYSICAL MODEL

*The studies in this chapter have the assumption of the Physical model as the common denominator. The first section deals with generalizing the notion of connectivity to this model, and the second section discusses throughput in a particular network scenario.*

#### 3.1 Connectivity

In this section, we discuss connectivity when studied under the Physical model. In this case, connectivity is subject to a more difficult choice of definition than under the Boolean model. Whereas in the latter all the variables that dictate whether two nodes can communicate directly can be woven into a single parameter, viz. the transmission range within which direct connections exist, in the Physical model successful communication also depends on the interference from all other nodes that is time-varying in nature. Thus, not only will the existence of a direct link between two nodes be governed by the locations of all nodes in the network, but the condition for its existence should be defined with this temporal aspect in mind. On the other hand, the dependence on the interference means that connectivity is also affected by medium access control.

The existing literature on connectivity under the Physical model is sparse in comparison to that under the Boolean model; we will begin with a review of earlier studies.

##### Review of existing results

The pioneering work in assessing the impact of interferences on connectivity has been made in [DBT03] (see also the journal version [DBT05]) where the authors study a CDMA network. In line with the Physical model, they assume that node  $j$  can successfully receive the signal transmitted by node  $i$  if and only if

$$\frac{P_i \cdot l(\|x_i - x_j\|)}{N_0 + \gamma \sum_{k \neq i, j} P_k \cdot l(\|x_k - x_j\|)} \geq T, \quad (3.1)$$

where the factor  $0 \leq \gamma \leq 1$  weighting the interference power sum is motivated by the partial orthogonality of CDMA codes and can be interpreted as the inverse of the processing gain of the system. Neglecting unidirectional links, an undirected edge is assumed to exist between nodes  $i$  and  $j$  if and only if (3.1) also applies with  $i$  and  $j$  interchanged; the resulting graph is named the Signal To Interference Ratio Graph (STIRG). Note that the case  $\gamma = 0$  reduces to the Boolean model.

Recently refined from the initial result presented in [DBT03, DBT05], it is then shown in [DFM<sup>+</sup>06] that in an infinite STIRG whose nodes are located on the points of a Poisson process on  $\mathbb{R}^2$  and transmit with some common power, i.e.  $P_i \equiv P$ , percolation occurs when the node density  $\lambda$  is greater than the much-studied critical intensity  $\lambda_c$  of the case  $\gamma = 0$  (with

the given  $T \cdot N_0/P$ ), provided that  $\gamma$  is sufficiently small (but not zero). More precisely, assuming that the attenuation function  $l(\cdot) \leq 1$  is continuous, strictly decreasing, and satisfies some additional non-degeneracy conditions, for any node density  $\lambda > \lambda_c(T \cdot N_0/P)$ , there exists  $\gamma^*(\lambda) > 0$  such that for  $\gamma \leq \gamma^*(\lambda)$ , the graph has an infinite connected component. In [DT04], it is further shown that as  $\lambda$  tends to infinity, the condition for the occurrence of percolation is that either  $\gamma^*(\lambda)$  or the SINR threshold  $T$  must decrease at least as fast as  $1/\lambda$ . Since decreasing  $\gamma$  requires using a wider frequency band and decreasing  $T$  implies a lower rate of communication, this means in either case that the available transport capacity per fixed bandwidth decreases. This establishes thus a trade-off between connectivity and capacity.

Also an infinite random network, but employing slotted Aloha instead of CDMA, is analyzed in [BBM04, BBM06]. In the presented medium access scheme, time is slotted and each node is allowed to transmit in any slot with a fixed medium access probability  $p$ . The condition for node  $j$  successfully receiving node  $i$ 's transmission in any time slot – producing what is chosen as the unit throughput from  $i$  to  $j$  over this time slot – now becomes

$$\frac{P_i \cdot l(\|x_i - x_j\|)}{N_0 + \sum_{k \neq i, j} e_k P_k \cdot l(\|x_k - x_j\|)} \geq T, \quad (3.2)$$

where  $e_k$  is the indicator variable of the event that node  $k$  is allowed to transmit in that time slot. The individual permissions to transmit are independent among both nodes and time slots, so that the variables  $\{e_k\}$  in every time slot are independent Bernoulli-distributed random variables with parameter  $p$ .

In this study, connectivity is defined as a dynamic property: the authors show that slotted Aloha allows the transmission of packets over time from any node to any other node, under simple independence and non-degeneracy assumptions on node mobility, and also in the case without mobility if  $N_0 = 0$ , i.e., if there is no background noise. However, this definition places no requirements on the long-term rate of successful communication and can therefore be regarded as coincident with the omnipresent information-theoretic connectivity discussed in Chapter 1.

Thus far, in the context of network models that account for interference – such as the Physical model – little has been done to address connectivity of finite networks as defined in graph theory, namely, the requirement that all node pairs be connected through the network. The aim in Publication [6] is to take a first step in this direction.

### Studying graph connectivity under the Physical model

In Publication [6], we generalize the notion of the threshold range for connectivity from the Boolean model to the Physical model. We do this under the two scenarios reviewed above, i.e., assuming first a CDMA network and then a network using slotted Aloha. As we will see, since there is more than one free parameter in these scenarios, unlike in the Boolean model characterized solely by the transmission range, the threshold range generalizes into a boundary in the space of these parameters. This boundary consists of segments of individual link connectivity constraints that are linear with

respect to at least some of the free parameters, and it can be traced in the parameter space by keeping track of the prevailing network topology and, in particular, the critical link for connectivity by applying a very simple rule.

Like in the definition of the threshold range under the Boolean model, we also assume throughout that all network nodes employ some common constant transmission power  $P$ . Further, although not at all necessary, we restrict ourselves to the commonly used power-law attenuation function

$$l(\|x_i - x_j\|) = (C\|x_i - x_j\|)^{-\alpha}, \quad (3.3)$$

to allow easy scaling of any configuration of network nodes to arbitrary physical node densities. Here,  $C > 0$  sets the scale and  $\alpha > 2$ .

### Connectivity boundary in a CDMA network

We begin with the conceptually simpler scenario with the CDMA network which allows easy and unambiguous determination of the connectivity boundary. We will derive this boundary for a STIRG graph representing a network with given node locations  $\{x_k\}$  and a fixed attenuation exponent  $\alpha$  in the above attenuation function. As in [DBT03], we assume that all nodes transmit constantly, representing a worst-case scenario. In this case, (3.1) is equivalent to

$$A \leq -B \sum_{k \neq i, j} \|x_k - x_j\|^{-\alpha} + \|x_i - x_j\|^{-\alpha}, \quad (3.4)$$

where we have defined the two free parameters  $A = C^\alpha \frac{T}{P/N_0}$  and  $B = T\gamma$ . Hence, in a given network, the condition for node  $j$  successfully receiving node  $i$ 's transmission is satisfied on and below the descending line in the  $B/A$ -plane with slope  $-\sum_{k \neq i, j} \|x_k - x_j\|^{-\alpha}$  and intercept  $\|x_i - x_j\|^{-\alpha}$ , both of which we may calculate (see Figure 3.1 for three instances of such lines). Since our attenuation function is reciprocal, i.e.  $l(\|x_i - x_j\|) = l(\|x_j - x_i\|)$ , which can also be said to hold in reality on a fixed frequency, the resulting condition for an edge existing between nodes  $i$  and  $j$  in the undirected STIRG is determined by this common intercept and the steeper one of the two slopes calculated in both directions (i.e., the receiver under more noise).

The domain in the  $B/A$ -plane in which the given network is connected then lies below a connected curve consisting of segments of descending lines, representing the constraints of different links that are critical for network connectivity at each point; this curve is the connectivity boundary (see Figure 3.1 for an example). It can be found as follows. Start at  $B = 0$ , whence  $\gamma = 0$ . This makes the model coincide with the Boolean model, meaning that the critical link is found as the longest edge in the Euclidean minimum spanning tree of the nodes. The descending line corresponding to this link determines the boundary of the connectivity region as long as no line of the other links is crossed. When this happens, we may determine the new critical link using the following simple rule:

- When a line is crossed from below (as in the first intersection encountered in Figure 3.1 when increasing  $B$  from 0), i.e. the connectivity

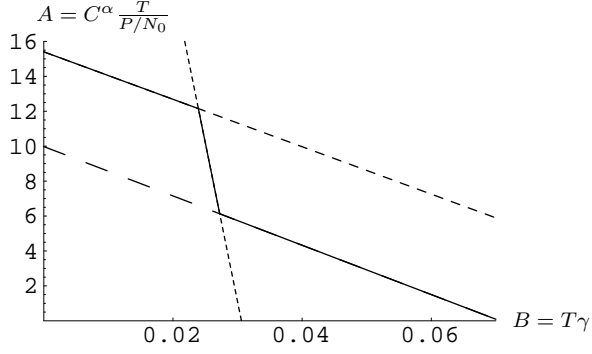


Figure 3.1: The connectivity domain of the network of Figure 3.2 when  $\alpha = 3$ . The solid line indicates the border of the domain. Taking  $A = 0$  implies neglecting background noise; the case  $\gamma = 0 \Rightarrow B = 0$  implies neglecting interference.

region of another link is left when tracing the current line, the corresponding disappearing link is the critical link from this point onwards if the remaining links no longer form a single connected graph. Otherwise, the critical link does not change.

- When a line is crossed from above (as in the second intersection of Figure 3.1), i.e. the connectivity region of another link is entered when tracing the current line, the corresponding appearing link is the critical link from this point onwards if it connects the two network partitions separated by the current critical link. Otherwise, the critical link does not change.

As an example, Figure 3.1 shows the resulting connectivity boundary of the network in a unit square shown in Figure 3.2.

In this model, we assumed that all other nodes transmit constantly. It would of course be more reasonable to assume that at least half of the nodes are receiving instead of transmitting at any instant. In general, if we assume that on average every  $k$ th node is transmitting, we should regard  $\gamma$  as a general interference thinning factor  $\gamma = \tilde{\gamma}/k$ , where  $\tilde{\gamma}$  is the actual code orthogonality factor.

### Connectivity in a slotted-Aloha network

We then extend the definition of the connectivity boundary to the slotted-Aloha network. In addition to the node locations and the attenuation exponent, we assume that the medium access probability  $p$  has been fixed. The equivalent of (3.4) is now

$$A \leq -T \sum_{k \neq i, j} e_k \|x_k - x_j\|^{-\alpha} + \|x_i - x_j\|^{-\alpha}, \quad (3.5)$$

where the sum  $\sum_{k \neq i, j} e_k \|x_k - x_j\|^{-\alpha}$  is a random variable independent in every time slot, having a discrete probability distribution with generally



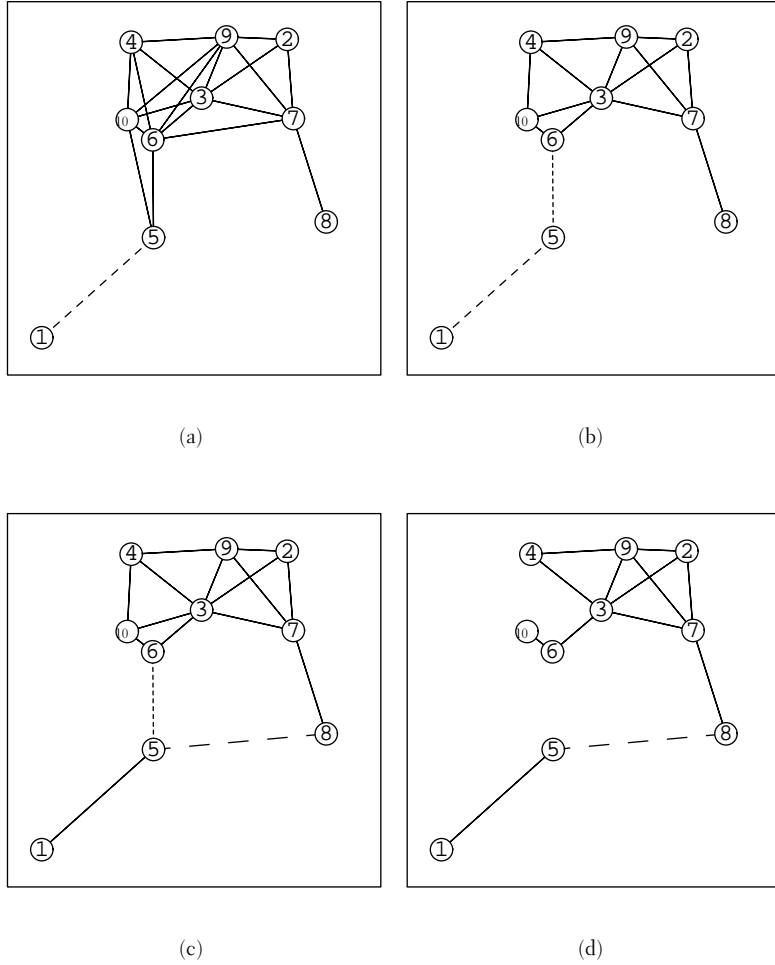


Figure 3.2: An example network in the unit square. The prevailing topology with the identified critical links have been drawn at each of the vertices of the border in Figure 3.1. The dasheding of the links corresponds to that used in Figures 3.1 and 3.3.

