TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Kommunikationsnetze

# Wireless Multi-Hop Networks with Beamforming Antennas and Multi-User Detection

Robert Vilzmann

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

| | | |
|---|---|---|
| Vorsitzender: | | Univ.-Prof. Dr.-Ing., Dr.-Ing. habil. Erwin Biebl |
| Prüfer der Dissertation: | 1. | Univ.-Prof. Dr.-Ing. Jörg Eberspächer |
| | 2. | Univ.-Prof. Dr.-Ing. Wolfgang Utschick |

Die Dissertation wurde am 29.09.2008 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 24.03.2009 angenommen.

**Abstract**

Multi-hop networking is considered to be an essential technology for future wireless network architectures. This thesis is concerned with the application of beamforming and multi-user detection in multi-hop networks. At first, centrally optimized, distributed and emergent methods for improving the connectivity using beamforming are proposed, leading to a considerably better connectedness of multi-hop networks. Then, this work contributes by a novel cross-layer approach to MAC layer design. In a concept called MUD-MAC, it facilitates multi-user detection in multi-hop networks, showing significant throughput improvements over a single-user system. Finally, the cross-layer approach is applied to beamforming. The resulting new MAC protocol, called BeamMAC, shows similar throughput improvements. The work provides an analytical and simulation-based evaluation of the different approaches, and discusses aspects of their implementation in practice.

**Zusammenfassung**

Multihop Networking wird in künftigen drahtlosen Netzarchitekturen eine wesentliche Rolle spielen. Diese Arbeit beschäftigt sich mit der Anwendung von Beamforming und Mehrnutzerdetektion in Multihop-Netzen. Zunächst werden zentral optimierte, verteilte und emergente Verfahren zur Verbesserung der Konnektivität mittels Beamforming vorgeschlagen, die zu deutlich besser verbundenen Multihop-Netzen führen. Anschließend liefert die Arbeit Beiträge durch einen neuartigen schichtenübergreifenden Ansatz für den Medienzugriff. In einem MUD-MAC genannten Konzept ermöglicht dieser die Anwendung von Mehrnutzerdetektion in Multihop-Netzen, und zeigt dabei erhebliche Durchsatzverbesserungen im Vergleich zu einem Einnutzersystem. Schließlich wird der schichtenübergreifende Ansatz auf Beamforming angewandt. Das sich daraus ergebende MAC-Protokoll, genannt BeamMAC, zeigt ähnliche Durchsatzverbesserungen. Die Arbeit liefert eine analytische und rechnergestützte Auswertung der verschiedenen Verfahren, und diskutiert Aspekte ihrer praktischen Realisierung.

Robert Vilzmann

# Wireless Multi-Hop Networks with Beamforming Antennas and Multi-User Detection

# Preface

In 2004, Professor Dr. Jörg Eberspächer gave me the opportunity to join the research and teaching staff of the Institute of Communication Networks (LKN) at the TUM. This thesis is the result of my research activities during these years. Thank you for giving me this great opportunity for conducting research, for providing a pleasant and productive working environment, and for advising my dissertation!

I would like to thank Professor Dr. Wolfgang Utschick for serving as second examiner. The seminars and guest lectures which he organized contributed to my academic experience at the TUM. I also thank Professor Dr. Erwin Biebl for chairing the examination committee.

This thesis would not have been possible without my fellow researchers and co-authors at the LKN and at DOCOMO Euro-Labs, in particular Imad Aad, Gerhard Bauch, Christian Bettstetter, Stephan Eichler, Christian Hartmann, Moritz Kiese, Katsutoshi Kusume, Silke Meister, Robert Nagel, Jörg Widmer and Hans-Martin Zimmermann. I always admired their devotion to scientific work.

I thank my fellow assistants at the LKN for fruitful technical discussions and reviewing the manuscript draft. Thanks also to my students. The work of Hauke Holtkamp, Mohammed Kamaruzzaman, Diego Herrero Marín, Daniel Medina, Andreas Müller and Christian Rössert directly contributed to this work. For everything else, thanks to Dr. Martin Maier!

In 2007, Christian Bettstetter invited me for a stay as a guest researcher at his institute in Klagenfurt. Thank you for this great experience!

My time at the institute was not only marked by research. It was also a time for making friends, and I hope that these friendships will last for many years to come.

Finally, I would like to express my deepest gratitude to my parents and family for their endless trust and support.

Munich, September 2008 Robert Vilzmann

# Contents

# 1 Introduction

Wireless radio is virtually the only technology for distant mobile communication. Its commercial success began with Marconi's "Wireless Telegraph and Signal Company" in 1897. At that time, wireless communication was tantamount to point-to-point communication. The biggest upturn in civil wireless communication took place because of the cellular principle. While introduced already in the 1950's, it started to boom with the introduction of fully digital systems in the 1990's, in particular GSM [EVB01b], allowing even for international roaming.

Besides cellular communication, so-called *packet radio networks* have been subject-matter of international research for decades. Advances in microelectronics and new applications have fostered work in this direction. Specific developments lead to new system architectures, such as ad hoc networks, wireless sensor networks, or relay systems for the extension of cellular networks. We can summarize and characterize such developments as *multi-hop* networks. Multi-hop means that, in between the source and the destination of a message, a chunk of data, or multimedia content, two or more continuous segments of the communication path are implemented by wireless technology.

The notion of multi-hop networking leads to a diversification of wireless system architectures. Instead of communicating only with basestations of cellular systems or access points of WLANs, mobile devices get more and more interconnected. Driving forces for multi-hop networks, in part reality already today, are the need for

- establishing a network where otherwise no (suitable) network would be available (e.g. wireless sensor networks for building automation, mesh networks for metropolitan area Internet access, tactical fault-tolerant networks),
- improving the performance of existing networks (e.g. cellular multi-hop for increased data rates in next-generation ("4G") cellular networks), and
- introducing new proximity-based services (e.g. vehicular communication for safety applications, social ad hoc networking).

A classification of different multi-hop architectures is provided in the following chapter.

In the mainpart of this thesis, signal processing techniques are studied and utilized from a networking perspective. Contributions to three major research topics on multi-hop networking are made:

- **Network layer based beamforming:** The absence of basestations raises the problem of disconnected devices in multi-hop networks. On the other hand, if many devices are co-located in a small area, mutual interference affects the reliability of transmissions. Methods and communication protocols are necessary to address both issues. The coupling of beamforming antennas with network-level

information has promise to this end.

- **Multi-User Detection (MUD) in multi-hop networks:** Multi-user detection techniques have been thoroughly investigated for the case of cellular systems. But the same is not the case for multi-hop networks. In particular, new networking mechanisms and protocols are necessary to exploit multi-user detection in multi-hop networks.
- **Medium Access Control (MAC) with beamforming antennas:** Using adaptive beamforming for interference suppression has long been envisaged for cellular systems. In the context of multi-hop networks, this is a quite novel idea. Several MAC protocols for beamforming in ad hoc and multi-hop networks have been proposed, but several open questions have not yet been addressed by the networking community.

The thesis is structured as follows.

Chapter 2 provides an overview of multi-hop networking, and highlights applications and research areas. It emphasizes two main limitations of multi-hop networks, namely connectivity and throughput performance. This chapter also describes different system models that will be used repeatedly in this thesis.

Chapter 3 discusses the coupling of network layer (NET) information with beamforming on the physical layer (PHY). We call this approach *persistent* beamforming. First, we recall different antenna array concepts and beamforming techniques, and review connectivity concepts in graph theory. Based on these fundamentals, a methodology for optimizing a connectivity-related measure by persistent beamforming is developed and applied. Distributed and emergent approaches for practical implementations are the following subject-matter. Their impact is examined in static and mobile networks. Finally, a persistent scheme to interference suppression using adaptive beamforming is proposed. Like the approaches for connectivity improvements, it is based on neighborhood information obtained from the network layer.

Chapter 4 covers the design of MAC protocols for multi-user detection. We begin with a review of interference cancellation, and derive requirements influencing MAC layer design. Based on these requirements, a novel MAC framework called *MUD-MAC* is designed. For a simulation-based study of MUD-MAC, means to reduce the computational complexity of simulations are discussed. Then, the chapter provides network simulations of the new protocol, evaluating the throughput as performance measure. The final focus of chapter 4 is on receiver heterogeneity. We analyze how receivers with different complexity can interact, and study effects on network and per-node throughput.

Chapter 5 provides a thorough survey of existing MAC proposals for beamforming antennas. Critical assumptions in connection with these protocols are highlighted, and requirements for MAC protocol design for beamforming are identified. Then, chapter 5 applies paradigms of MUD-MAC to beamforming, resulting in a novel MAC protocol called *BeamMAC*. As opposed to the persistent schemes of chapter 3, it is intended to allow for beamforming on a per-packet basis. The protocol description is followed by a simulative analysis in different network scenarios. Finally, chapter 5 sheds some light

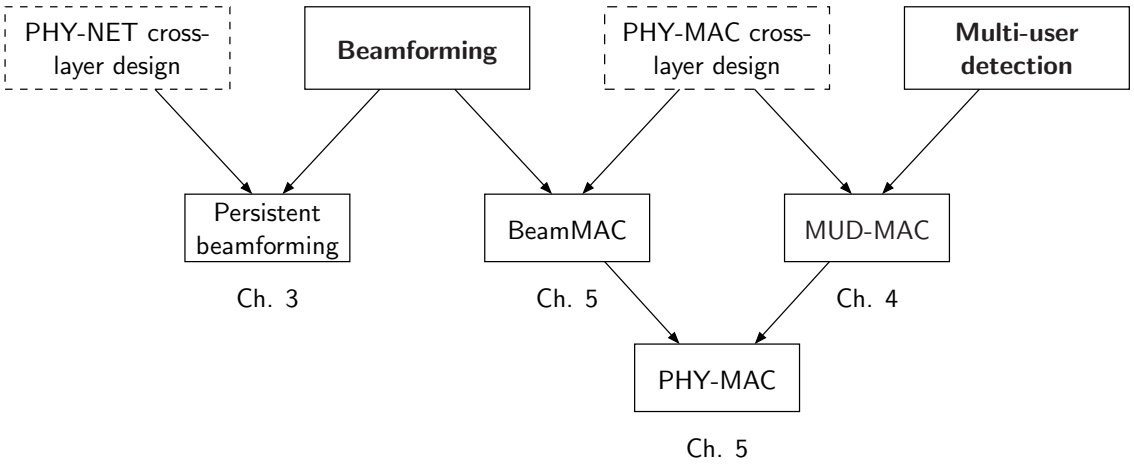



Figure 1.1: Structure of this thesis.

on the integration of beamforming and multi-user detection by discussing MAC layer issues of an integrated approach called *PHY-MAC*.