$2^{n-2}$  distinct possible values, with  $n$  denoting the number of all the network nodes. Given the assumed information, we may calculate this distribution.

In addition to the free parameters  $A = C^\alpha \frac{T}{P/N_0}$  and  $T$ , we then define a third parameter, *link confidence*  $q$ . For a given  $q$ , the  $q$ -quantile of the distribution of the above random sum describing the scaled interference gives the level below which the latter remains with confidence (probability)  $q$  for each directed link  $i \rightarrow j$ . The difference from the previous section is that instead of the function  $f_{ij}(\gamma) = \gamma \sum_{k \neq i,j} \|x_k - x_j\|^{-\alpha}$  that can be seen in (3.4), which is linear in its argument  $0 \leq \gamma \leq 1$ , we now have the nonlinear function  $F_{ij}^{-1}(q)$  where  $F_{ij}(\cdot)$  is the cumulative distribution function of the random sum  $\sum_{k \neq i,j} e_k \|x_k - x_j\|^{-\alpha}$ . (To be exact, we

define the inverse function of this discrete-valued cumulative distribution as  $F_{ij}^{-1}(q) = \min\{t : F_{ij}(t) \geq q\}$ .) Because of this nonlinearity, the parameter  $q$  can no longer be incorporated into the second free parameter with  $T$  but has to be treated as a separate, third parameter. Note however that the two functions  $f_{ij}(\gamma)$  and  $F_{ij}^{-1}(q)$  coincide at  $\gamma = q = 1$ .

The connectivity boundary is a surface in the space of the three free parameters  $A = C^\alpha \frac{T}{P/N_0}$ ,  $T$ , and  $q$ , a cross-section of which with fixed  $q$  looks similar to Figure 3.1: (3.5) can be written in the form

$$A \leq -F_{ij}^{-1}(q) \cdot T + \|x_i - x_j\|^{-\alpha}, \quad (3.6)$$

which, with fixed  $q$ , is satisfied on and below the descending line in the  $T/A$ -plane with slope  $-F_{ij}^{-1}(q)$  and intercept  $\|x_i - x_j\|^{-\alpha}$ . As in the previous section, we define the condition for nodes  $i$  and  $j$  being bidirectionally connected to be determined by the steeper slope, i.e., the greater interference with the given confidence  $q$ .

With any fixed  $q$ , the cross-section of the connectivity boundary in the  $T/A$ -plane is found exactly as in the previous section, by tracing along the critical links. The reason why the longest edge in the Euclidean minimum spanning tree is again the critical link as  $T \rightarrow 0$  is that in this limit, the above condition (3.6) for every link is dominated by the intercept  $\|x_i - x_j\|^{-\alpha}$  which is a monotonically decreasing function of the link distance.

To demonstrate the connectivity boundary in this case, we examine again the example network of Figure 3.2. We assume that  $\alpha = 3$  and take  $p = 0.1$ , the latter representing a magnitude found suitable for the medium access probability in [BBM04]. Figure 3.3 shows the connectivity domain in the  $T/A$ -plane with  $q$  fixed to different values; it is easy to see that  $q = 1$  leads to the domain of Figure 3.1 with  $T$  in the place of  $B$ . The connectivity surface of the network in the space of all three parameters is depicted in Figure 3.4.

Next, we will show that the connectivity requirement in the slotted-Aloha network constitutes a boundary condition for a tradeoff between the delay and throughput of the network links and allows for an optimization between the two as desired.

To this end, recall that the link confidence  $q$  defines when a pair of nodes is considered directly connected: we say that there is a link from any node  $i$  to any other node  $j$  only if, with given parameters  $A = C^\alpha \frac{T}{P/N_0}$  and  $T$ , the probability  $\tilde{q}$  that (3.2) holds in a random time slot is at least  $q$ . Assume that the conditions for a completely successful transmission from  $i$  to  $j$  are that (i): node  $i$  is allowed to transmit, (ii): node  $j$  does not transmit in the same time slot, and (iii): the states of the remaining nodes are such that (3.2) holds. Then, due to the nodes' independent operation, the number of time slots needed for one successful transmission from  $i$  to  $j$  obeys a geometric distribution with parameter  $p(1-p)\tilde{q}$ . Furthermore, for the critical link in a network with many nodes it is reasonable to assume that  $\tilde{q} \approx q$ . Hence, requiring a higher link confidence means requiring a lower maximum average link delay in the network, whereas allowing a lower link confidence means allowing a higher maximum average link delay.

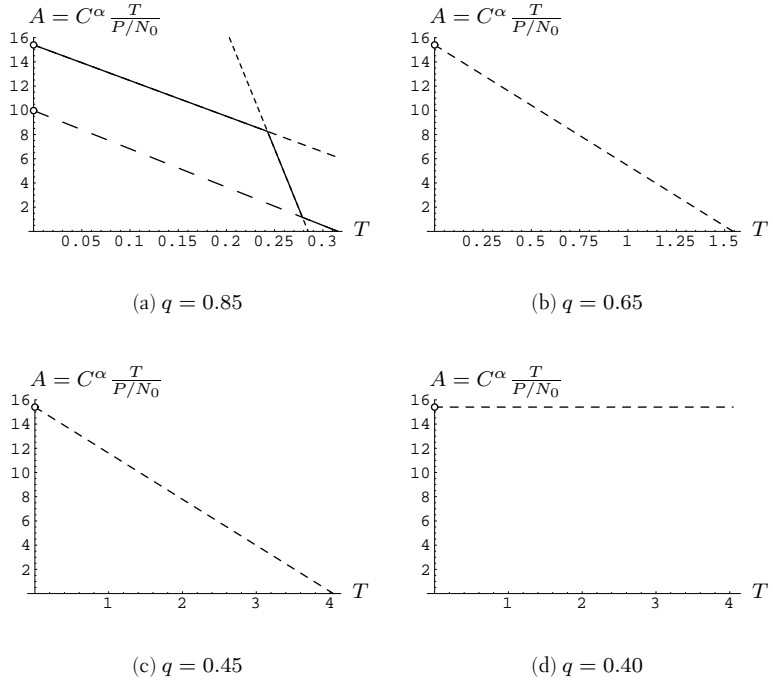


Figure 3.3: The connectivity domain of the network of Figure 3.2 when  $\alpha = 3$  and  $p = 0.1$ , with the boundary indicated by a solid line where needed. Note that in this case, there is no interference in an arbitrary time slot with probability  $(1 - 0.1)^8 \approx 0.43$ , hence the zero-slope with  $q = 0.40$ .

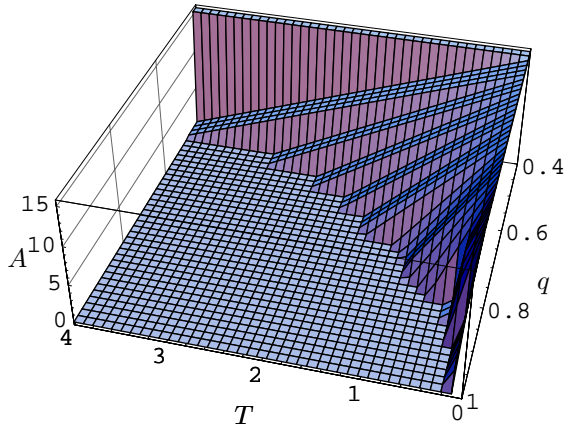


Figure 3.4: The surface below which the network of Figure 3.2 is connected when  $\alpha = 3$  and  $p = 0.1$ . The cross-section of the surface is as in Figure 3.3(d) for all  $q \leq (1 - 0.1)^8 \approx 0.43$ .

On the other hand, the probability  $p(1-p)\bar{q}$  is also the proportion of time slots with successful transmissions from node  $i$  to  $j$  over time. The defined unit throughput of each successful transmission in turn depends on the SINR threshold  $T$ : if we take as a reference the Shannon capacity of a channel with Gaussian noise and interference and a given SINR, this is proportional to  $\log(1 + \text{SINR})$ . Then the minimum time-averaged link throughput in the network is proportional to  $q \log(1 + T)$ .

The two parameters  $q$  and  $T$  that together determine the maximum average link delay and the minimum link throughput are bound together by the connectivity constraint, which dictates that the greatest achievable  $T$  depends on the required  $q$ :  $T_{\max} = T_{\max}(q)$ . For a given network, the required link confidence  $q$  can be increased from zero to some positive value without sacrificing minimum link throughput (with our example network, we know that this value is at least  $q = (1 - 0.1)^8$ ). The maximum of the minimum link throughput with respect to  $q$  marks the beginning of Pareto-optimal combinations of maximum link average delay and minimum link throughput, meaning that neither quantity can be improved without making the other quantity worse: beyond this maximizing value of  $q$ , a delay-throughput tradeoff must then be made according to design preferences.

The computational task of determining the distribution of the interference from a given network of  $n$  nodes for a given transmitter-receiver pair grows exponentially with  $n$ , as it entails evaluating all the possible  $2^{n-2}$  values of the interference and their probabilities. The end of Publication [6] is devoted to presenting a computationally efficient method for determining an approximation for this distribution.

## 3.2 Throughput

This section addresses throughput in networks under the assumptions of the Physical model, both as an aggregate quantity in the whole network and as experienced by an individual node. The most fundamental existing results, which we will discuss first, concern the highest achievable throughput under the restrictions of this model.

### Review of existing results

The aim in the study [GK00] that first presented the Physical model is to determine bounds and scaling laws for the achievable throughput under that model. For this purpose, the authors define bit-meter as the basic unit of information transfer in the network, and the *transport capacity* of the network as the bitrate-distance products summed over all concurrent single-hop transmissions taking place in the network, measured in bit-meters per second, in analogy with, e.g., the passenger-kilometers-per-year metric used by airlines.

The authors then derive bounds for the best-case transport capacity achievable in arbitrary networks, where the various parameters such as node locations, transmission powers and the traffic pattern can be arbitrarily optimized. The main result, further refined in [AK04], is that for a network of  $n$  nodes in a domain with area  $A$  under the Physical model and the power-law attenuation function (3.3), when the maximum transmission

power is bounded by  $\Theta(n^\alpha)$ , the transport capacity has the upper bound  $\Theta(\sqrt{An})$ . This bound is sharp, in the sense that it is also an achievable lower bound. Furthermore, in the case without power constraints, the upper bound  $O\left(n^{\frac{\alpha-1}{\alpha}}\right)$  holds.

In [BBM04, BBM06], Baccelli et al. analyze infinite random slotted-Aloha networks where the transmission powers are assumed independent random variables with identical probability distribution. The authors define an alternative throughput measure, the *spatial density of progress*, which is measured in bits per second crossing a meter of the line perpendicular to the direction of transfer in the planar network. The mean density of progress is proposed as the optimization criterion in selecting the medium access probability  $p$ ; the optimal value depends only on the required SINR threshold  $T$  and, more importantly, not on the density of nodes. This makes a decentralized implementation possible, provided that nodes have some local information on the location of other nodes. In contrast, the authors point out that for optimal spatial reuse and hence optimal throughput scaling in a network using CSMA, the range within which one transmission should be prohibited from another depends on the density of nodes, which is an impediment to the decentralized implementation of CSMA in wireless multihop networks.

Translating the mean density of progress to Gupta and Kumar's transport capacity, it is then shown in [BBM04] that a network with node density  $\lambda$  in the infinite plane transports  $\Theta(\sqrt{\lambda})$  bit-meters per second, per unit area, which is equivalent to the above bound  $\Theta(\sqrt{An})$ . Thus, the upper bound for the scaling of transport capacity in multihop networks is in fact reached by random networks using slotted Aloha. This, along with the possibility for decentralized implementation, certainly makes this simple random-access scheme seem appealing.

Most of the quantitative analysis in [BBM04] assumes that the random transmission powers of the nodes are exponentially distributed and hence unbounded. In Publication [7], we extend the analysis by departing from this assumption, and evaluate the probability of successful transmission – or, more aptly, reception – when all nodes transmit with a common constant transmission power. As the basis of the more evolved analysis also in [BBM04], this probability is one of the most fundamental performance quantities in the network and, as pointed out using the concept of link confidence in the previous section, also determines the long-term throughput of any given link. We discuss this in more detail in what follows.

### Probability of successful transmission in a random slotted-Aloha network

We begin by briefly going through the network modeling assumptions. As in [BBM04], the network that we study is infinite, with nodes located at the points of a Poisson point process  $\Phi = \{X_i\}$  with intensity  $\lambda$  on the plane  $\mathbb{R}^2$ . As with most results in [BBM04], we also assume the power-law attenuation function (3.3) and that the power of the background noise  $N_0 = 0$ , so that the condition for successful reception reduces to the one for the Signal-to-Interference Ratio (SIR). We remark, however, that since the evaluation of the probability of successful transmission reduces to that of the

distribution of the random interference, the assumption  $N_0 = 0$  is made merely to allow compact results; in a practical scenario with numerical parameter values available, the following work could easily accommodate a non-zero noise power.

This setting where we neglect the background noise is equivalent to the extreme limit of an interference-dominated case, which results when transmission activity is high and the transmission powers or the spatial density of the network nodes is increased: these changes all have the same effect of diminishing the parameter  $A$  introduced in the previous section. This can be contrasted with the noise-dominated case which results under the opposite conditions and whose extreme limit is equivalent to the Boolean model.

### Problem statement

Under the above assumptions, by the properties of the homogeneous Poisson process the probability of node  $i$  at  $x_i$  successfully receiving node  $j$ 's transmission only depends on the distance  $d = \|x_j - x_i\|$  and not on the specific locations  $x_i, x_j$ . Thus, without loss of generality we may select  $x_i$  as the origin, whence the condition for successful reception can be written in the following equivalent forms:

$$\begin{aligned} \frac{P(Cd)^{-\alpha}}{I} \geq T &\Leftrightarrow I \leq P(CT^{1/\alpha}d)^{-\alpha} \\ \Leftrightarrow \frac{I}{P(CT^{1/\alpha}d)^{-\alpha}} &= \sum_k e_k \left( \frac{\|X_k\|}{T^{1/\alpha}d} \right)^{-\alpha} \leq 1, \end{aligned} \quad (*)$$

where  $I = \sum_k e_k P(C\|X_k\|)^{-\alpha}$  is the interference power sum at the recipient. In Publication [7], we are interested in the probability that this condition holds given  $d$ . However, as will next be explained, we take a deeper look at this quantity than was done in [BBM04].

First of all, we may ask what is the probability, accounting for both the random locations  $\{X_k\}$  and medium access states  $\{e_k\}$  of all other nodes, that (\*) holds in a random time slot. In other words, if a random configuration  $\{X_k\}$  is observed in a random time slot, what is the probability that (\*) holds, given the distance  $d$ ? Let us write this probability as  $\Pr_{\{X_k\}, \{e_k\}}[(*) \text{ holds} | d]$ . This is the probability that was derived analytically for exponentially distributed transmission powers in [BBM04], and which we will discuss first.

Then we study the probability that (\*) holds in a random time slot for a given configuration of surrounding nodes  $\{X_k\} = \{x_k\}$  representing – and completely characterizing – the interference environment of one receiving node in the network; we write this conditional probability as  $\Pr_{\{e_k\}}[(*) \text{ holds} | d; \{X_k\} = \{x_k\}]$ . This probability is different for different configurations  $\{x_k\}$ , but it is fully determined once given the  $\{x_k\}$ . It is thus a function of the random node locations  $\{X_k\}$  and therefore itself a random variable with a probability distribution over  $\{X_k\}$ . This distribution describes how different nodes in the network are in different positions with regard to the success of communication. In fact, the probability  $\Pr_{\{X_k\}, \{e_k\}}[(*) \text{ holds} | d]$  discussed in the next section can be seen to be the expected value of this distribution over  $\{X_k\}$ .

### Probability of successful reception: expected value over $\{X_k\}$

We begin with the probability that (\*) holds for a random configuration  $\{X_k\}$  in a random time slot, given the distance  $d$  of the transmitting node from the receiver. Determining this probability reduces to knowing the distribution of the random interference power sum  $I$ ; as will become evident shortly, the Laplace transform of such a power sum is known in closed form, up to a certain integral. In our case, this integral requires resorting to numerical evaluation, making the inversion of the Laplace transform difficult. For this reason, we utilize a decomposition of the total interference power into two parts. By deriving the distribution of the interference power from some neighborhood of the receiver exactly, we are left with the remaining interference whose distribution can be deduced to be approximately Gaussian; the larger that neighborhood, the smaller the difference. The Gaussian approximation can be improved by utilizing the known Laplace transform with the aid of so-called Bahadur-Rao approximation. This way, combining the exact distribution of the near-by interference power and the approximate one of the remaining interference, we obtain a numerical approximation for the probability we are interested in that can be improved to an arbitrary level of accuracy. We sketch out the details in what follows.

Because we are interested in any random configuration for only one time slot, we may limit our attention in any configuration to the nodes that transmit in that time slot. By the properties of the Poisson process, we may then write the interference power sum in (\*) as  $I = I_{\Phi(\lambda p)} = P \sum_k (C||Y_k||)^{-\alpha}$ , the shot noise of a Poisson process  $\{Y_k\}$  with intensity  $\lambda p$  at the origin. As pointed out in [BBM04], the Laplace transform of a general Poisson shot noise  $I_{\Phi(\lambda)}$  with i.i.d. transmission powers  $\mathcal{P}_k \sim \mathcal{P}$ , calculated at the origin, is

$$I_{\Phi(\lambda)}^*(s) = \exp \left( -\lambda \int_{\mathbb{R}^2} 1 - \mathbb{E}_{\mathcal{P}} [\exp(-s\mathcal{P}(C||x||)^{-\alpha})] dx \right). \quad (3.7)$$

The generalization of the above to the interference at the origin from transmitters in some arbitrary domain simply amounts to integrating over that domain instead of  $\mathbb{R}^2$ .

Let  $P_r$  denote the power received from a transmitter at distance  $r$ , i.e.  $P_r = P(Cr)^{-\alpha} = P_{T^{1/\alpha}d} \left( \frac{r}{T^{1/\alpha}d} \right)^{-\alpha}$ . By (3.7), the Laplace transform of  $I$  is now

$$I_{\Phi(\lambda p)}^*(s) = \exp \left[ \lambda p 2\pi \int_0^\infty \left( e^{-sP_{T^{1/\alpha}d} \left( \frac{r}{T^{1/\alpha}d} \right)^{-\alpha}} - 1 \right) r dr \right],$$

from which it follows that

$$I_{\Phi(\lambda p)}^* \left( \frac{s}{P_{T^{1/\alpha}d}} \right) = \exp \left[ \lambda p 2\pi (T^{1/\alpha}d)^2 \int_0^\infty \left( e^{-st^{-\alpha}} - 1 \right) t dt \right], \quad (3.8)$$

where the last expression can also be seen as the Laplace transform of  $I/P_{T^{1/\alpha}d}$ . Note that it is precisely  $I/P_{T^{1/\alpha}d}$  whose distribution we are interested in, since (\*) is also equivalent to  $I/P_{T^{1/\alpha}d} \leq 1$ . (With a positive background noise power  $N_0$ , the corresponding condition would be  $I/P_{T^{1/\alpha}d} \leq 1 - N_0/P_{T^{1/\alpha}d}$ , the probability of which reduces to zero as

the distance  $d$  reaches the value with which  $P_{T^{1/\alpha}d} = N_0$ ; this value equals the transmission range of the Boolean model.)

A closer look reveals that  $I/P_{T^{1/\alpha}d}$  has infinite expectation and variance. This is a side effect caused by the assumed power-law attenuation function (3.3) which has a singularity at zero distance. This shortcoming is implicitly dealt with by the following decomposition of the interference.

We will treat the total interference power  $I$  as the sum of two parts. The distribution of one part is approximated using its Laplace transform, whereas the distribution of the other part is calculated exactly. The key observation allowing this division is that for (\*) to hold, there may be at most  $m$  active transmitters at distances  $y$  satisfying

$$\frac{P(Cd)^{-\alpha}}{(m+1)P(Cy)^{-\alpha}} < T \quad \Leftrightarrow \quad y < [(m+1)T]^{1/\alpha}d \stackrel{\text{def}}{=} r_m,$$

i.e.,  $m+1$  active transmitters alone at distance  $r_m$  would still satisfy the condition (\*), but moving them any closer would violate this condition. Now, we will partition the total interference power into two parts originating from different zones, i.e.  $I_{\Phi(\lambda p)} = I = I_{\text{in}} + I_{\text{out}}$ , where  $I_{\text{in}}$  denotes the interference power originating from distances up to  $r_m$  and  $I_{\text{out}}$  denotes that from distances beyond  $r_m$  (see Figure 3.5).

#### Distribution of $I_{\text{in}}$

Because of the above limitation to at most  $m$  active transmitters within  $r_m$ , we may determine the distribution of  $I_{\text{in}}$  exactly, as follows. By the properties of the Poisson process, given the number of nodes in the inner zone, their locations in that zone are i.i.d. uniformly distributed. Thus, the distribution of the interference  $I_1$  from a single node in this zone can be easily determined: denoting by  $R$  the distance of the node from the origin

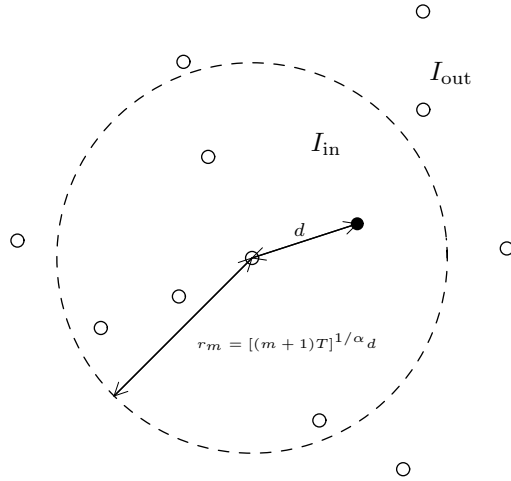


Figure 3.5: Division of the interference into that originating from two zones, for some  $(m+1)T > 1$



and noting that  $I_1 = P_R = P_{T^{1/\alpha_d}} \left( \frac{R}{T^{1/\alpha_d}} \right)^{-\alpha} \stackrel{def}{=} P_{T^{1/\alpha_d}} Q$  where  $P_{T^{1/\alpha_d}}$  is constant, we easily obtain the probability density of  $Q$  as

$$f_Q(q) = \begin{cases} \frac{2}{\alpha(m+1)^{2/\alpha}} q^{-(2+\alpha)/\alpha}, & q > \frac{1}{m+1}, \\ 0, & q \leq \frac{1}{m+1}. \end{cases}$$

The total interference power from  $i$  nodes in the inner zone is then the sum of i.i.d. random variables  $P_{T^{1/\alpha_d}} \cdot \sum_{j=1}^i Q_j$ , where the probability density of  $\sum_{j=1}^i Q_j$  is obtained as the convolution of  $i$  instances of the above density  $f_Q(q)$ , which we denote by  $f_Q^{*i}(q)$ . Given the condition that there are at most  $m$  active transmitters within  $r_m$ , the conditional distribution of  $I_{in}$  is then obtained by conditioning on  $i$  with the Poisson distribution with parameter  $\lambda p \pi r_m^2$ , truncated at  $m$ :

$$f_{I_{in}/P_{T^{1/\alpha_d}}}(q) = \frac{\sum_{i=0}^m \frac{(\lambda p \pi r_m^2)^i}{i!} f_Q^{*i}(q)}{\sum_{k=0}^m \frac{(\lambda p \pi r_m^2)^k}{k!}}.$$

For example, with  $m = 2$  this conditional distribution can still be calculated analytically (see Publication [7]).