Chapter 6 concludes this thesis by summarizing its contributions and potential further research.

Appendix A illustrates sample antenna patterns obtained from methods used in this work. Appendix B details the linear formulation of an optimization problem in connection with chapter 3. Appendix C describes the simulation software that was developed in the course of this work. It is an integrated tool that was used for nearly all the numerical work of this thesis, and is vigorously used in further research projects. Appendix D summarizes symbols, notations and abbreviations used in the text.

The materials of this thesis have been published in part before, as indicated in subsequent chapters. These publications contain research results that have been achieved together with the respective co-authors.

# 2 Communication in Multi-Hop Networks

## 2.1 Multi-Hop Network Architectures

Present and future multi-hop networks come in different flavors and architectures. An attempt to a classification was made in [EEH+07], where the following multi-hop concepts, illustrated in Fig. 2.1, are identified:

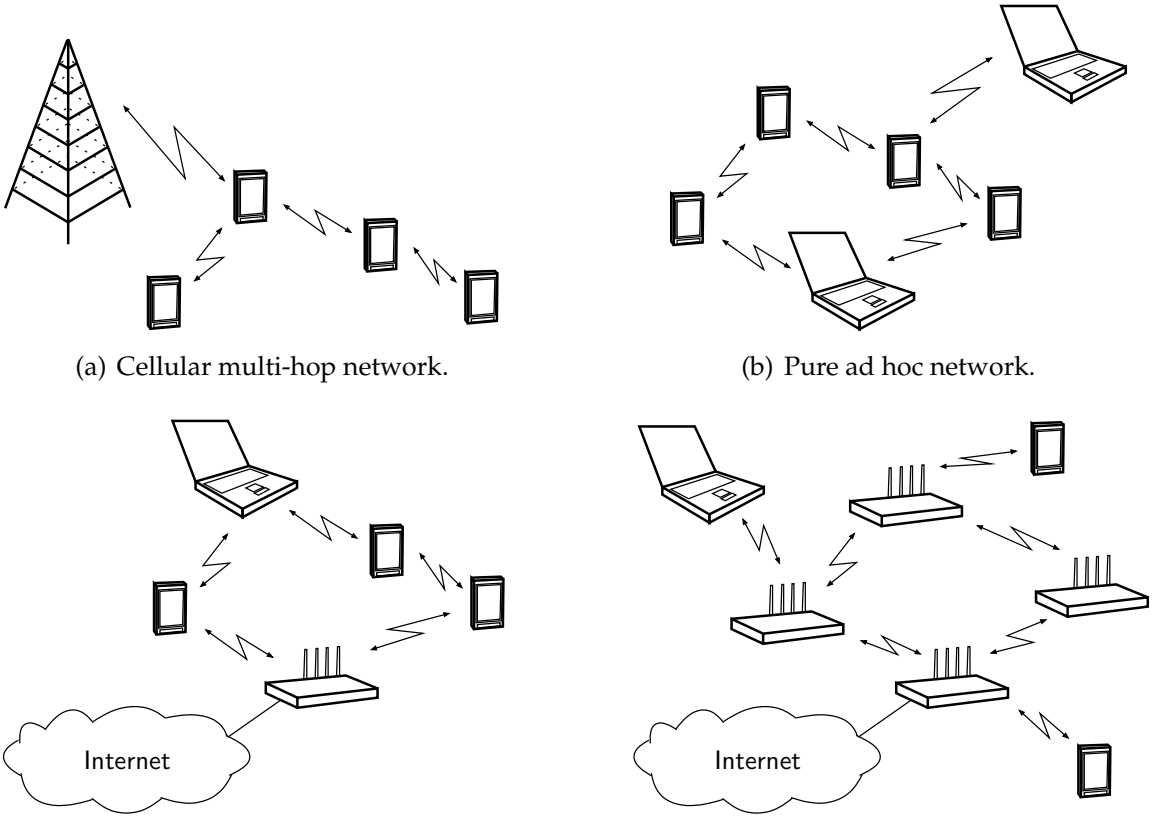
In **cellular multi-hop networks**, a mobile device (typically a mobile phone) is not necessarily communicating directly with the basestation. Instead, it can use another device in the cell as relay. Initially, such an architecture was supposed to extend the range of the basestation, thereby improving network coverage. While full coverage has been achieved – at least in most inhabited parts of developed countries – even without multi-hop extensions, it is still an interesting option for future 4G systems. Due to the high frequencies and high data rate demands, it is envisaged as means to achieve sufficient Quality of Service (QoS) also at the cell edge, thereby improving the "coverage-by-data-rate". Although devices communicate directly with each other, the network operator is still in full control of the system.

In **pure ad hoc networks**, devices interconnect without any supporting infrastructure. The devices, typically all mobile, have to establish a network in a completely self-organizing manner. The term *ad hoc* often implies point-to-point communication, for instance when looking at todays WLAN card configurations or consumer market solutions for home multimedia. But pure ad hoc networks are classical multi-hop networks. While not yet deployed widely, pure ad hoc networks find some application already today, e.g. in sensor networks or military systems. Another major future application is vehicular communication.

**Ad hoc access networks** are connected to the Internet. Today, there is still a coverage issue with low-cost broadband Internet access. Ad hoc access networks could be used to circumvent this issue. New problems such as gateway discovery arise in these networks, and have received ample attention by the research community. Sensor networks and car-to-car networks may be connected to the Internet or some other sort of backbone. In this case they also fall in this multi-hop category.

**Mesh networks** are characterized by an immobile backbone of access points. This backbone may be set up in a planned or unplanned manner. The use case of mesh networks is similar to ad hoc access networks, but mesh networks are expected to provide higher reliability and better QoS. This comes along with the need for maintaining the backbone, and installing the access points including power supply.

The findings of this thesis is not restricted to a particular architecture. It attends to distributed self-organizing methods and algorithms that may be applied in any type of




(a) Cellular multi-hop network.

(b) Pure ad hoc network.

(c) Ad hoc access network.

(d) Mesh network.

Figure 2.1: Different classes of multi-hop networks [EEH+07].

multi-hop network. With cellular multi-hop networks, however, there is no inevitable need for network self-organization. This is because the basestation can control the usage of operator-owned radio resources in a centralized manner. Also, the number of hops is expected to be very low.

The concepts of this work are well applicable to ad hoc and mesh networks. The concepts are fully decentralized, and have particular merit in larger networks. Some approaches work best in networks with low to moderate mobility, and may therefore be of special interest in mesh networks.

## 2.2 Research Challenges

The different multi-hop architectures raise specific research challenges. With cellular multi-hop networks, the efficient use of operator-owned radio resources is critical. When multi-hop networking relies on fixed relays between the user equipment and the basestation, the (possibly adaptive) assignment of relay nodes to basestations is of particular importance. In pure ad hoc networks, major research issues are efficient

spectrum sharing, scalable routing, energy efficiency, and connectivity aspects. Questions specific to ad hoc access networks are with respect to gateway organization, and gateway discovery. In mesh networks, the provisioning of quality of service and security aspects are important topics.

> The most severe limitations of multi-hop networks are with respect to *connectivity* and *throughput*.

In single-hop networks, the question of connectivity is simply a question of the network coverage. A mobile device can send and receive data as soon as it is within communication range of a basestation or access point. In multi-hop networks, this question cannot be simply answered by the notion of coverage. Instead, the probability of the connectedness of two devices is a function of the location of other nodes. For a given node, it is *per se* unclear how many neighboring nodes it can directly communicate with, and whether or not there is a multi-hop path to other devices.

Even with sufficient connectivity, limited data *throughput* can render a multi-hop network useless. Large distances between mobile devices affect the throughput in sparse networks. Mutual interference between devices limits the throughput in dense networks. In such cases, multi-hop networks have difficulties sustaining applications with high data rate demands, such as multimedia applications.

For addressing the issues of connectivity and throughput – and ultimately for improving over the current state of art, smart antennas and multi-user detection play an important role in this thesis. These are wide and diversified research areas which have received ample attention in the signal processing research community. The focus of this work is strongly put on involved *networking* aspects.

## 2.3 Modeling and Analysis

Today, the analysis of multi-hop networks is mostly based on mathematical methods and computer-based simulation. Prototypes and testbeds are often prohibitive due to time and cost efforts. As an interesting meta-solution, some researchers use hardware-in-the-loop devices which operate like real devices of a wireless network, while other parts are emulated by software (e.g. [RSO+05]).

For mathematical and simulative analysis, the considered multi-hop network has to be represented by models that are suitable for mathematical methods and computer programs, respectively. Such models should be accurate enough to capture relevant aspects of real networks, yet simple enough to be useful. Besides, they should be generic enough to allow for general conclusions. The choice of a model typically depends on the desired sort of analysis. A detailed model may make a problem too complex for its analysis. On the contrary, a simplistic model may not allow for meaningful results.

Splitting up a complete wireless network into models is often guided by the functional separation of real devices and software. Most notably, the Open Systems Interconnection (OSI) layers are used for both implementation and modeling. But

not the complete stack of hardware, communication protocols and applications has to be modeled for a given problem. For instance, for research on signal processing it may suffice to model the characteristics of the wireless channel, while a researcher working on security issues may be able to make contributions with virtually no channel assumptions.

In this thesis, models span from channel and mobility models to teletraffic models and queuing. The resulting plethora of modeling options, parameters and possible interdependencies between models suggests to use models that are as general as possible, omitting details of secondary importance, and still capturing essential effects and aspects. The used models and parameters are summarized in the following.

### 2.3.1 Scenarios and Node Distributions

In Sec. 2.1 different architectures of present and future wireless systems have been discussed. This thesis considers the special case of multi-hop networks which are pure ad hoc networks without any infrastructure, and all devices are assumed to be basically equal. Still, the major ideas of this work are well applicable to other network architectures in which multi-hop communication plays a major role.

Let a multi-hop network be composed by $N$ devices, or *nodes*, forming a node set $\mathcal{N}$. In the network topologies of this thesis, these nodes are placed independently in a two-dimensional square system area, according to a certain stochastic process.

For homogeneous node distributions in space, this stochastic process follows a uniform distribution. A sample of the resulting node distribution is used e.g. in Fig. 4.13 on page 98. It has properties similar to humans located in a pedestrian area, sensor devices in habitat monitoring, or networked buildings or cars on a city scale.

In many cases, inhomogeneous node distributions are of interest as model for realistic scenarios. Then, nodes form clusters in space with more or less tight interconnection. Such distributions are generated in this thesis by first placing $c$ cluster centers according to the above uniform distribution, and then placing $\frac{N}{c}$ nodes per cluster with a two-dimensional Gaussian offset from the respective cluster center. The standard deviation of the Gaussian offset is chosen as 10% of the edge length of the square system area. The resulting Gaussian mixture is truncated such that all nodes are located within the system area. This is achieved by simply re-placing nodes according to the Gaussian distribution around the cluster center until its coordinates are within the system area. We choose $c = 5$ since a number of clusters close to $1$ or close to $N$ would not result in a substantially clustered network. A sample network for the inhomogeneous case is used e.g. in Fig. 3.11 on page 38.

### 2.3.2 Mobility Model

Mobile scenarios are generated by placing nodes in a uniform fashion, and then letting them move according to the *random direction mobility* model. Nodes reaching the border of the system area are "reflected" back so as to keep the number of nodes in the network

constant. As shown in detail in [Bet03a], the random direction mobility model with this so-called *bounce back* border behavior results in a stationary uniform node distribution.

According to the random direction mobility model, nodes choose a random direction, move for a random time, then pause for a certain time and randomly select a new direction. This procedure is continued indefinitely. In the simulations of this work, the random durations in between direction changes are uniformly distributed in $[0;\ 60\,s]$, and pause times are set to $0$. When re-selecting a direction, it is chosen uniformly in $[0;\ 2\pi]$ at random, i.e. there are no correlations between consecutive motion segments. The node velocity is fixed as indicated in the description of the respective simulation.

### 2.3.3 Teletraffic and Routing

The parts of this thesis concerned with medium access and throughput aspects require a model for teletraffic to be specified.

The traffic model assumes that data is generated as fixed size packets with negative exponentially distributed inter arrival times, which is a commonly used model. The offered traffic of a data source is varied by varying the inter arrival time of packets. This model is quite generic and not bound to the statistics of a specific application.

In case of transmission backlog, packets are queued on a First In First Out (FIFO) basis without any kind of packet prioritization.

The data packets are transmitted in a unicast fashion from the source node to the specified destination node. Unless traffic relationships are specified differently, each node $n \in \mathcal{N}$ generates traffic destined to a randomly chosen destination $n_D \in \mathcal{N} \setminus \{n\}$. If the destination cannot be reached by a single-hop transmission, it is assumed that an underlying routing protocol discovers the shortest path in terms of the number of hops between the source and the destination, and that all intermediate nodes along the shortest path forward packets accordingly. Thus, this thesis is not concerned with details of routing, and network coding is not considered.

### 2.3.4 Antenna Models

Although its focus is on networking, antenna models play an important role in this thesis. Antenna design allows the dissemination of energy into the propagation medium on transmission, and the collection of energy on reception, to be direction-dependent. Examples for directional antennas are horn, dish, PCB or Yagi antennas. Even a dipole antenna does not have perfectly *omnidirectional*, i.e. direction-independent characteristics.

In contrast to the aforementioned antenna types, *antenna arrays* can be steered electronically instead of mechanical steering. This is achieved using a beamforming circuitry, controlling the signals applied to or tapped from the antenna elements. The terms *adaptive* or *smart* antennas are also used for such multi-element antennas.

It is assumed here that the individual antenna elements are ideal isotropic radiators.

Two types of configurations are considered: Uniform Linear Array (ULA) antennas, where $m$ antenna elements are equally spaced on a line, and Uniform Circular Array (UCA) antennas, where $m$ antenna elements are placed in equidistant manner on a circle. Both configurations are repeatedly used in the literature. More details on the antenna models are given in Sec. 3.1.

### 2.3.5 Channel and Link Model

This section describes the basic channel and link model used in this thesis. Some particular parameters will be given in subsequent chapters.

The channel is modeled as Line of Sight (LOS) channel without shadowing or fast fading. As transmitted signals propagate, their attenuation is modeled by a modified free space path loss model with path loss coefficient $\alpha$, with $\alpha = 3$ throughout this thesis. It should be noted that we do not model the multi-path characteristics of the channel, i.e. angular signature and delay characteristics.

Given a transmission power $p_\mathrm{t}$ and a distance $d_{n_1,n_2}$ between a transmitter $n_1$ and a receiver $n_2$, the received power at the receiver is modeled as

$$p_r = p_\mathrm{t} \cdot \left(\frac{\lambda}{4\pi \cdot d_{n_1,n_2}}\right)^2 \left(\frac{d_0}{d_{n_1,n_2}}\right)^{\alpha-2}, \tag{2.1}$$

where $\lambda$ is the carrier wavelength, and $d_0$ is a reference distance. The path loss coefficient $\alpha = 3$ accounts for attenuation effects of reflection, diffraction and scattering [Rap02].

We have to account for the antenna gain of the transmitter and the receiver, as illustrated in Fig. 2.2. The antenna gain is effective both on transmission and reception. The gain of the transmitter $n_1$ in the direction of the receiver $n_2$ is denoted as $g_1$, and $g_2$ is the gain of $n_2$ in the direction of $n_1$.

Unfortunately, it is not possible to find a visual pattern representation where links exist as soon as patterns overlap. Instead, Fig. 2.2 illustrates linear antenna gains. As throughout this thesis, the antenna patterns are not stylized, but actual gain patterns resulting from the respective antenna model.

Considering antenna gains and a reference distance $d_0 = 1\,\mathrm{m}$, the power of the received signal is

$$p_r = p_\mathrm{t} \cdot g_1\ g_2 \cdot \left(\frac{\lambda}{4\pi}\right)^2 \left(\frac{1}{d_{n_1,n_2}}\right)^{\alpha}. \tag{2.2}$$

The models for the channel, antennas and nodes lead to the simplifying assumption of having only symmetric, bidirectional links in the network.

So far, we have only considered one transmitter-receiver pair. In multi-hop networks, we have to account for interference as well. Let node $n_1$ transmit a signal $s_1(t)$ to node $n_2$, and a node $n_3$ transmit a signal $s_2(t)$ to a node $n_4$. The received signals at $n_2$ and $n_4$ are $r_1(t)$ and $r_2(t)$, respectively. The mathematical model for the interference channel used in this thesis is an Additive White Gaussian Noise (AWGN)

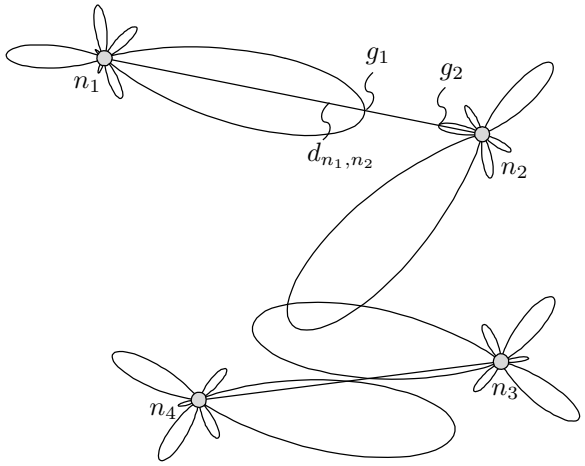



Figure 2.2: Illustration of directional transmission and directional reception.

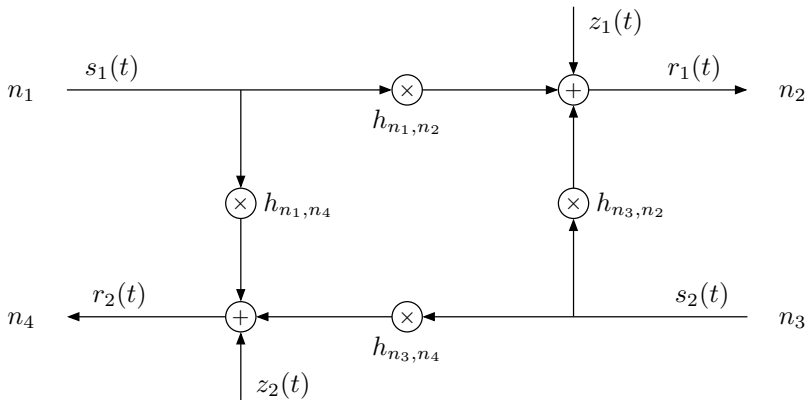



Figure 2.3: Interference channel for two transmitter-receiver pairs.

channel [Pro00], which abstracts the channel as shown in Fig. 2.3. The signal $r_1(t)$ received at $n_2$ has a noise component $z_1(t)$ and an interference component $s_2(t)$ originating from $n_3$, attenuated by a channel coefficient $h_{n_3,n_2}$, i.e.

$$r_1(t) = h_{n_1,n_2} s_1(t) + h_{n_3,n_2} s_2(t) + z_1(t) \; . \tag{2.3}$$

This model can be extended straight forward to more than two transmitters.

Threshold models are then used to decide on whether or not a signal can be successfully received and decoded. This is the case when the received signal power exceeds the receiver sensitivity $p_{r,\min} = -81$ dBm, and the threshold $\mathrm{SINR}_0$ is exceeded:

$$\mathrm{SINR} = \frac{p_r}{\sum p_i + p_z} \geq \mathrm{SINR}_0 \; , \tag{2.4}$$

where $p_r$ is the power of the desired signal, the $p_i$ are the interference powers at the receiver, and $p_z$ is the prevailing noise power.

Using this interference channel model, we can compute Signal to Interference and Noise Ratio (SINR) values for each transmission in the multi-hop network. Error-free transmissions are assumed when $\mathrm{SINR}_0 = 10$ dB is exceeded. Below this value, the collision model depends on the modulation and coding scheme, as will be discussed in more detail in the material on multi-user detection.

### 2.3.6 Numerical Analysis

Unless stated otherwise, the data points of the simulation results in this work have a confidence level of 95% for a confidence interval of $\pm 5\%$ around the simulation average. The underlying assumptions are independent and identically distributed samples, and the assumption that the central limit theorem holds. At least five samples were gathered before deciding on confidence.