#### Distribution of $I_{out}$

Let us next turn to approximating the distribution of  $I_{out}$ . Applying (3.7) again yields the Laplace transform

$$\begin{aligned} I_{out}^*(s) &= \exp \left[ \lambda p 2\pi \int_{r_m}^{\infty} \left( e^{-s P_{r_m} \left( \frac{r}{r_m} \right)^{-\alpha}} - 1 \right) r dr \right] \\ \Rightarrow I_{out}^* \left( \frac{s}{P_{r_m}} \right) &= I_{out}^* \left( \frac{s}{P_{T^{1/\alpha_d}} / (m+1)} \right) \\ &= \exp \left[ \lambda p 2\pi r_m^2 \int_1^{\infty} \left( e^{-st^{-\alpha}} - 1 \right) t dt \right], \end{aligned}$$

where the last expression can also be seen as the Laplace transform of  $J \stackrel{def}{=} I_{out}/P_{r_m} = (m+1)I_{out}/P_{T^{1/\alpha_d}}$ . After a change of variables  $u = t^{-(\alpha-2)}$ , we then have for the logarithmic moment generating function  $\varphi(\beta) = \log E[e^{\beta J}] = \log J^*(-\beta)$  of  $J$ ,

$$\begin{aligned} \varphi(\beta) &= 2\lambda p \pi r_m^2 \int_1^{\infty} (e^{\beta t^{-\alpha}} - 1) t dt = \frac{2\lambda p \pi r_m^2}{\alpha - 2} \int_0^1 \frac{e^{\beta u^{\alpha/(\alpha-2)}} - 1}{u^{\alpha/(\alpha-2)}} du, \\ \varphi'(\beta) &= \frac{2\lambda p \pi r_m^2}{\alpha - 2} \int_0^1 e^{\beta u^{\alpha/(\alpha-2)}} du, \\ \varphi''(\beta) &= \frac{2\lambda p \pi r_m^2}{\alpha - 2} \int_0^1 u^{\alpha/(\alpha-2)} e^{\beta u^{\alpha/(\alpha-2)}} du, \end{aligned} \quad (3.9)$$

which yields, e.g., the mean and variance of  $I_{out}$  as  $E[I_{out}] = P_{r_m} \varphi'(0)$  and  $\text{Var}[I_{out}] = P_{r_m}^2 \varphi''(0)$ .

Now, consider generating a random realization of  $I_{out}$ , taking only nodes within some maximum distance into account for conceptual simplicity. This can be done by drawing the Poisson-distributed number of

interfering nodes, placing these nodes independently and uniformly at random on the considered domain, neither closer than  $r_m$  nor further than the maximum distance, and calculating  $I_{\text{out}}$  as the sum of the individual interference powers. Thus,  $I_{\text{out}}$  is the sum of i.i.d. random variables, and hence, by the Central limit theorem (see, e.g., [Lin22]), should obey a distribution that tends to the Gaussian as the node density tends to infinity. More precisely, provided that the density is so large that there is likely to be many nodes at the smaller distances with nearly equal contributions to  $I_{\text{out}}$ , the distribution of  $I_{\text{out}}$  – and hence that of the scaled quantity  $J$  – should be close to Gaussian.

Note that given  $\alpha$ , the distribution of  $J$  is fully characterized by the product  $\lambda p r_m^2$ . In fact, the quantity  $\lambda p \pi r_m^2$  is the expected number of transmitting nodes within distance  $r_m$  from an arbitrary reference point. This quantity also determines how close to Gaussian the distribution of  $I_{\text{out}}$  – and hence  $J$  – is: if this number is small, then the total interference power is likely to be dominated by few terms, resulting in a distribution far from the Gaussian. Accordingly, the larger the value, the better the Gaussian approximation. Figure 3.6 shows how the Gaussian approximation agrees with the simulated distribution of  $J$  with different values of  $\lambda p \pi r_m^2$ . To facilitate the simulation, interference from distances beyond  $k \cdot r_m$ , with  $k$

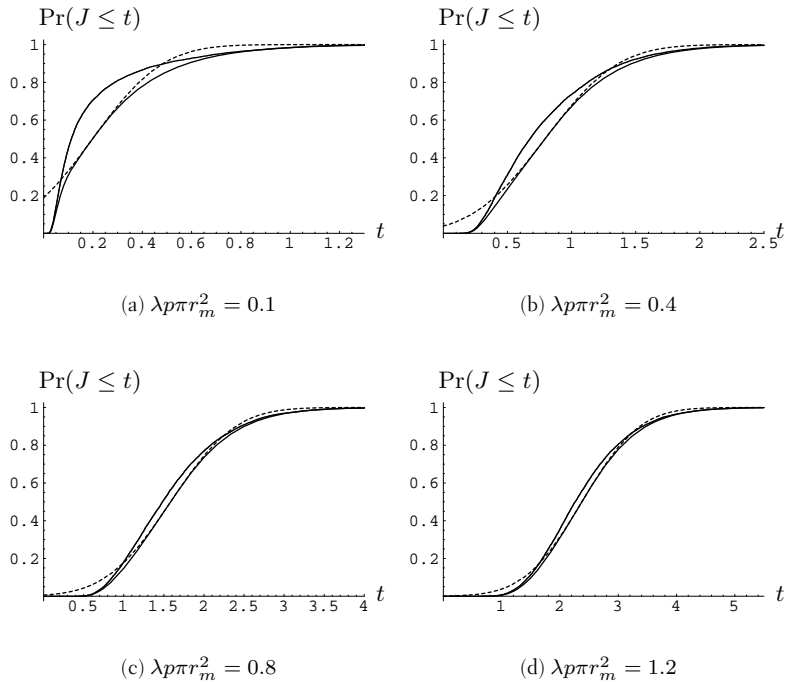


Figure 3.6: Cumulative distribution of  $J = I_{\text{out}}/P_{r_m}$  with different values of  $\lambda p \pi r_m^2$  when  $\alpha = 3$ . The simulation domain has been defined using  $k = 100$ . Upper line: simulated distribution; dashed line: Gaussian approximation; lower line: Bahadur-Rao approximation.

chosen to be some large number, was neglected; this changes the results in (3.9) so that the upper limit  $\infty$  in any of the integrals now becomes  $k$  and the lower limit 0 becomes  $k^{-(\alpha-2)}$ . One can see in the figure that the accuracy of the approximation indeed improves as  $\lambda p \pi r_m^2$  increases, but the fit at the tails of the distribution remains poor. This can be remedied by adopting an approximation from large deviations theory (see [BR60, Section 6]) and applying it also below the mean; this approximation is introduced in the appendix of Publication [7]. This Bahadur-Rao (BR) approximation significantly improves the fit in the tails, while it coincides with the Gaussian approximation at the mean of the distribution: it is also shown in Figure 3.6.

*Adding up  $I_{\text{in}}$  and  $I_{\text{out}}$*

We may now combine the means to evaluate the distributions of  $I_{\text{in}}$  and  $I_{\text{out}}$  to obtain an approximation for the probability that (\*) holds: this can be written as

$$\begin{aligned}
& \Pr [(I_{\text{in}} + I_{\text{out}})/P_{T^{1/\alpha}d} \leq 1] \\
&= \Pr(\text{at most } m \text{ active transmitters within } r_m) \\
&\quad \times \int_0^1 f_{I_{\text{in}}/P_{T^{1/\alpha}d}}(q) \Pr [I_{\text{out}}/P_{T^{1/\alpha}d} \leq 1 - q] dq \\
&= e^{-\lambda p \pi r_m^2} \sum_{k=0}^m \frac{(\lambda p \pi r_m^2)^k}{k!} \int_0^1 \frac{\sum_{i=0}^m \frac{(\lambda p \pi r_m^2)^i}{i!} f_Q^{*i}(q)}{\sum_{k=0}^m \frac{(\lambda p \pi r_m^2)^k}{k!}} \\
&\quad \times \Pr[(m+1)I_{\text{out}}/P_{T^{1/\alpha}d} \leq (m+1)(1-q)] dq \\
&= e^{-\lambda p \pi r_m^2} \int_0^1 \sum_{i=0}^m \frac{(\lambda p \pi r_m^2)^i}{i!} f_Q^{*i}(q) \Pr[J \leq (m+1)(1-q)] dq. \quad (3.10)
\end{aligned}$$

The accuracy of different approximations for  $\Pr_{\{X_k\}, \{e_k\}}[(*) \text{ holds} | d]$  is demonstrated in Figure 3.7. In comparison, we point out that the Laplace transform (3.8) of  $I/P_{T^{1/\alpha}d}$  can also be used directly, by applying the BR approximation; this method has also been included in the figure and can be seen to result in a very poor approximation. On the other hand, using (3.10) with  $m = 2$  already proves to be notably accurate and gives a significant improvement from using  $m = 0$ . One may also note in Figure 3.7(a) how the probability behaves differently under the assumption of exponentially distributed transmission powers. In particular, the probability makes a sharper transition with increasing distance  $d$  in our case. This is because the signal-to-interference ratio at reception varies less due to the lack of randomness in the transmission powers.

Note that we can improve the approximation to any level of accuracy by choosing sufficiently large  $m$ . The gain from increasing  $m$  is twofold. First, through increasing  $\lambda p \pi r_m^2$ , it makes it possible to approximate the distribution of  $J$  more accurately, as shown by Figure 3.6. Second, it decreases the share of  $I_{\text{out}}$  in the total interference, thus mitigating the effect of the remaining inaccuracy. The cost of increasing  $m$  is the added numerical labor in computing further convolutions  $f_Q^{*i}(q)$ .

We conclude by noting that a similar decomposition of interferers into near-by and distant ones is used in [WYAd05], in determining upper and

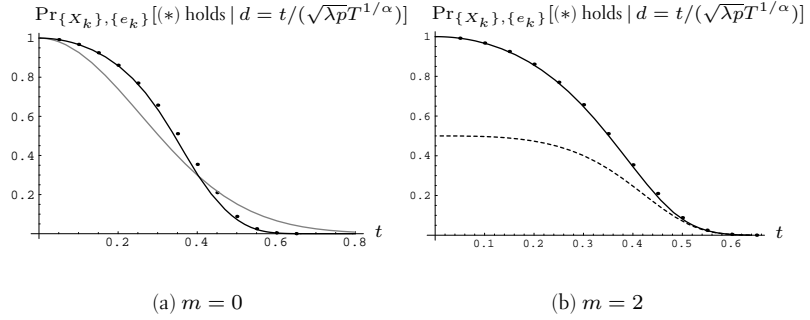


Figure 3.7: The probability of (\*) holding when  $\alpha = 3$ , determined using (3.10) and applying the BR approximation with different values of  $m$  (solid lines), and by simulation (points). For comparison, the corresponding probability in the case of exponentially distributed transmission powers as derived in [BBM04] (gray line) and the result of applying the BR approximation directly to the Laplace transform (3.8) of  $I/P_{T^{1/\alpha}d}$  (dashed line) are also shown. The interference from distances beyond  $100T^{1/\alpha}d$  has been neglected in all cases.

lower bounds for what is called the *transmission capacity* of a CDMA network under the Physical model. To bound the probability of outage, the complement event of (\*), the authors utilize the probability that a certain neighborhood of the receiver is free of interferers (which results in a lower bound for the outage probability), along with a bound for the remaining interference power obtained using the Chebyshev inequality. Although the resulting bounds were quite crude, the authors consider the cases of both one common and individual transmission powers: in our terms, the used decompositions correspond to some real-valued  $m \in [0, 1]$  optimized in each case.<sup>1</sup>

### Probability of successful reception: distribution over $\{X_k\}$

We now turn to evaluating the probability that (\*) holds in a random time slot, given the distance  $d$  and the configuration of surrounding nodes  $\{X_k\} = \{x_k\}$ . As we mentioned earlier, this probability is a function of  $\{X_k\}$  and therefore itself a random variable; for brevity, we will use the notation  $\Pr_{\{e_k\}}[(*) \text{ holds} \mid d; \{X_k\}] = \Pi(\{X_k\})$ . We are interested in the distribution of  $\Pi(\{X_k\})$ .

With  $\alpha$  and  $T$  fixed, the last form in (\*) can be seen to be a condition for the transmission indicators  $\{e_k\}$  and the distances to the other nodes from the recipient, relative to the distance  $T^{1/\alpha}d$ . Thus, with  $\{X_k\}$  fixed,  $\Pi(\{X_k\})$  only depends on the medium access probability  $p$ . Under the assumption that  $\{X_k\}$  is a realization of a Poisson process, the distribution of  $\Pi(\{X_k\})$  then depends only on  $p$  and the average number of nodes

<sup>1</sup>In fact, the preliminary version [WYdA04] of this reference appeared before our work in Publication [7], but until pointed out by Dr. Błaszczyszyn during the pre-examination of this thesis, we were unaware of [WYdA04, WYAd05].

within distance  $T^{1/\alpha}d$ , equal to  $\lambda\pi(T^{1/\alpha}d)^2$ . This should be contrasted with the fact that the averaged probability  $\Pr_{\{X_k\},\{e_k\}}[(*) \text{ holds} \mid d]$  studied above, i.e. the mean of the distribution of  $\Pi(\{X_k\})$ , only depends on the product  $\lambda p\pi(T^{1/\alpha}d)^2$ , a scaling result pointed out already in [BBM04].

In what follows, we will concentrate on the tail probability  $\Pr_{\{X_k\}}[\Pi(\{X_k\}) > \hat{P}]$ . Let  $I(\mathcal{S})$  denote the random interference power originating from the set  $\mathcal{S} \subseteq \mathbb{R}^2$  and observed by the recipient at the origin in any time slot, with the convention  $I = I(\mathbb{R}^2)$ . Also, let  $\mathcal{B}_r$  denote the disk with radius  $r$  centered at the origin, i.e.  $\mathcal{B}_r = \{x : \|x\| \leq r\}$ , and denote its complement with  $\bar{\mathcal{B}}_r$ .

With this notation, let us assume that the locations of nodes in some neighborhood  $\mathcal{B}_r$  and hence the probability distribution of  $I(\mathcal{B}_r)$  in any time slot are given, and focus on the conditional tail probability of  $\Pi(\{X_k\})$ ,

$$\Pr_{\{X_k \in \bar{\mathcal{B}}_r\}} \left\{ \Pi(\{X_k\}) > \hat{P} \mid F_{I(\mathcal{B}_r)}(t) \right\}, \quad (3.11)$$

where we have denoted the distribution of  $I(\mathcal{B}_r)$  by its cumulative distribution function  $F_{I(\mathcal{B}_r)}(t)$ . We will now derive a method to approximately evaluate this conditional probability. Let us partition the exterior of  $\mathcal{B}_r$  into an annulus with inner radius  $r$  and some outer radius  $\hat{r} > r$  and the rest, as shown in Figure 3.8, and make the approximation that the interference from every transmitting node in the annulus (drawn as black points in the figure) is equal to  $P_{\hat{r}}$ , for some  $r \leq \tilde{r} \leq \hat{r}$ . Then, given the number of nodes  $N$  in the annulus, the interference originating from the annulus in each time slot is  $I(\bar{\mathcal{B}}_r \cap \mathcal{B}_{\hat{r}}) \mid N = X \cdot P_{\hat{r}}$  with  $X \sim \text{Bin}(N, p)$ , and we have  $I(\mathcal{B}_{\hat{r}}) \mid N = I(\mathcal{B}_r) + I(\bar{\mathcal{B}}_r \cap \mathcal{B}_{\hat{r}}) \mid N$ . Since this makes it possible to compute the distribution of  $I(\mathcal{B}_{\hat{r}}) \mid N$ , conditioning on  $N$  – which in

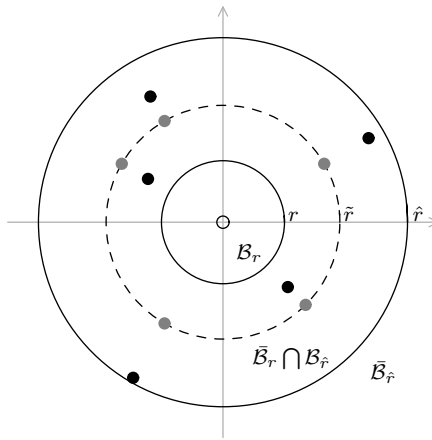


Figure 3.8: Schematic representation of the partitioning of  $\bar{\mathcal{B}}_r$  and the approximation made in the annulus  $\bar{\mathcal{B}}_r \cap \mathcal{B}_{\hat{r}}$

our model is Poisson-distributed with mean  $\lambda\pi(\hat{r}^2 - r^2)$  – now leads to the following recursion for the conditional probability (3.11):

$$\Pr_{\{X_k \in \bar{\mathcal{B}}_r\}} \left\{ \Pi(\{X_k\}) > \hat{P} \mid F_{I(\mathcal{B}_r)}(t) \right\} = \sum_{n=0}^{\infty} \Pr(N = n) \cdot \Pr_{\{X_k \in \bar{\mathcal{B}}_{\hat{r}}\}} \left\{ \Pi(\{X_k\}) > \hat{P} \mid F_{I(\mathcal{B}_{\hat{r}})|N}(t|n) \right\}.$$

In fact, by starting with  $r = 0$  and  $F_{I(\mathcal{B}_0)}(0) = 1$ , this recursion can be used to evaluate the tail probability  $\Pr_{\{X_k\}}[\Pi(\{X_k\}) > \hat{P}]$  to an arbitrary level of accuracy, by partitioning the plane into sufficiently many thin annuli.

Of course, this recursion in itself is infinite, through the infinitely many possible values of  $N$  on the one hand and through the partitioning of  $\mathbb{R}^2$  into an infinite number of annuli on the other. Proper pruning and termination conditions are therefore needed. The first and obvious termination condition is that (3.11) equals 0 for such a distribution of  $I(\mathcal{B}_r)$  for which  $\Pr[I(\mathcal{B}_r) \leq P_{T^{1/\alpha}d}] \leq \hat{P}$ . Since a high enough value of  $N$  gives  $I(\mathcal{B}_{\hat{r}}) \mid N$  such a distribution, we only need to add new terms to the above sum as long as the conditional probability  $\Pr_{\{X_k \in \bar{\mathcal{B}}_{\hat{r}}\}} \left\{ \Pi(\{X_k\}) > \hat{P} \mid F_{I(\mathcal{B}_{\hat{r}})|N}(t|n) \right\}$  on the right-hand side differs from zero by this termination condition.

As for dealing with the infinite plane, we may, for some  $r_{\max}$ , ignore how different configurations of nodes in  $\bar{\mathcal{B}}_{r_{\max}}$  result in different distributions of  $I(\bar{\mathcal{B}}_{r_{\max}})$ , and instead use the distribution averaged over all possible configurations, as if the configuration of nodes producing the interference was different in every time slot. This amounts to approximating the distribution of interference from transmitters in  $\bar{\mathcal{B}}_{r_{\max}}$ , located according to a Poisson process with intensity  $\lambda p$ , by utilizing its Laplace transform exactly as done earlier. With such an approximation for the distribution of  $I(\bar{\mathcal{B}}_{r_{\max}})$ , the final level of recursion simply gives

$$\Pr_{\{X_k \in \bar{\mathcal{B}}_{r_{\max}}\}} \left\{ \Pi(\{X_k\}) > \hat{P} \mid F_{I(\mathcal{B}_{r_{\max}})}(t) \right\} = \begin{cases} 1, & \Pr [I(\bar{\mathcal{B}}_{r_{\max}}) + I(\mathcal{B}_{r_{\max}}) \leq P_{T^{1/\alpha}d}] > \hat{P}, \\ 0, & \Pr [I(\bar{\mathcal{B}}_{r_{\max}}) + I(\mathcal{B}_{r_{\max}}) \leq P_{T^{1/\alpha}d}] \leq \hat{P}, \end{cases}$$

where the probability is calculated by conditioning on  $I(\mathcal{B}_{r_{\max}})$ , which has a discrete distribution with a finite number of values.

Because of the scaling result that applies to the distribution of  $\Pi(\{X_k\})$ , the parameter that completely characterizes the above recursion for evaluating the tail probability  $\Pr_{\{X_k\}}[\Pi(\{X_k\}) > \hat{P}]$  is an increasing sequence  $\{r/(T^{1/\alpha}d)\}$  of distances  $r$ , starting with zero and ending with  $r_{\max}$ , given relative to  $T^{1/\alpha}d$ . These are the outer radii of the nested annuli to consider at the successive levels of recursion. Because no other node may transmit within distance  $T^{1/\alpha}d$  for (\*) to hold, it is sensible to choose the first two distances as  $\{r/(T^{1/\alpha}d)\} = \{0, 1\}$ . Our method of choosing the remaining distances is to fix  $r_{\max}$  and the number of annuli to divide the distances  $[T^{1/\alpha}d, r_{\max}]$ , and select the annuli so that the expected interference from each annulus is equal, i.e. the integral  $\int_r^{\hat{r}} \lambda p 2\pi t P(Ct)^{-\alpha} dt$  is the same for each annulus.

The choice of  $\tilde{r}$  with which the interference from every node in an annulus with inner and outer radius  $r$  and  $\hat{r}$ , respectively, is taken to be  $P_{\tilde{r}}$ , determines the nature of our approximation: choosing  $\tilde{r} = r$  naturally results in a conservative approximation, whereas setting  $\tilde{r} = \hat{r}$  results in underestimating the interference. Aiming at an approximation as accurate as possible, we may choose  $P_{\tilde{r}}$  as the expected interference power from a node placed uniformly at random in the annulus; this is our choice in the demonstration that follows. The fact that this expected interference from the inmost, degenerate annulus is infinite does not affect the final result, since the condition (\*) in any case prohibits all nodes in this annulus from transmitting.

We conclude with a validation of this recursion. In Figure 3.9 we compare the results given by the recursion with simulated distributions of  $\Pi(\{X_k\})$ . Each simulated sample represents the proportion of 1000 time slots in which (\*) was satisfied in a given configuration of nodes, and 10000 random configurations were considered. The two subplots show how the accuracy of the recursion improves as the range covered by the annuli is increased and a larger number of annuli is used. This is particularly clear in the latter subplot, where the most accurate setting already required rather extensive computation time from the recursion, due to the high value of  $\lambda\pi(T^{1/\alpha}d)^2$ , i.e. wide ranges of numbers of nodes to consider in each annulus. For comparison, upper and lower bounds obtained by choosing  $\tilde{r} = r$  and  $\tilde{r} = \hat{r}$  have also been plotted; these points are connected with dashed lines in the figure.