## 2.4 Synchronization of Multi-Hop Networks

Several methods of this thesis require that the wireless devices forming a network are synchronized in time. The synchronization of networks can be categorized into

- synchronization to an absolute global time [EGE02, GKS03, MKSL04], and
- slot synchronization [HS05a, HS05b, TAB06].

A synchronization to an absolute time can be important to furnish sensor data with time labels. Slot synchronization is a prerequisite e.g. for slotted MAC protocols.

The synchronization can be guided by one or several pacemaker nodes (master nodes), which define a global time or – for slot synchronization – the slot beginnings. This requires a node hierarchy defining pacemaker nodes. A different approach is distributed synchronization, where synchronization is achieved in a peer-to-peer manner without pacemaker nodes. The references given above can thus be further categorized into

– hierarchical synchronization [GKS03, HS05b], and
– distributed synchronization [EGE02, MKSL04, HS05a, TAB06].

In this thesis, we always suppose *slot synchronization*. Since the methods of this thesis are all distributed, we prefer *distributed* self-organizing synchronization approaches.

A promising scheme in this respect is inspired by a role model in biology. This role model has been described mathematically, and adopted for wireless network synchronization (Fig. 2.4). In the course of this work, MAC layer adaptations have been made to this synchronization method, which enabled an implementation on wireless sensor modules (Crossbow MICAz). In experiments, a network synchronization of up to $3\,\mu$s was achieved, which would be sufficient for the methods of this thesis. Throughout the subsequent chapters, we assume that the network is already synchronized to slots.

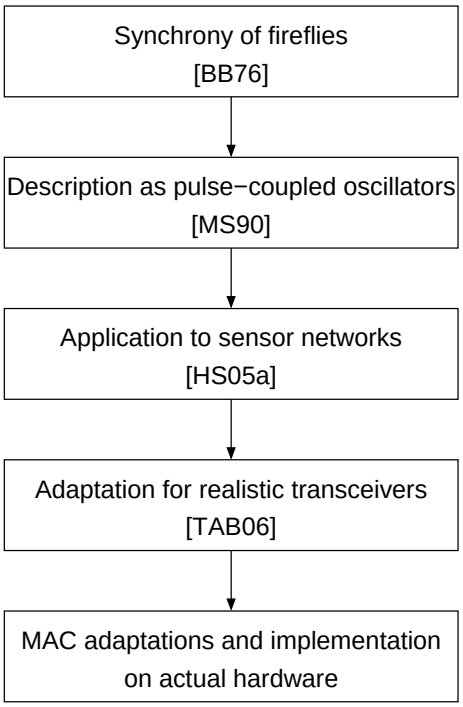
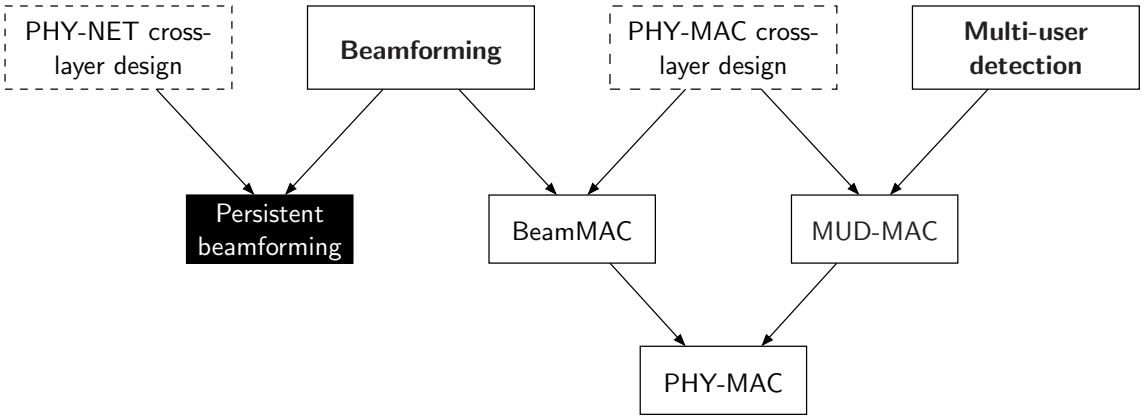



Figure 2.4: Toward distributed slot synchronization.

# 3 Persistent Beamforming




The use of beamforming antennas in packet-based wireless networks can be categorized into two approaches. The first approach is to carry out beamforming on a packet basis. Then, transmitting and/or receiving nodes adjust their beamforming pattern for the duration of each data packet, separately. This means that the beamforming may change in the order of milliseconds, since subsequently transmitted (received) packets can be destined to (received from) different neighboring nodes. According to the second approach, the beamforming pattern of the antennas is maintained on a larger time scale, e.g. for the duration of a data flow, a connection between a client and a server, or for even longer time periods.

In this chapter we look at the second approach, which is in the following referred to as *persistent beamforming*. With persistent beamforming, the beamforming is adjusted to traffic relationships on a networking level, rather than to instantaneous channel conditions between transmitters and receivers. We can thus categorize all methods and approaches of this chapter as cross-layer design between the physical layer and the network layer.

The work presented in this chapter is one of the first steps toward this kind of cross-layer interaction for beamforming antennas. Most publications on beamforming in wireless multi-hop networks focus on the above-mentioned first approach. It was shown that standard MAC approaches do not work well with per-packet beamforming [KSV00, Ram01, CYRV06], which is why the vast majority of these publications only considers MAC protocol design (cf. survey in Sec. 5.2).

Besides medium access, some previous research addresses the design of neighbor discovery [Ste03], routing [Rho04, NC06, RSB+03, NMMH00, SJ04], broadcasting and multi-casting [HHH03, WNE02], power control [AKM04], and capacity issues [PS03, YPK03, SR03b, SR03a] with beamforming antennas. Interestingly, the actual changes

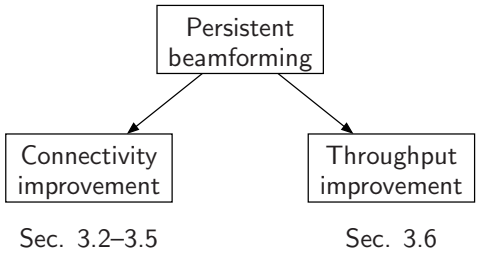
in the network topology – i.e. the properties of the network graph resulting from beamforming in comparison to the properties of the graph resulting from omnidirectional antennas – have long not been studied. This lack of research is very surprising since the network topology has significant impact on various performance measures, such as route lengths, end-to-end delay, reachability, and capacity. An adequate understanding of these changes in topological properties is therefore crucial for assessing the benefits of beamforming in multi-hop networks.

This chapter provides a thorough analysis of connectivity-related aspects with beamforming antennas. It demonstrates that significant improvements in terms of connectivity – but also throughput – can be achieved with persistent beamforming.

Since the system dynamics introduced by persistent beamforming are significantly lower than with per-packet beamforming, persistent beamforming can be deployed without changing network functionalities such as medium access control, neighbor discovery, or routing.

Sec. 3.1 of this chapter reviews signal processing methods for adaptive beamforming, used in subsequent sections as well as in chapter 5. Sec. 3.2 summarizes connectivity aspects and terms. Sec. 3.3 is devoted to optimizing the node degree of multi-hop networks assuming full network knowledge. Algorithms for improving the connectivity of practical multi-hop networks are subject-matter of Sec. 3.4. Improving connectivity can be accomplished by interconnecting disconnected network partitions. Sec. 3.5 considers such partitioning. It diversifies the material of this work by discussing a quite unorthodox class of algorithm: emergent behavior. Persistent beamforming can also be used to increase the throughput of multi-hop networks. Sec. 3.6 suggests means for throughput improvements by adaptive, yet still persistent, interference suppression. Sec. 3.7 reviews related work in the field. Finally, Sec. 3.8 summarizes this chapter.

This chapter is in part based on research previously published in [VBMH05, VBH05, VWAH06, VWAH07, HKV08].




## 3.1 Signal Processing for Digital Beamforming

Signal processing methods for multi-element antennas can roughly be categorized into

- beamforming methods,
- diversity schemes (e.g. antenna selection), and
- spatial multiplexing.

In this thesis, multi-element antennas are used for beamforming only. In subsequent sections, beamforming methods are used for a) range extension and b) interference suppression. Two pertinent signal processing methods from the literature are presented in this section. The first one is a standard approach for phased arrays, maximizing the antenna gain in a given direction (Sec. 3.1.3). The second one is an approach to nullify the antenna gain in one or several directions (Sec. 3.1.4). The simulation tool developed in the course of this work comprises an implementation of both methods.

It is important to note that the PHY-NET cross-layer interaction in this chapter is not bound to these two methods. The same holds for chapter 5, where the methods are applied as well. The methods are rather exemplary, revealing general aspects of cross-layer interaction with beamforming antennas in a multi-hop context that would arise in a similar way with different, more advanced beamforming schemes. In this regard, the reader is referred to [TKU05] for a comprehensive collection of state-of-the-art methods and algorithms for smart antennas.

### 3.1.1 Digital Beamforming

The general structure of a digital beamformer is illustrated in Fig. 3.1. It consists of a digital signal processor and one transceiver unit for each of the $m$ antenna elements. On transmission, the transmit signal is supplied to the digital signal processor. The processor creates $m$ binary streams (baseband signal composed of in-phase (I) and quadrature-phase (Q) component), one for each transceiver. On reception, the processor receives $m$ signals from the transceivers and outputs a combined receive signal.

Each transceiver performs frequency up-/down-conversion, filtering and amplification, and thus the conversion between the baseband signal and the narrow band band-pass signal applied to/tapped from the antenna element. The conversion between the digital and the analog regime is performed using an ADC at the transceiver side of the digital signal processor.

With ideal filters, sufficient sampling rate and high ADC accuracy, the digitally processed signals comprise all information that is available at the antenna array. The concept of digital beamforming can then be used for a wide range of beamforming methods (subsuming analog methods), allowing for shaping the antenna response with *beams* and *nulls* in a relatively arbitrary way.

Two methods that may be implemented in the digital signal processor will be described shortly, after a brief summary of antenna array concepts and terms.

### 3.1.2 Antenna Array Concepts and Terms

The most important concepts of antenna array theory for this work are summarized in this section [Bal97, LL96]. The *reciprocity* of antenna arrays makes the following definitions for transmitting antennas applicable to receiving antennas as well.