The  $(T, p)$ -pairs selected for the two validation cases represent values of  $p(T)$  maximizing the mean density of progress with exponentially distributed transmission powers as presented in [BBM04]. The quantity  $\lambda\pi d^2$

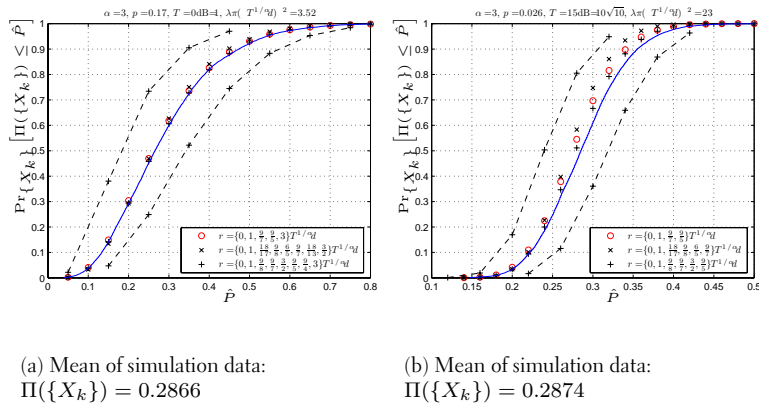


Figure 3.9: Cumulative distribution functions of 10000 simulated estimates of  $\Pi(\{X_k\})$  (solid curves) and  $1 - \Pr_{\{X_k\}} \left[ \Pi(\{X_k\}) > \hat{P} \right]$  as determined by the recursion using different radius sequences (see legends), with  $\alpha = 3$  and  $\lambda p \pi (T^{1/\alpha} d)^2 = 0.598$ . To ease simulations, interference from distances beyond  $10T^{1/\alpha}d$  was neglected in all cases.

was chosen to make  $\lambda p \pi (T^{1/\alpha} d)^2$  the same in both cases, implying the same means for the two distributions. The means obtained from the simulation data can be compared with Figure 3.7, considering that here  $\sqrt{\lambda p} T^{1/\alpha} d \approx 0.44$ . The slightly lower value predicted by the figure is due to the fact that here we have only taken interference from distances up to  $10T^{1/\alpha} d$  into account, as opposed to  $100T^{1/\alpha} d$  in Figure 3.7; substituting these distances in the place of  $r_m$  along with the assumed  $\alpha = 3$  in the mean  $E[I_{\text{out}}] = P_{r_m} \varphi'(0)$  obtained from (3.9), we see that we have here neglected 1/10 of the expected interference from distances beyond  $T^{1/\alpha} d$ , whereas only 1/100 was neglected in Figure 3.7.

The fact that the latter distribution with a higher value of  $\lambda \pi (T^{1/\alpha} d)^2$  has a smaller variance can be explained by the fact that the number of nodes in any annulus with given inner and outer radii (relative to  $T^{1/\alpha} d$ ) is Poisson-distributed with parameter proportional to  $\lambda \pi (T^{1/\alpha} d)^2$ , whereby a higher value implies a lower coefficient of variation for this number. Therefore, the number of nodes located within any distance interval has the smaller relative variance the greater the  $\lambda \pi (T^{1/\alpha} d)^2$ , resulting in a smaller variance for  $\Pi(\{X_k\})$ .

### 3.3 Summary and Conclusions

In this chapter, we studied the properties of networks modeled using the Physical model, which also takes into account the interference from simultaneous transmissions in the network.

We began by addressing the graph connectivity of finite networks under this model, which has seen few or no prior contributions. To this end, we generalized the notion of the threshold range for connectivity to the Physical model. In this case, connectivity also depends on medium access control through the time-varying interference: we focused on two previously-studied scenarios, a CDMA network and a network employing slotted Aloha.

In the former setting, we made the worst-case assumption that all nodes transmit constantly and thereby ignored the time dependence. In this case, the connectivity condition of each link is a linear constraint on the two free parameters, and the threshold range for connectivity generalizes into a boundary in this parameter space, consisting of connectivity constraint segments of individual links that are critical for connectivity at each point. We showed how to trace this boundary using a simple rule.

With the slotted-Aloha network, we defined a third free parameter, link confidence, which is the required minimum probability of successful transmission. The link connectivity constraint is no longer linear with respect to link confidence, but the connectivity boundary can be determined with any fixed link confidence according to the same principle as with the CDMA network.

Overall, the connectivity boundary can be seen to imply tradeoffs between different performance quantities in the network. Our results make it possible to extend the approach of studying connectivity of wireless multihop networks by simulation to interference-dominated scenarios such as



networks under a high traffic load, as well as studying the sensitivity of connectivity to different network parameters.

In the remainder of this chapter, we studied the temporal probability of successful transmission (or, more precisely, reception) in an infinite random slotted-Aloha network employing constant transmission power: this medium access scheme has recently been shown to achieve the optimal throughput scaling under the Physical model, but the existing quantitative results were based on exponentially distributed and hence unbounded transmission powers [BBM04].

We first focused on the probability of successful reception averaged over all configurations of other nodes surrounding the receiver. This probability only depends on the distribution of a certain Poisson shot noise. By deriving the distribution of interference power from the proximity of the receiver exactly, the distribution of the remaining interference can be approximated using its Laplace transform, making the evaluation of the overall averaged success probability possible.

We also addressed the distribution of the temporal probability of successful reception over different configurations of surrounding nodes and hence over different recipients. Dividing the neighborhood of the receiver into zones where each transmitter is assumed to produce equally strong interference and taking the effect of different configurations outside this neighborhood as an average, we obtained a recursion for evaluating the tail probability of this distribution.

Thus, both results are numerical approximations that can be improved to an arbitrary level of accuracy, at the cost of added numerical computations; we validated these approximations with the aid of simulations. As a potential direction for future work, it might be interesting to utilize these results in studying how the assumption of a common transmission power changes the values of the medium access probability that maximize the mean density of progress, determined for the exponentially distributed powers in [BBM04].



## 4 SUMMARIES OF PUBLICATIONS AND AUTHOR'S CONTRIBUTIONS

Publication [1] presents algorithms for determining the threshold ranges for  $k$ -connectivity for a given set of nodes under the assumption of the Boolean model with a constant transmission range. In addition, purely empirical models fitted to simulation data obtained by using these algorithms are presented. We conclude the publication by pointing out that these models are not consistent with existing asymptotic results.

Publication [1] is the sole work of the present author.

Publication [2] augments Publication [1] by taking the known asymptotic distributions of the threshold ranges as the bases of the empirical models. The models thus describe the convergence of the distributions to the known asymptotic ones. In addition, independently of the exact asymptotic distributions for  $k$ -connectivity with  $k > 1$  derived at the time of writing, we derive approximations for these distributions. Finally, we demonstrate the ability of these models to predict these distributions for finite networks.

Publication [2] is the sole work of the present author.

Publication [3] applies the recently derived exact expression for the stationary spatial distribution of nodes in the Random Waypoint mobility model to predicting the connectivity of networks of  $n$  nodes moving according to this model. Motivated by asymptotic properties, the distribution of the threshold range for  $k$ -connectivity is approximated by that of the maximum of  $n$  i.i.d.  $k$ -nearest-neighbor distances. In addition, approximations for the mean durations of connectivity periods are presented, also based on recently derived properties of the RWP model.

The present author pointed out the way to utilize the stationary distribution to approximate the probability of connectivity and contributed the analytical motivation for this approximation to the paper. In addition, the efficient way of validating the approximation, by comparison against the empirical cumulative distribution of the threshold range for  $k$ -connectivity, was suggested by the present author.

Publication [4] focuses on the coverage of random networks in a bounded domain when using a Boolean coverage disk model. The covered fraction of the target domain is defined as the random variable of interest, whose expectation is determined in a simple circular domain. It is pointed out that the problem of full coverage is analogous to that of connectivity, in that it also reduces to knowing the distribution of a well-defined threshold range that can easily be determined. Existing asymptotic results are interpreted as a limit distribution of this range, and an approximation for this distribution in the case dominated by border effects is derived. Finally, the applicability of empirical models in predicting these distributions for finite configurations is demonstrated.

Publication [4] is the sole work of the present author.

Publication [5] presents the problem of making a given network connected by adding as few additional nodes as possible to the network. The connections of the problem to existing NP-hard problems are shown, and increasingly advanced heuristic algorithms are proposed for the problem, together with their complexity analysis. Finally, the performance of the algorithms is compared by simulation.

The problem was originally presented by Dr. Karvo. The contributions of the present author include pointing out the connection with the Euclidean minimum Steiner tree problem, the idea of utilizing the Voronoi diagram, the exhaustive optimizations used in the most advanced algorithm, and the complexity analysis of all the algorithms. The present author is also the first author of this publication, having written most of the paper.

The optimization presented in the Appendix was omitted from the publication because of space restrictions, although it was included in the submission evaluated in the review process.

Publication [6] presents a generalization of the notion of threshold range for connectivity to the Physical model. Because connectivity is now also affected by medium access control through the time-varying interference, two scenarios from existing studies are considered. In contrast to the Boolean model, there is more than one free parameter, and the threshold range generalizes to a boundary in the space of these parameters that implies tradeoffs between different performance quantities; we show how to determine this boundary for a given network.

Publication [6] is the sole work of the present author.

Publication [7] extends the analysis of random slotted-Aloha networks by assuming that all nodes in the network transmit with some common constant power. We evaluate the probability of successful transmission in a random time slot. As a function of the random node locations, this temporal probability is a random variable with its own distribution. We develop numerical approximations for evaluating both the mean and the tail probability of this distribution. The accuracy of our approximations can be improved indefinitely, at the cost of added numerical computations.

In studying the success probability averaged over different configurations, the present author noticed that the constant transmission power can be translated into zones with maximum numbers of active transmitters, whereas Prof. Virtamo contributed the idea that the interference power from outside these zones should be approximately normally distributed, making it possible to apply the Bahadur-Rao approximation from large-deviations theory to evaluate this distribution. Following Virtamo's observation that the temporal probability is a random variable with its own distribution, the present author derived the recursion for evaluating the tail probability of this distribution. The present author is the first author of this publication, having composed the paper.

## APPENDIX A SUPPLEMENTARY MATERIAL FOR PUBLICATIONS

### Optimal placement of two relay nodes to connect three terminal nodes

For completeness, we present here the optimization of the placement of two relay nodes to connect three terminal nodes, utilized in the Greedy Triangle algorithm in Publication [5] but omitted from the final publication due to space restrictions.

By the optimal placement we mean that the transmission range required from the relay nodes is minimized. Recall that the case of one relay node is simple: if the point equally distant from the three terminal nodes falls inside the triangle spanned by the terminal nodes, then that is the optimal place for the relay node, else the midpoint of the longest side of the triangle is.

Then consider the problem of connecting three terminal nodes at given locations with the jointly optimal placement of two relay nodes. Let us name the locations of the three terminal nodes as points  $A$ ,  $B$ , and  $C$ , forming the vertices of a triangle  $ABC$ . Without loss of generality, we assume that  $|AB| \leq |CA| \leq |CB|$ . We will choose Euclidean coordinates in  $\mathbb{R}^2$  so that point  $C$  is chosen as the origin and the first dimension is in the direction  $\overrightarrow{CA}$ , so that the position vector  $\bar{r}_A = (a \ 0)$ . Let  $\bar{r}_B$  denote the position vector of terminal node  $B$ . Given the assumption  $|AB| \leq |CA| \leq |CB|$ , the triangle can always be flipped and rotated so that point  $B$  is located inside the bounded set depicted in Figure 4.1(a). Also, note that any triplet of points in  $\mathbb{R}^d$  is located on a  $\mathbb{R}^2$  plane, and thus this algorithm generalizes easily to  $\mathbb{R}^d$ .

The optimal solution for the locations of the two relay nodes is either

- A** to use the relay nodes to split in half the two shortest sides ( $CA$ ,  $AB$ ) of the triangle  $ABC$  (as in the final two example figures in Table 4.1 at the end of this section), or
- B** first, to connect two terminal nodes ( $A$  and  $B$ ) with one relay node and then place the second relay node midway between the first relay node and the remaining terminal node  $C$  to be connected (the other example figures in Table 4.1).

It can be deduced that whenever case **B** is optimal, the first relay node must connect the two terminal nodes closest to each other. This is because by the possibility of case **A**, the least required transmission range cannot exceed half the second-shortest distance between two of the three terminal nodes.

For now, assume that case **B** above is optimal (we will derive the conditions for this later). Under this assumption, our task is to optimize the location of the first relay node, point  $P$ , so as to minimize the required range  $\max\{f_1(P) = |CP|/2, f_2(P) = |AP|, f_3(P) = |BP|\}$ . For an explanation of  $f_1(P)$ , recall that in case **B**, the second relay node is placed in the middle of edge  $CP$ . We know that the optimal  $P$  must lie inside the

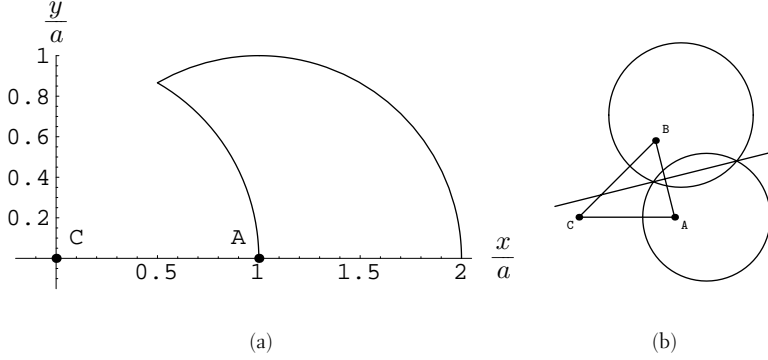


Figure 4.1: The set of possible locations of terminal node  $B$  (a); example triangle (b)

triangle  $ABC$ , for otherwise it would be possible to decrease the distance to all the points  $A$ ,  $B$ , and  $C$  by moving  $P$ . Furthermore, in the optimum we must have  $f_i(P) = f_j(P) \geq f_k(P)$ , for some  $i \neq j \neq k; i, j, k \in \{1, 2, 3\}$ , otherwise  $P$  could again be improved. Here, strict inequality applies if the two equal functions have attained their least possible common value.

The equation  $f_1(P) = |CP|/2 = f_2(P) = |AP|$  is satisfied by points  $P$  located on the circle with radius  $\frac{2}{3}|CA|$ , centered at  $\frac{4}{3}\bar{r}_A$ . Another circle is defined accordingly by the equation  $f_1(P) = f_3(P)$ . The solutions of the equation  $f_1(P) = f_2(P) = f_3(P)$  are hence the intersections of these two circles, which are on the line  $|AP| = |BP|$ , as shown by Figure 4.1(b). It is easy to show that under the assumption  $|AB| \leq |CA| \leq |CB|$ , these intersections always exist. In light of the above, the optimal  $P$  is located at such an intersection if it is inside the triangle, otherwise it is at the intersection of the line or circle  $f_i(P) = f_j(P)$  and the side of the triangle where the value  $f_i(P) = f_j(P)$  is smallest.

We know that one of the intersections  $f_1(P) = f_2(P) = f_3(P)$  always falls outside the triangle  $ABC$ . In order that the other intersection not fall outside the side  $AB$ , the midpoint of  $AB$  must lie inside the two circles. In fact, it suffices to write this condition for one circle only, since it implies the other, so we get

$$\left| \frac{\bar{r}_A + \bar{r}_B}{2} - \frac{4}{3}\bar{r}_A \right| < \frac{2}{3}|CA| \Leftrightarrow \left| \bar{r}_B - \frac{5}{3}\bar{r}_A \right| < \frac{4}{3}|CA|, \quad (4.1)$$

i.e. point  $B$  must lie inside the circle with radius  $\frac{4}{3}|CA|$ , centered at  $\frac{5}{3}\bar{r}_A$ . If this condition is not satisfied, the optimal  $P$  — given the assumption that case **B** really is optimal — is located midway between  $A$  and  $B$ , at  $(\bar{r}_A + \bar{r}_B)/2$ . On the other hand, the condition for the other intersection not falling outside the side  $CB$  is

$$\left| \frac{2}{3}\bar{r}_B - \frac{4}{3}\bar{r}_A \right| > \frac{2}{3}|CA| \Leftrightarrow |\bar{r}_B - 2\bar{r}_A| > |CA|, \quad (4.2)$$

i.e. point  $B$  must lie outside the circle with radius  $|CA|$ , centered at  $2\bar{r}_A$ . If

this condition is not satisfied, the optimal  $P$  — given the assumption that case **B** really is optimal — is located on the segment  $CB$  at  $\frac{2}{3}\bar{r}_B$ . (Under the assumption  $|AB| \leq |CA| \leq |CB|$ , the other intersection cannot fall outside the side  $CA$ .)

In general, case **B** is optimal if the optimal  $P$  presented above satisfies  $|CP| < |CA|$ . It is easy to check that whenever the optimal  $P$  is at  $(\bar{r}_A + \bar{r}_B)/2$ , this condition is always satisfied. When it is at  $\frac{2}{3}\bar{r}_B$  (i.e. when  $|\bar{r}_B - 2\bar{r}_A| < |CA|$ ), the condition becomes

$$\frac{2}{3}|CB| < |CA| \Leftrightarrow |CB| < \frac{3}{2}|CA|. \quad (4.3)$$

Finally, let us derive the condition  $|CP| < |CA|$  for the intersection  $f_1(P) = f_2(P) = f_3(P)$  falling inside the triangle  $ABC$ . Let  $\bar{r}_B = (x \ y)$ . A general point  $P$  on the line  $|AP| = |BP|$  is then at  $\bar{r}_P = ((a \ 0) + (x \ y))/2 + t \cdot (y - 0 \ - (x - a))$ . The value of the scalar  $t$  corresponding to the intersection  $f_1(P) = f_2(P) = f_3(P)$  falling inside the triangle  $ABC$  can be solved to be

$$t = \frac{\{4ay - [16a^2y^2 - (3a^2 - 10ax + 3x^2 + 3y^2) \times (12a^2 - 24ax + 12x^2 + 12y^2)]^{1/2}\}}{(12a^2 - 24ax + 12x^2 + 12y^2)}.$$

On the other hand, the value of  $t$  that makes  $|CP| = |CA|$  and results in the greater  $x$ -coordinate for  $P$ , is

$$t = \frac{-2ay + (3a^4 - 8a^3x + 6a^2x^2 - x^4 + 6a^2y^2 - 2x^2y^2 - y^4)^{1/2}}{2(a^2 - 2ax + x^2 + y^2)}.$$

We get the boundary for the condition of interest by setting these two values equal. The resulting equation is satisfied on the circles with radius  $|CA|/2$ , centered at  $(\frac{7}{8}a \pm \frac{\sqrt{15}}{8}a)$ . Only the upper one of these creates a boundary in the domain of interest. Simple experimenting shows that given the conditions (4.1) and (4.2), case **B** is optimal if

$$\left| \bar{r}_B - \left( \frac{7}{8}a \quad \frac{\sqrt{15}}{8}a \right) \right| < |CA|/2. \quad (4.4)$$

The labels of the conditions (4.1) to (4.4) have been placed near their respective boundaries in Figure 4.2, so that each label is on the side of the boundary where the condition is satisfied. The resulting five subsets of the possible locations of terminal node  $B$  have been labelled using Roman numerals, and the optimal solution in each subset is summarized in Table 4.1.

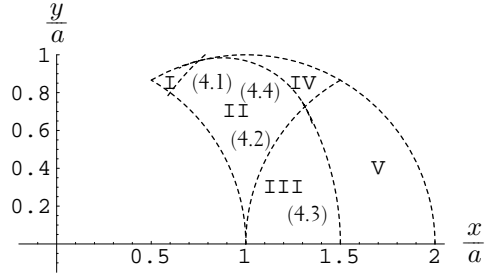
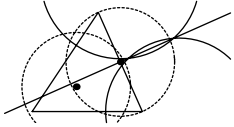
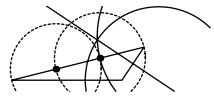
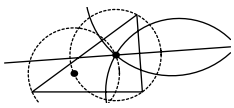
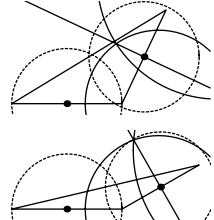


Figure 4.2: The division of the possible locations of terminal node  $B$  according to the optimal placement of two relay nodes

Table 4.1: Optimal placement of relay nodes according to the locations of terminal node  $B$  in Figure 4.2

Subset	Solution and example	Subset	Solution and example
I	Place first node $P$ at $\bar{r}_P = (\bar{r}_A + \bar{r}_B)/2$ , place second node midway between $C$ and $P$ 	III	Place first node $P$ at $\bar{r}_P = \frac{2}{3}\bar{r}_B$ , place second node midway between $C$ and $P$ 
II	Place first node $P$ in the intersection $f_1(P) = f_2(P) = f_3(P)$ inside the triangle, place second node midway between $C$ and $P$ 	IV,V	Place one node midway between $C$ and $A$ and the other node midway between $A$ and $B$ 







## REFERENCES

- [3GP05] 3rd Generation Partnership Project Technical Specification 04.60: General Packet Radio Service; Mobile Station - Base Station System interface; Radio Link Control / Medium Access Control protocol, September 2005. <http://www.3gpp.org>.
- [Abr70] N. Abramson. The ALOHA system – another alternative for computer communications. In *Proceedings of the AFIPS Fall Joint Computer Conference*, volume 37, pages 281–285. AFIPS Press, 1970.
- [Abr73] N. Abramson. Packet switching with satellites. In *Proceedings of the AFIPS National Computer Conference*, volume 42, pages 695–702. AFIPS Press, 1973.
- [AK04] Ashish Agarwal and P. R. Kumar. Capacity bounds for ad hoc and hybrid wireless networks. *SIGCOMM Comput. Commun. Rev.*, 34(3):71–81, 2004.
- [ASSC02] Ian F. Akyildiz, Weilian Su, Yogesh Sankarasubramaniam, and Erdal Cayirci. A survey on sensor networks. *IEEE Communications Magazine*, 40(8):102–114, August 2002.
- [Aur91] Franz Aurenhammer. Voronoi diagrams – a survey of a fundamental geometric data structure. *ACM Comput. Surv.*, 23(3):345–405, 1991.
- [BBM04] François Baccelli, Bartłomiej Błaszczyszyn, and Paul Mühlethaler. An Aloha protocol for multihop mobile wireless networks. In *Proceedings of 16th ITC Specialist Seminar*, pages 28–39, August 2004.
- [BBM06] F. Baccelli, B. Błaszczyszyn, and P. Mühlethaler. An Aloha protocol for multihop mobile wireless networks. *IEEE Transactions on Information Theory*, 52(2):421–436, February 2006.
- [Bet02] C. Bettstetter. On the minimum node degree and connectivity of a wireless multihop network. In *Proc. 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'02)*, pages 80–91, June 2002.
- [Bet04] C. Bettstetter. On the connectivity of ad hoc networks. *The Computer Journal, Special Issue on Mobile and Pervasive Computing*, 47(4):432–447, 2004.
- [BHM05] Christian Bettstetter, Christian Hartmann, and Clemens Moser. How does randomized beamforming improve the connectivity of ad hoc networks? In *Proceedings of IEEE International Conference on Communications (ICC)*, volume 5, pages 3380–3385, May 2005.