**Radiation pattern and antenna gain** The *radiation pattern* of an antenna array describes its relative distribution of radiated power as a function of angular direc-

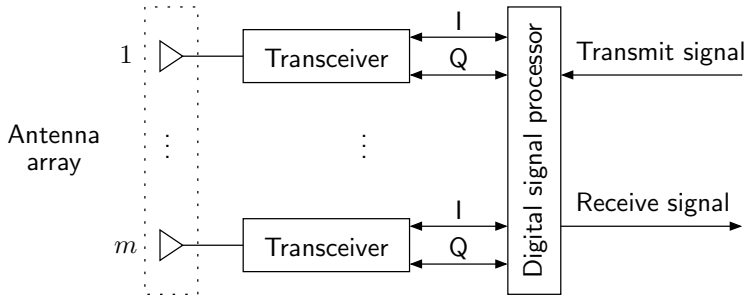



Figure 3.1: Digital beamforming [LL96].

tion. This pattern is, in the far-field of the array, quantified by the *antenna gain*, which is defined as

$$g(\theta, \phi) = \frac{4\pi \cdot \text{power radiated per unit solid angle in direction } \theta, \phi}{\text{Total input power to antenna}}, \tag{3.1}$$

where $\theta$ is the azimuthal angle and $\phi$ is the elevation angle. In this thesis, the *antenna efficiency* is assumed to be 1, which means that the total power radiated by the antenna is equal to the total input power. The antenna gain $g(\phi, \theta)$ is then equal to another quantity from antenna theory, namely the *directive gain*. In practice, the antenna efficiency is smaller than 1.

**Array factor and pattern multiplication** The far-field radiation pattern of an array depends on the geometric arrangement of its antenna elements. The notion of the *array factor* $F(\theta, \phi)$ is used to represent the radiation pattern for an array with isotropically radiating antenna elements. For non-isotropical elements – all realistic elements are non-isotropical – the *pattern multiplication principle* states that the overall antenna gain is obtained by multiplying the array factor $F(\theta, \phi)$ with the radiation pattern $f(\theta, \phi)$ of the individual antenna elements,

$$g(\theta, \phi) = f(\theta, \phi) \cdot F(\theta, \phi). \tag{3.2}$$

In this thesis, the antenna elements are assumed to be perfectly isotropical with $f(\theta, \phi) = 1$. The pattern multiplication principle is still important since it provides that the methods of this thesis can be applied to realistic arrays as well.

**Main lobe and side lobes** The radiation pattern of a beamforming antenna typically exhibits noticeable lobes (or *beams*). The lobe containing the maximum radiation power is called *main lobe*. All other lobes are called *side lobes*.

### 3.1.3 Array Factor and Main Lobe Steering

#### Linear Arrays

The far-field antenna pattern of an array is described by the array factor $F(\theta, \phi)$. For the ULA geometry with $m$ antenna elements as shown in Fig. 3.2, the array factor is given by

$$F(\theta, \phi)_{\mathrm{ULA}} = \sum_{i=0}^{m-1} x_i \, e^{ji\pi \sin\theta}, \tag{3.3}$$

i.e. by the inner product of the complex-valued *weight vector* $\mathbf{x}$ and the array propagation vector describing the Direction of Arrival (DOA) information of the incident wave front.

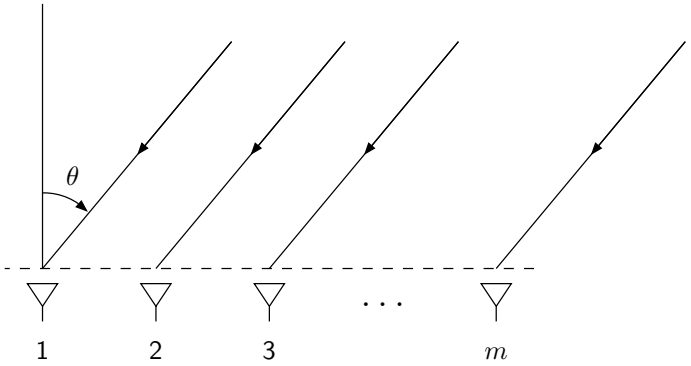



Figure 3.2: Geometric configuration of a ULA antenna with far-field wavefront.

For a ULA, the antenna pattern is rotationally symmetric, its array factor therefore independent of the elevation angle $\phi$. The above array vector holds for an element spacing of half the carrier wavelength. This is assumed throughout this work. The array factor for a general element spacing can be found in [LL96].

When assuming an equal phase difference $\Delta$ between the complex-valued entries of $\mathbf{x}$, these $x_i$ with magnitude $A_i$ can be expressed as

$$x_i = A_i \, e^{ji\Delta}, \quad i \in 0, \ldots m-1. \tag{3.4}$$

With this phase shift $\Delta$ between consecutive antenna elements, the array factor becomes

$$F(\theta, \phi)_{\mathrm{ULA}} = \sum_{i=0}^{m-1} A_i \, e^{j\,(i\pi \sin\theta + i\Delta)}. \tag{3.5}$$

Consequently, the antenna response can be maximized in a direction $\theta_0$ by providing that

$$\Delta = -\pi \sin\theta_0, \tag{3.6}$$

and $\Delta$ thereby becomes a means for steering the main lobe of the ULA.

### Circular Arrays

The array factor of an UCA antenna can be derived in an analogous manner. With UCA antennas, the $m$ radiators are evenly spaced on a circle. This is exemplified in Fig. 3.3 for $m = 8$. Since there is no rotational symmetry as with linear arrays, an elevation angle $\phi$ has to be considered in addition to the azimuthal angle $\theta$ in order to obtain the three-dimensional antenna characteristic. The following description is simplified by only looking at the lateral cut of the antenna pattern. This is sufficient for the numerical evaluations of this work, where planar multi-hop networks are assumed, with antenna arrays perfectly perpendicular to the ground.

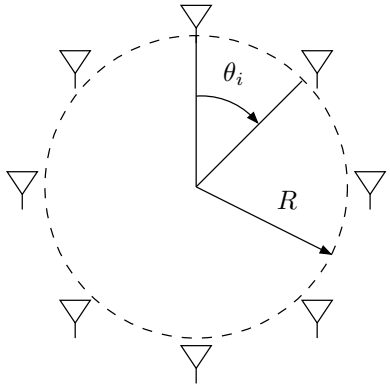



Figure 3.3: Geometric configuration of a UCA antenna.

With this simplification, the array factor expressing the phase shift with respect to the center of the UCA is

$$F\left(\theta,\, \phi{=}\frac{\pi}{2}\right)_{\mathrm{UCA}} = \sum_{i=0}^{m-1} x_i\, e^{-j\frac{2\pi}{\lambda}R\cos(\theta-\theta_i)}, \tag{3.7}$$

where $\lambda$ is the carrier wavelength, $R$ is the radius of the circle on which the antenna elements are arranged, and $\theta_i$ is the angular location of the $(i{+}1)$-th element on the circle.

Again, it is possible to steer the main lobe into a direction $(\theta_0, \phi_0)$. In our case, $\phi_0 = \frac{\pi}{2}$, since we want the main lobe to be aligned with the planar network. The azimuthal main lobe direction $\theta_0$ can be set by choosing

$$x_i = A_i\, e^{j\frac{2\pi}{\lambda}R\cos(\theta_0-\theta_i)} \tag{3.8}$$

for the elements of the weight vector.

## 3.1.4 Adaptive Nulling

The concept of adaptive *nulling* is used to suppress co-channel interference by nullifying the antenna gain in the angular direction of incident interference. A variety of nulling

methods can be found in the literature. Looking at early work, [App76] has considered nulling where the degrees of freedom remaining after nulling are used for maximizing an SIR index. A different scheme in [DM69] maximizes the antenna gain with nulling constraints. In [Ste82], Steyskal determines weight vectors minimizing the mean square difference between the antenna pattern before nulling and the antenna pattern after nulling, subject to desired nulling directions.

The latter method is adopted in this work. With this method, placing and removing nulls leads to only minor changes of the remaining antenna pattern, at least when there is a reasonable number of antenna elements available. Consequently, as nulls are added and removed by the MAC layer (in PHY-MAC cross-layer design) or networking layer (in PHY-NET cross-layer design), the behavior of the antenna array is "predictable" for these layers.

With persistent beamforming approaches described in this chapter, the beamforming pattern is maintained for different communication partners, and not adapted per-packet. Therefore, nulling methods maximizing an SIR index with respect to one single communication partner are inappropriate, while the *synthesis method* of [Ste82] is well-suited for this purpose.

### Synthesis Method

Assume that an adaptive array antenna is used to form an antenna gain pattern $g(\theta)$, where $\theta$ is the azimuthal angle. In the following, we assume that an array vector $\mathbf{x}_0$ has already been chosen to form a pattern $g_0(\theta)$. According to [Ste82], we call this the *quiescent* pattern. A special case of the quiescent pattern is a pattern without a directive characteristic, i.e. an omnidirectional pattern, where the antenna gain is equal to one in all directions. This pattern can be achieved by setting one entry of $\mathbf{x}_0$ to one, and the remaining entries to zero.

The synthesis method is used to impose antenna gain nulls on the quiescent pattern (no matter whether it is omnidirectional or directive), resulting in an *approximate* pattern $g_a(\theta)$. The term *approximate* describes that we are interested in the closest approximation to the quiescent pattern, i.e. we want to change the original pattern as little as possible when imposing nulls.

The approximate pattern is subject to the nulling constraints. A nulling constraint specifies that the antenna gain shall be zero in a given angular direction. In addition, we can demand the $\nu$-th derivative of the pattern to be zero. This may be necessary to suppress wide-band interference in a given angular direction, and thereby counteract the frequency dependency of the antenna array. In this thesis, it is always assumed that the interference source is narrow-band enough such that imposing a zero-order null is sufficient to suppress its interference. Still, the following formulations include the general case $\nu \geq 0$.

For imposing $M_1$ nulls up to the $M_2$-th derivative, the corresponding constraints can be formalized as

$$\frac{\mathrm{d}^\nu}{\mathrm{d}\theta^\nu}\, g_a(\theta_\mu) \stackrel{!}{=} 0, \quad \mu = 1, \ldots M_1, \quad \nu = 0, \ldots M_2, \tag{3.9}$$

where $\theta_\mu$ is the desired angular direction of the $\mu$-th null. Since the Degrees of Freedom DOF of the antenna array are limited to $m$, we have to ensure that

$$M_1 (M_2 + 1) \leq m, \tag{3.10}$$

i.e. given $m$ antenna elements, the higher the order $M_2$ of the antenna gain nulls, the less nulls can be formed.

In [Ste82], it is shown that finding the best approximation $g_a(u)$ minimizing the error $\epsilon(g_a)$,

$$\epsilon(g_a) = \frac{1}{2} \int_0^{2\pi} |g_0(\theta) - g_a(\theta)|^2 \, \mathrm{d}\theta, \tag{3.11}$$

is equivalent to minimizing the mean square error of the corresponding weight vectors, i.e.

$$\min : \quad \|\mathbf{x}_0 - \mathbf{x}_a\|^2 . \tag{3.12}$$

The main idea in [Ste82] to determine the weight vector of the approximate pattern, $\mathbf{x}_a$, is that the weight vector of the quiescent pattern, $\mathbf{x}_0$, has two projections on orthogonal subspaces $Y$ and $Z$, respectively. The projection on the subspace $Y$ represents the constraints given by (3.9). The subspace $Y$ is of dimension $M_1 (M_2 + 1)$, one dimension per constraint. By decomposing $\mathbf{x}_0$ into the subspaces $Y$ and $Z$, i.e.

$$\mathbf{x}_0 = \mathbf{y}_0 + \mathbf{z}_0, \quad \mathbf{y}_0 \in Y, \ \mathbf{z}_0 \in Z, \tag{3.13}$$

it follows that $\mathbf{x}_a$ must be in subspace $Z$, and orthogonal to $Y$. This is illustrated in Fig. 3.4. The projection of $\mathbf{x}_0$ into subspace $Z$ is the best approximation of $\mathbf{x}_0$ under the nulling constraints, and therefore the desired solution $\mathbf{x}_a$.

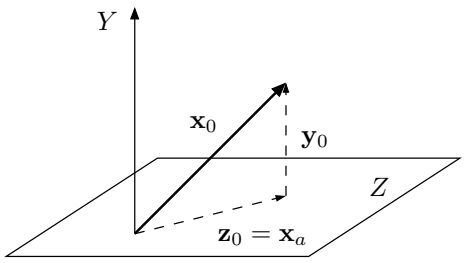



Figure 3.4: Illustration of the approximation problem [Ste82].

Speaking in terms of the desired antenna pattern, $g_a(\theta)$ consists of the quiescent pattern $g_0(\theta)$, and $M_1 (M_2 + 1)$ superimposed *cancellation beams*.

The problem of finding, finally, the desired weights $\mathbf{x}_a$ requires a linear equation system corresonding to Fig. 3.4 to be solved, as detailed in [Ste82].

### Application to Linear and Circular Arrays

The steps for computing the sought-after antenna pattern $g_a(\theta)$ for a ULA or UCA antenna are

- determining, from the quiescent weight vector $\mathbf{x}_0$ and the desired nulling directions $\theta_\mu$, $\mathbf{x}_a$ according to the synthesis method described above,
- substituting $\mathbf{x} = \mathbf{x}_a$ in the array factor (3.3) or (3.7), respectively, and
- normalizing the resulting pattern by the integral over all angular directions, such that the same energy is radiated with nulls as it is without nulls.

Sample approximate patterns $g_a(\theta)$ for different quiescent patterns and nulling directions for ULA and UCA antennas are given in appendix A. In the examples for ULA antennas, it can be seen that the antenna pattern is always symmetric to the antenna axis, and that imposing a null at $\theta_\mu$ causes a null at $-\theta_\mu$ as well.

## 3.2 Connectivity Concepts in Graph Theory

Multi-hop networks are modeled as graphs in this thesis. This has the merit that connectivity-related concepts from graph theory can be applied to characterize the connectedness of wireless devices forming a network. The following terms and concepts [DA93, GM95] will be used in subsequent sections:

- A graph is *connected* if there exists at least one path from each node to all other nodes. Otherwise, the graph is *disconnected*.
- If there exist at least $k$ node-disjoint paths between each pair of nodes, then the graph is *$k$-connected*. For instance, if a graph is 2-connected, all intermediate nodes forming the route between a source and a destination may fail and it is still provided that there is a second path that can be used as backup route. A routing protocol that is capable of discovering disjoint routes is called *multi-path routing protocol*.
- The *path probability* $p_{\mathrm{Path}}$ is defined as the probability that there exists a path between two randomly selected nodes. It can be determined by averaging over a large number $\Omega$ of random topologies:

$$p_{\mathrm{Path}} = \lim_{\Omega\to\infty} \frac{1}{\Omega} \sum_{i=1}^{\Omega} \frac{\text{number of connected node pairs}}{\text{number of node pairs}} \tag{3.14}$$

  The path probability is a relaxed measure as compared to the graph theoretical notion of connectivity, the latter defining a graph to be either connected or disconnected. The path probability is therefore a very useful measure to characterize the connectedness of a multi-hop network, and closely relates to the experience of a user wishing to connect to another device.
- The *node degree* $\delta$ of a node is equal to the number of neighbors of the node. For instance, a node with *links* (or *edges*) to three other nodes has a node degree of $\delta = 3$, while an *isolated* node has $\delta = 0$.

– The *Penrose Theorem* [Pen97, Pen99] interrelates node degrees to connectivity. Informally speaking, it states that a random geometric graph is connected with high probability when there is no isolated node in the graph.

*Random geometric graphs* are of great importance in this thesis. When placing nodes randomly in a system area and adding links between these nodes based on a maximum allowed node distance (= transmission range), then this results in a random geometric graph. A theorem analogous to the Penrose Theorem exists for purely *random graphs*, where the probability for the existence of a link is independent of the node distance.

The analysis of multi-hop networks with beamforming antennas is extremely complicated by the fact that such networks are neither strictly geometric graphs, nor strictly random graphs: The existence of links correlates with the node distance, but is furthermore dependent on the beamforming patterns of the nodes. Most of the numerical results in this chapter are therefore based on simulations.

### Significance of Node Degree and Path Probability

The notion of the path probability is obviously related to the usability of a multi-hop network. The higher the path probability, the better the user experience. Many applications may even require $p_{\mathrm{Path}} = 1$.

The significance of the node degree is more complex. On the one hand, higher node degrees are beneficial:

– There are more one-hop neighbors which can be communicated with very efficiently (no routing necessary, short delay),
– higher node degrees result in the existence of more node-disjoint paths, which can be important for fault-tolerant multi-path routing in mobile applications, and
– there is a higher potential for network coding techniques in case multi-casting is a communication requirement.

But there are also disadvantages:

– Higher node degrees mean higher contention on a MAC level, and thus reduced per-node channel access opportunities because of MAC layer blocking, and
– redundant message transmissions may occur excessively when flooding a message, leading to large flooding overhead.

For these disadvantages of high node degrees, Sec. 3.6 discusses means to reduce the node degree by nulling undesired neighboring nodes, resulting in higher data throughput. Among the primal and most limiting problems with multi-hop networks are, however, a lack of connectivity, large hop distances that disallow for high-quality multimedia applications for delay reasons, and short path life times in mobile scenarios. Unless the node density is very high and mobility is low, we should therefore still aim for rather high node degrees.

## 3.3 Optimal Beamforming with Node Degree Criterion

In this section, we control the beamforming directions of wireless devices so as to maximize the node degree, i.e. the number of neighbors a wireless device can directly communicate with. It is known from graph theory that there is a close relationship between the node degree and the overall connectedness of a network (Penrose Theorem), so maximizing the node degree is a major step toward optimizing connectivity.

The basic question answered in this section is:

> In which direction should the mobile devices point their main lobe so as to maximize the sum of the node degrees of all nodes in the multi-hop network?

This problem is formulated as a Mixed Integer Linear Program (MIP) [Dan63, BT97]. Even though such a central optimization approach with full network knowledge is not feasible in self-organizing multi-hop networks, the results obtained through optimization are of great interest: By comparison with the global optimum, we are able to assess distributed algorithms which can be applied to realistic multi-hop networks, and estimate their performance gap to the optimal solution. Such distributed algorithms will be discussed in later sections.

### 3.3.1 Modeling Assumptions

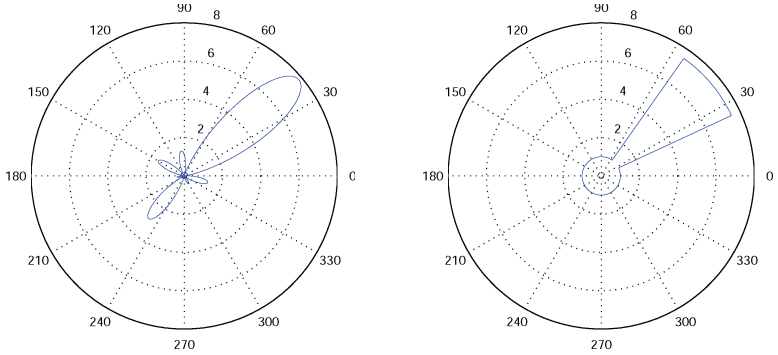
It is the intention to formulate the problem of maximizing node degrees as a MIP. This is virtually impossible for realistic antenna patterns as exemplified for a UCA with $m = 8$ antenna elements in Fig. 3.5(a).

To make the antenna model accessible for a linear formulation, its characteristic is in the following simplified and abstracted by the so-called *keyhole* model. It has been used repeatedly in the literature (e.g. [Ram01]). We assume that the mobile devices are located in a plane, and antennas are perfectly perpendicular to the ground. The keyhole model can thus be limited to two dimensions. It is then characterized by an angular range with high antenna gain $G_{\mathrm{M}}$ (main lobe), and a low antenna gain $G_{\mathrm{S}}$ in the remaining directions (Fig. 3.5(b)). $G_{\mathrm{S}}$ models the side lobes of a realistic antenna array. In the following, we assume $G_{\mathrm{M}} = 10$ dBi within an antenna aperture of $\alpha = 30°$, and $G_{\mathrm{S}} = -10$ dBi outside the aperture.