- [Bol85] Béla Bollobás. *Random Graphs*. Academic Press, 1985.
- [BR60] R. R. Bahadur and R. Ranga Rao. On deviations of the sample mean. *The Annals of Mathematical Statistics*, 31(4):1015–1027, December 1960.
- [BR04] P. Basu and J. Redi. Movement control algorithms for realization of fault-tolerant ad hoc robot networks. *IEEE Network*, 18(4):36–44, July 2004.
- [BRS03] C. Bettstetter, G. Resta, and P. Santi. The node distribution of the Random Waypoint mobility model for wireless ad hoc networks. *IEEE Transactions on Mobile Computing*, 2(1):25–39, 2003.
- [BZ02] C. Bettstetter and J. Zangl. How to achieve a connected ad hoc network with homogeneous range assignment: an analytical study with consideration of border effects. In *Proc. 4th IEEE International Conference on Mobile and Wireless Communication Networks (MWCN'02)*, pages 125–129, September 2002.
- [DBT03] O. Dousse, F. Baccelli, and P. Thiran. Impact of interferences on connectivity in ad hoc networks. In *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, volume 3, pages 1724–1733, April 2003.
- [DBT05] O. Dousse, F. Baccelli, and P. Thiran. Impact of Interferences on Connectivity in Ad Hoc Networks. *IEEE/ACM Trans. on Networking*, 13(2):425–436, 2005.
- [DFM<sup>+</sup>06] Olivier Dousse, Massimo Franceschetti, Nicolas Macris, Ronald Meester, and Patrick Thiran. Percolation in the signal to interference ratio graph. *Journal of Applied Probability*, 43(2), 2006.
- [DH89] H. Dette and N. Henze. The limit distribution of the largest nearest-neighbor link in the unit  $d$ -cube. *Journal of Applied Probability*, 26:67–80, 1989.
- [DH90a] H. Dette and N. Henze. Some peculiar boundary phenomena for extremes of  $r$ th nearest neighbor links. *Statistics & Probability Letters*, 10:381–390, 1990.
- [DH90b] D.-Z. Du and F.K. Hwang. An approach for proving lower bounds: solution of Gilbert-Pollak’s conjecture on Steiner ratio. In *Proceedings of 31st Annual Symposium on Foundations of Computer Science*, volume 1, pages 76–85, October 1990.
- [DRL03] R. D’Souza, S. Ramanathan, and D. Lang. Measuring performance of ad hoc networks using timescales for information flow. In *Proceedings of the 22nd Annual Joint Conference of*

the *IEEE Computer and Communications Societies (INFOCOM)*, April 2003.

- [DT96] G. Di Battista and R. Tamassia. On-line maintenance of triconnected components with SPQR-trees. *Algorithmica*, 15:302–318, 1996.
- [DT04] O. Dousse and P. Thiran. Connectivity vs capacity in dense ad hoc networks. In *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, volume 1, March 2004.
- [DTH02] O. Dousse, P. Thiran, and M. Hasler. Connectivity in ad-hoc and hybrid networks. In *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, volume 2, pages 1079–1088, 2002.
- [ER60] Pál Erdős and Alfréd Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.
- [Far05] András Faragó. New analytical results on ad hoc network connectivity. In *Proceedings of the Third IASTED International Conference on Communications and Computer Networks (CCN)*, pages 126–131, October 2005.
- [GGJ77] M. R. Garey, R. L. Graham, and D. S. Johnson. The complexity of computing Steiner minimal trees. *SIAM J. Appl. Math.*, 32(4):835–859, June 1977.
- [Gil61] E. N. Gilbert. Random plane networks. *SIAM J.*, 9:533–543, 1961.
- [GK98] P. Gupta and P. R. Kumar. Critical power for asymptotic connectivity in wireless networks. *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*, pages 547–566, 1998.
- [GK00] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000.
- [GM00] C. Gutwenger and P. Mutzel. A linear time implementation of SPQR-trees. In *Proceedings of 8th International Symposium on Graph Drawing (GD 2000)*, pages 77–90, September 2000.
- [Gup00] P. Gupta. *Design and Performance Analysis of Wireless Network*. PhD thesis, University of Illinois, Urbana-Champaign, 2000.
- [Hal88] P. Hall. *Introduction to the theory of coverage processes*. John Wiley & Sons, New York, 1988.

- [HLV06] Esa Hyttiä, Pasi Lassila, and Jorma Virtamo. Spatial node distribution of the Random Waypoint mobility model with applications. *IEEE Transactions on Mobile Computing*, 2006. To appear.
- [HM04] R. Hekmat and P. Van Mieghem. Study of connectivity in wireless ad hoc network with an improved radio model. In *Proceedings of the 2nd Workshop on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt'04)*, March 2004.
- [Hsu93] Tsan-sheng Hsu. *Graph Augmentation and Related Problems: Theory and Practice*. PhD thesis, University of Texas at Austin, 1993.
- [HT03] Chi-Fu Huang and Yu-Chee Tseng. The coverage problem in a wireless sensor network. In *WSNA '03: Proceedings of the 2nd ACM international conference on Wireless sensor networks and applications*, pages 115–121, New York, NY, USA, 2003. ACM Press.
- [HV05] E. Hyttiä and J. Virtamo. Random waypoint model in  $n$ -dimensional space. *Operations Research Letters*, pages 567–571, 2005.
- [IEE99] IEEE 802.11 standard: Wireless LAN Medium Access Control and Physical Layer Specifications , 1999. <http://standards.ieee.org/>.
- [Jan86] S. Janson. Random coverings in several dimensions. *Acta Mathematica*, 156:83–118, 1986.
- [JM96] David B Johnson and David A Maltz. Dynamic source routing in ad hoc wireless networks. In Imielinski and Korth, editors, *Mobile Computing*, volume 353. Kluwer Academic Publishers, 1996.
- [KKP03] George Kesidis, Takis Konstantopoulos, and Shashi Phoha. Surveillance coverage of sensor networks under a random mobility strategy. In *Proceedings of IEEE Sensors Conference*, volume 2, pages 961–965, October 2003.
- [KLB04] Santosh Kumar, Ten H. Lai, and József Balogh. On  $k$ -coverage in a mostly sleeping sensor network. In *MobiCom '04: Proceedings of the 10th annual international conference on Mobile computing and networking*, pages 144–158, New York, NY, USA, 2004. ACM Press.
- [Kru56] Joseph Kruskal. On the shortest spanning subtree of a graph and the traveling salesman problem. In *Proceedings of the American Mathematical Society*, volume 7, pages 48–50, 1956.

- [KT75] L. Kleinrock and F. Tobagi. Packet switching in radio channels: Part I – Carrier Sense Multiple-Access modes and their throughput-delay characteristics. *IEEE Transactions on Communications*, 23(12):1400–1416, December 1975.
- [LBD<sup>+</sup>05] Benyuan Liu, Peter Brass, Olivier Dousse, Philippe Nain, and Don Towsley. Mobility improves coverage of sensor networks. In *MobiHoc '05: Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing*, pages 300–308, New York, NY, USA, 2005. ACM Press.
- [Le 05] Jean-Yves Le Boudec. On the stationary distribution of speed and location of random waypoint. *IEEE Transactions on Mobile Computing*, 4(4):404–405, 2005.
- [Lin22] Jarl Waldemar Lindeberg. Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 15:211–225, 1922.
- [LT03] B. Liu and D. Towsley. On the coverage and detectability of large-scale wireless sensor networks. In *Proceedings of the Workshop on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt'03)*, pages 201–203, March 2003.
- [LV05] Jean-Yves Le Boudec and Milan Vojnović. Perfect simulation and stationarity of a class of mobility models. In *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, volume 4, pages 2743–2754, March 2005.
- [LWF02] Xiang-Yang Li, Peng-Jun Wan, and O. Frieder. Coverage in wireless ad-hoc sensor networks. In *Proceedings of IEEE International Conference on Communications (ICC)*, volume 5, pages 3174–3178, 2002.
- [MAN05] MANET (mobile ad-hoc networks) working group, IETF Secretariat, October 2005. <http://www.ietf.org/html.charters/manet-charter.html>.
- [MKPS01] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M.B. Srivastava. Coverage problems in wireless ad-hoc sensor networks. In *Proceedings of IEEE Infocom*, pages 1380–1387, April 2001.
- [MR96] R. Meester and R. Roy. *Continuum percolation*. Cambridge University Press, Cambridge, 1996.
- [NC04] W. Navidi and T. Camp. Stationary distributions for the Random Waypoint mobility model. *IEEE Transactions on Mobile Computing*, 3(1):99–108, January-March 2004.

- [NMN01] Neseřil, Milkova, and Neseřilova. Otakar Boruvka on minimum spanning tree problem: translation of both the 1926 papers, comments, history. *DMATH: Discrete Mathematics*, 233, 2001.
- [Pen97] Mathew D. Penrose. The longest edge of the random minimal spanning tree. *Annals of Applied Probability*, 7(2):340–361, 1997.
- [Pen98] M. D. Penrose. Extremes for the minimal spanning tree on normally distributed points. *Advances in Applied Probability*, 30(3):628–639, 1998.
- [Pen99] M. D. Penrose. On  $k$ -connectivity for a geometric random graph. *Random Structures and Algorithms*, 15(2):145–164, 1999.
- [PPT89] T. K. Philips, S. S. Panwar, and A. N. Tantawi. Connectivity properties of a packet radio network model. *IEEE Transactions on Information Theory*, 35(5):1044–1047, September 1989.
- [Sha49] Claude E. Shannon. Communication in the presence of noise. In *Proceedings of the Institute of Radio Engineers (IRE)*, volume 37, pages 10–21, January 1949.
- [SMH99] M. Sánchez, P. Manzoni, and Z. J. Haas. Determination of critical transmission range in Ad-Hoc Networks. In *Proceedings of Multiaccess Mobility and Teletraffic for Wireless Communications 1999 Workshop (MMT'99)*, October 1999.
- [TG02] Di Tian and Nicolas D. Georganas. A coverage-preserving node scheduling scheme for large wireless sensor networks. In *WSNA '02: Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, pages 32–41, New York, NY, USA, 2002. ACM Press.
- [WCP04] Guiling Wang, Guohong Cao, and Tom La Porta. Movement-assisted sensor deployment. In *Proceedings of Twenty-Third Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, volume 4, pages 2469–2479, March 2004.
- [WY04] Peng-Jun Wan and Chih-Wei 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*, pages 1–8. ACM Press, 2004.
- [WYAd05] Steven P. Weber, Xiangying Yang, Jeffrey G. Andrews, and Gustavo de Veciana. Transmission capacity of wireless ad hoc networks with outage constraints. *IEEE Transactions on Information Theory*, 51(12):4091–4102, December 2005.



- [WYdA04] S. P. Weber, X. Yang, G. de Veciana, and J. G. Andrews. Transmission capacity of CDMA ad-hoc networks. In *Proceedings of Eighth IEEE International Symposium on Spread Spectrum Techniques and Applications (ISSSTA)*, pages 245–249, September 2004.
- [XK04a] Liang-Liang Xie and P. R. Kumar. A network information theory for wireless communication: scaling laws and optimal operation. *IEEE Transactions on Information Theory*, 50(5):748–767, May 2004.
- [XK04b] Feng Xue and P. R. Kumar. The number of neighbors needed for connectivity of wireless networks. *Wireless Networks*, 10(2):169–181, March 2004.
- [XXK05] Feng Xue, Liang-Liang Xie, and P. R. Kumar. The transport capacity of wireless networks over fading channels. *IEEE Transactions on Information Theory*, 51(3):834–847, March 2005.
- [ZH04] Honghai Zhang and Jennifer Hou. On deriving the upper bound of  $\alpha$ -lifetime for large sensor networks. In *MobiHoc '04: Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, pages 121–132, New York, NY, USA, 2004. ACM Press.
- [ZH05] Honghai Zhang and Jennifer C. Hou. Maintaining sensing coverage and connectivity in large sensor networks. *Ad Hoc & Sensor Wireless Networks*, 1(1-2):89–124, March 2005.