We assume that each node in the network's node set $\mathcal{N}$ is equipped with the same beamforming antenna characterized by $(G_{\mathrm{M}}, G_{\mathrm{S}}, \alpha)$, and that all nodes $n_i \in \mathcal{N}$ can adjust the direction $\gamma_{n_i}$ of the main lobe independently.

### 3.3.2 Linear Problem Formulation

We can start developing a linear formulation of the optimization problem by looking at two nodes $\{n_1, n_2\} \subset \mathcal{N}$ only. An illustration is given in Fig. 3.6, where two nodes $n_1$




(a) Realistic phased array. (b) Keyhole model.

Figure 3.5: Gain patterns.

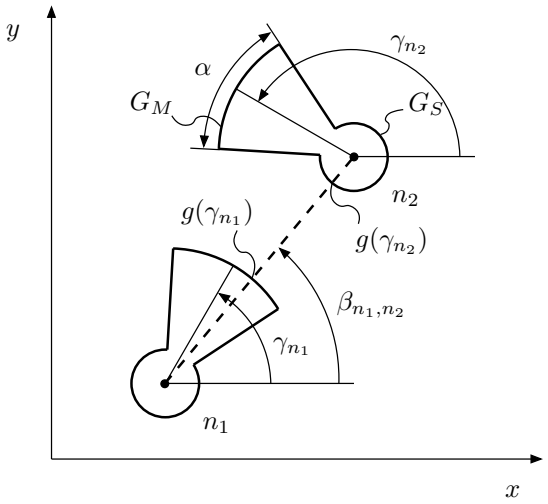



Figure 3.6: Illustration of optimization variables and parameters.

and $n_2$ with node coordinates $(x_{n_1}, y_{n_1})$ and $(x_{n_2}, y_{n_2})$, and main lobe steering angles $\gamma_{n_1}, \gamma_{n_2} \in [0; 2\pi[$ are shown.

The two nodes may or may not be able to establish a link $l_{n_1,n_2}$ between them, depending on the respective antenna gains $g(\gamma_{n_1})$ and $g(\gamma_{n_2})$, and depending on their Euclidean distance

$$d_{n_1,n_2} = \sqrt{(x_{n_2} - x_{n_1})^2 + (y_{n_2} - y_{n_1})^2}. \tag{3.15}$$

Using the path loss model (2.2) and a minimum required receive signal strength $P_{r,\min}$ (in decibel), we can formulate the link budget inequation in logarithmic notation by

$$0 \leq P_t + G(\gamma_{n_1}) + G(\gamma_{n_2}) - P_{L,n_1,n_2} - P_{r,\min}\,, \tag{3.16}$$

where $P_t$ is the transmission power, $G(\gamma_{n_i})$ are the antenna gains in decibel with

$$G = \begin{cases} G_{\mathrm{M}} & \text{within aperture } \alpha, \\ G_{\mathrm{S}} & \text{else,} \end{cases} \tag{3.17}$$

and $P_{L,n_1,n_2}$ is the logarithmic path loss given by

$$P_{L,n_1,n_2} = -10 \cdot \log\left(\frac{\lambda}{4\pi}\right)^2 - 10 \cdot \log\left({d_{n_1,n_2}}^{-\alpha}\right) . \tag{3.18}$$

The two nodes $n_1$ and $n_2$ are able to establish a link $l_{n_1,n_2}$ *iff* the received power $P_r$ is large enough to fulfill (3.16). This is the case within a maximum transmission range

$$d_{\mathrm{r}} = 10^{\frac{1}{10 \cdot \alpha} \cdot \left(-P_{r,\min} + P_{\mathrm{t}} + G\left(\gamma_{n_1}\right) + G\left(\gamma_{n_2}\right) + 10 \cdot \log\left(\frac{\lambda}{4\pi}\right)^2\right)} , \tag{3.19}$$

which is a function of the antenna gains. Looking at the keyhole model more closely, we can distinguish three basic cases: The nodes $n_1$ and $n_2$ are

- always out of communication range,
- always within communication range, or
- within communication range in case of proper main lobe steering angles $\gamma_{n_i}$.

### Always out of Communication Range

When $d_{n_1,n_2}$ is large, then $P_r$ may be smaller than $P_{r,\min}$ even if the two nodes point their main lobe toward each other. The critical distance for this is, with (3.19),

$$d_{\max} = d_{\mathrm{r}}\big|_{G\left(\gamma_{n_1}\right)=G_{\mathrm{M}},\, G\left(\gamma_{n_2}\right)=G_{\mathrm{M}}} . \tag{3.20}$$

Obviously, node pairs $\{n_1, n_2\}$ with $d_{n_1,n_2} > d_{\max}$ are not of interest for the optimization.

### Always within Communication Range

In case $d_{n_1,n_2}$ is so small that $P_r$ is sufficient independent of $\gamma_{n_i}$ and $\gamma_{n_2}$, i.e. not exceeding

$$d_{\min} = d_{\mathrm{r}}\big|_{G\left(\gamma_{n_1}\right)=G_{\mathrm{S}},\, G\left(\gamma_{n_2}\right)=G_{\mathrm{S}}} , \tag{3.21}$$

then there is always a link between $n_1$ and $n_2$. All node pairs $\{n_1, n_2\}$ with $d_{n_1,n_2} \leq d_{\min}$ can safely be ignored in the optimization.

### Within Communication Range in Case of Proper Steering

For all nodes pairs $\{n_1, n_2\}$ with

$$d_{\min} < d_{n_1,n_2} \leq d_{\max} \tag{3.22}$$

the existence of a link depends on the main lobe steering angles $\gamma_{n_1}$ and $\gamma_{n_2}$. In essence, these are the variables of the optimization.

According to Fig. 3.6, we can determine $\beta_{n_1,n_2} \in [0;\ 2\pi[$ by

$$\tan\left(\beta_{n_1,n_2}\right) = \frac{y_{n_2} - y_{n_1}}{x_{n_2} - x_{n_1}}\,. \tag{3.23}$$

For a given aperture $\alpha$ and an antenna orientation $\gamma_{n_1}$, $G(\gamma_{n_1})$ equals $G_\mathrm{M}$ (i.e. the signal is amplified by the high antenna gain) *iff*

- for $\frac{\alpha}{2} \leq \gamma_{n_1} \leq 2\pi - \frac{\alpha}{2}$:

$$\beta_{n_1,n_2} \leq \gamma_{n_1} + \frac{\alpha}{2} \quad \wedge \quad \beta_{n_1,n_2} \geq \gamma_{n_1} - \frac{\alpha}{2}, \tag{3.24a}$$

- for $2\pi - \frac{\alpha}{2} \leq \gamma_{n_1} \leq 2\pi$:

$$\beta_{n_1,n_2} - \frac{\alpha}{2} \leq \gamma_{n_1} \leq 2\pi \quad \vee \quad \beta_{n_1,n_2} \geq \gamma_{n_1} - \frac{\alpha}{2} + 2\pi, \tag{3.24b}$$

- and for $0 \leq \gamma_{n_1} \leq \frac{\alpha}{2}$:

$$0 \leq \gamma_{n_1} \leq \beta_{n_1,n_2} + \frac{\alpha}{2} \quad \vee \quad \beta_{n_1,n_2} - \frac{\alpha}{2} - 2\pi \leq \gamma_{n_1} \leq 2\pi. \tag{3.24c}$$

In the following, we will repeatedly use *indicator variables* with range $]-\infty;\ 1] \subset \mathbb{R}$. In the context of these indicator variables, a value of $1$ will always be used to indicate that a constraint is fulfilled, a link is active, etc. Using an indicator variable $a'_{n_1,n_2} \in ]-\infty;\ 1] \subset \mathbb{R}$, we can reformulate the first inequality of (3.24a) as

$$a'_{n_1,n_2} \leq \gamma_{n_1} + \frac{\alpha}{2} - \beta_{n_1,n_2} + 1\,, \tag{3.25a}$$

where $a'_{n_1,n_2}$ can be $1$ *iff* the first part of (3.24a) is fulfilled. Furthermore, we can put a tighter bound on $a'_{n_1,n_2}$, because the right-hand side of (3.25a) can only be as small as $-2\pi + 1 + \frac{\alpha}{2}$ with positive-valued angles. Consequently, $a'_{n_1,n_2} \in \left[-2\pi + \frac{\alpha}{2} + 1;\ 1\right]$ is sufficient to keep (3.25a) feasible.

For link budget calculations, we use $a'_{n_1,n_2}$ to set a binary variable $b'_{n_1,n_2}$ to $1$ *iff* $a'_{n_1,n_2} \leq 0$, and to $0$ otherwise. This can be achieved by the constraint

$$\left(-2\pi + 1 + \frac{\alpha}{2}\right) \cdot b'_{n_1,n_2} - a'_{n_1,n_2} + 1 \leq 0. \tag{3.25b}$$

The second inequality of (3.24a) can be formulated in a similar manner using a second indicator variable $a''_{n_1,n_2} \in \left[-2\pi + \frac{\alpha}{2} + 1;\ 1\right] \subset \mathbb{R}$ and a binary variable $b''_{n_1,n_2}$, resulting in

$$a''_{n_1,n_2} \leq \beta_{n_1,n_2} - \gamma_{n_1} + \frac{\alpha}{2} + 1 \tag{3.25c}$$

$$\left(-2\pi + \frac{\alpha}{2} + 1\right) \cdot b''_{n_1,n_2} - a''_{n_1,n_2} + 1 \leq 0. \tag{3.25d}$$

The variables $b''_{n_1,n_2}$ and $b'_{n_1,n_2}$ can become $0$ *iff* both constraints are fulfilled for a node pair $\{n_1, n_2\}$, allowing the binary variable $b_{n_1,n_2} \in \{0, 1\}$ to become $1$ in the constraint

$$b_{n_1,n_2} \leq \frac{1}{2}\left(1 - b'_{n_1,n_2}\right) + \frac{1}{2}\left(1 - b''_{n_1,n_2}\right). \tag{3.26}$$

Due to the angle discontinuity at $2\pi$ as apparent from (3.24b) and (3.24c), we still need to take care of those cases where $\beta_{n_1,n_2}$ is in the range of $\left[2\pi - \frac{\alpha}{2};\ 0\right]$ and $\left[0;\ \frac{\alpha}{2}\right]$. To this end, we use the following set of constraints in case of $\beta_{n_1,n_2} \geq 2\pi - \frac{\alpha}{2}$ for an indicator variable $c'_{n_1,n_2} \in \left[-\beta_{n_1,n_2} + \frac{\alpha}{2} + 1;\ 1\right]$:

$$c'_{n_1,n_2} \leq \gamma_{n_1} - \left(\beta_{n_1,n_2} - \frac{\alpha}{2}\right) + 1. \tag{3.27a}$$

Again, a binary variable $b'_{n_1,n_2}$ is generated by

$$\left(-\beta_{n_1,n_2} + \frac{\alpha}{2} + 1\right) \cdot b'_{n_1,n_2} - c'_{n_1,n_2} + 1 \leq 0. \tag{3.27b}$$

While the last constraint detects the obvious case where $\gamma_{n_1} \geq \beta_{n_1,n_2} - \frac{\alpha}{2}$, it is also possible that $\gamma_{n1} \leq \beta_{n_1,n_2} + \frac{\alpha}{2} - 2\pi$, which is detected by

$$c''_{n_1,n_2} \leq \left(\beta_{n_1,n_2} + \frac{\alpha}{2} - 2\pi\right) - \gamma_{n_1} + 1, \tag{3.27c}$$

and a binary variable is analogously set according to the fulfillment of the previous constraints in

$$\left(-4\pi + \beta_{n_1,n_2} + \frac{\alpha}{2}\right) \cdot b''_{n_1,n_2} - c''_{n_1,n_2} + 1 \leq 0. \tag{3.27d}$$

The opposing node $n_2$ is in the keyhole of the antenna *iff* one of the two binary indicator variables $b'_{n_1,n_2}, b''_{n_1,n_2}$ is $0$, which we indicate by

$$b_{n_1,n_2} \leq 1 - b'_{n_1,n_2} + 1 - b''_{n_1,n_2}. \tag{3.28}$$

The case where $\beta_{n_1,n_2} \leq \frac{\alpha}{2}$ (shown in (3.24c)) is formulated in a similar manner:

$$c'_{n_1,n_2} \leq \beta_{n_1,n_2} + \frac{\alpha}{2} - \gamma_{n_1} + 1 \tag{3.29a}$$

$$\left(-2\pi + \beta_{n_2} n_2 + \frac{\alpha}{2}\right) \cdot b'_{n_1,n_2} - c''_{n_1,n_2} + 1 \leq 0 \tag{3.29b}$$

$$c''_{n_1,n_2} \leq \gamma_{n_1} - \left(\beta_{n_1,n_2} - \frac{\alpha}{2} + 2\pi\right) + 1 \tag{3.29c}$$

$$\left(-\beta_{n_1,n_2} + \frac{\alpha}{2} - 2\pi + 1\right) \cdot b''_{n_1,n_2} - c''_{n_1,n_2} + 1 \leq 0 \tag{3.29d}$$

$$b_{n_1,n_2} \leq 1 - b'_{n_1,n_2} + 1 - b''_{n_1,n_2}. \tag{3.30}$$

At this point, we are able to formulate the conditioned link budget according to (3.16) as a linear constraint

$$\begin{aligned} 0 \leq & P_t + (1 - b_{n_1,n_2}) \cdot G_{\mathrm{S}} + b_{n_1,n_2} \cdot G_{\mathrm{M}} + \\ & + (1 - b_{n_2,n_1}) \cdot G_{\mathrm{S}} + b_{n_2,n_1} \cdot G_{\mathrm{M}} \\ & - l_{n_1,n_2} \cdot P_{L,n_1,n_2} - P_{r,\min} \, . \end{aligned} \tag{3.31}$$

All node pairs $\{n_1, n_2\} \subset \mathcal{N}$ can satisfy the above constraint by setting the binary link indicator variable $l_{n_1,n_2}$ to $0$. In order for a link to become active (i.e. $l_{n_1,n_2} = 1$), the link budget has to be sufficient. Maximizing the number of active links,

$$\max \sum_{\{n_1;n_2\} \subset \mathcal{N}} l_{n_1,n_2} \, , \tag{3.32}$$

causes the MIP solver to try to set as many $l_{n_1,n_2}$ to $1$ as possible. This requires nodes that are only within communication range in case of proper main lobe steering to turn their beamforming direction such that $b_{n_1,n_2}$ can be set to $1$ (representing a node in the main lobe), thereby finding a configuration with as many links as theoretically possible.

### Complete Formulation

For the link budget given by (3.16), a set of nodes $\mathcal{N}$, corresponding (precomputed) distances $d_{n_1,n_2} \in \mathbb{R}_0^+ \; \forall \{n_1, n_2\} \subset \mathcal{N}$, and angles between nodes $\beta_{n_1,n_2} \in [0;\, 2\pi[ \; \forall \{n_1; n_2\} \subset \mathcal{N}$, the complete mixed-integer program reads as follows:

$$\max \sum_{\{n_1;n_2\} \subset \mathcal{N}} l_{n_1,n_2} \tag{3.33}$$

subject to

– $\forall \{n_1; n_2\} \subset \mathcal{N}: \quad \beta_{n_1,n_2} < \frac{\alpha}{2} \;\; \wedge \;\; d_{\min} < d_{n_1,n_2} \leq d_{\max}$

$$c'_{n_1,n_2} \leq \beta_{n_1,n_2} + \frac{\alpha}{2} - \gamma_{n_1} + 1 \tag{3.34a}$$

$$\left(-2\pi + \beta_{n_1,n_2} n_2 + \frac{\alpha}{2}\right) \cdot b'_{n_1,n_2} - c''_{n_1,n_2} + 1 \leq 0 \tag{3.34b}$$

$$c''_{n_1,n_2} \leq \gamma_{n_1} - \left(\beta_{n_1,n_2} - \frac{\alpha}{2} + 2\pi\right) + 1 \tag{3.34c}$$

$$\left(-\beta_{n_1,n_2} + \frac{\alpha}{2} - 2\pi + 1\right) \cdot b''_{n_1,n_2} - c''_{n_1,n_2} + 1 \leq 0 \tag{3.34d}$$

$$b_{n_1,n_2} \leq 1 - b'_{n_1,n_2} + 1 - b''_{n_1,n_2} \tag{3.34e}$$

with

$$c'_{n_1,n_2} \in \left[-2\pi + \beta_{n_2} n_2 + \frac{\alpha}{2}; 1\right] \subset \mathbb{R},$$
$$c''_{n_1,n_2} \in \left[-\beta_{n_1,n_2} + \frac{\alpha}{2} - 2\pi + 1; 1\right],$$
$$b'_{n_1,n_2}, b''_{n_1,n_2}, b_{n_1,n_2} \in \{0,1\}.$$

– $\forall \{n_1; n_2\} \subset \mathcal{N}: \quad \frac{\alpha}{2} \leq \beta_{n_1,n_2} \leq 2\pi - \frac{\alpha}{2} \quad \wedge \quad d_{\min} < d_{n_1,n_2} \leq d_{\max}$

$$a'_{n_1,n_2} \leq \gamma_{n_1} + \frac{\alpha}{2} - \beta_{n_1,n_2} + 1 \tag{3.35a}$$
$$\left(-2\pi + 1 + \frac{\alpha}{2}\right) \cdot b'_{n_1,n_2} - a'_{n_1,n_2} + 1 \leq 0 \tag{3.35b}$$
$$a''_{n_1,n_2} \leq \beta_{n_1,n_2} - \gamma_{n_1} + \frac{\alpha}{2} + 1 \tag{3.35c}$$
$$\left(-2\pi + \frac{\alpha}{2} + 1\right) \cdot b''_{n_1,n_2} - a''_{n_1,n_2} + 1 \leq 0 \tag{3.35d}$$
$$b_{n_1,n_2} \leq \frac{1}{2}\left(1 - b'_{n_1,n_2}\right) + \frac{1}{2}\left(1 - b''_{n_1,n_2}\right) \tag{3.35e}$$

with

$$a'_{n_1,n_2} \in \left[-2\pi + \frac{\alpha}{2} + 1; 1\right] \subset \mathbb{R},$$
$$a''_{n_1,n_2} \in \left[-2\pi + \frac{\alpha}{2} + 1; 1\right] \subset \mathbb{R},$$
$$b'_{n_1,n_2}, b''_{n_1,n_2}, b_{n_1,n_2} \in \{0,1\}.$$

– $\forall \{n_1; n_2\} \subset \mathcal{N}: \quad 2\pi - \frac{\alpha}{2} < \beta_{n_1,n_2} \quad \wedge \quad d_{\min} < d_{n_1,n_2} \leq d_{\max}$

$$c'_{n_1,n_2} \leq \gamma_{n_1} - \left(\beta_{n_1,n_2} - \frac{\alpha}{2}\right) + 1 \tag{3.36a}$$
$$\left(-\beta_{n_1,n_2} + \frac{\alpha}{2} + 1\right) \cdot b'_{n_1,n_2} - c'_{n_1,n_2} + 1 \leq 0 \tag{3.36b}$$
$$c''_{n_1,n_2} \leq \left(\beta_{n_1,n_2} + \frac{\alpha}{2} - 2\pi\right) - \gamma_{n_1} + 1 \tag{3.36c}$$
$$\left(-4\pi + \beta_{n_1,n_2} + \frac{\alpha}{2}\right) \cdot b''_{n_1,n_2} - c''_{n_1,n_2} + 1 \leq 0 \tag{3.36d}$$
$$b_{n_1,n_2} \leq 1 - b'_{n_1,n_2} + 1 - b''_{n_1,n_2} \tag{3.36e}$$

with

$$c'_{n_1,n_2} \in \left[-\beta_{n_1,n_2} + \frac{\alpha}{2} + 1; 1\right] \subset \mathbb{R},$$
$$c''_{n_1,n_2} \in \left[-4\pi + \beta_{n_1,n_2} + \frac{\alpha}{2}; 1\right] \subset \mathbb{R},$$
$$b'_{n_1,n_2}, b''_{n_1,n_2}, b_{n_1,n_2} \in \{0,1\}.$$

– $\forall \{n_1; n_2\} \subset \mathcal{N}: \quad d_{\min} < d_{n_1,n_2} \leq d_{\max}$

$$\begin{aligned} 0 \leq & P_t + (1 - b_{n_1,n_2}) \cdot G_{\mathrm{S}} + b_{n_1,n_2} \cdot G_{\mathrm{M}} + \\ & + (1 - b_{n_2,n_1}) \cdot G_{\mathrm{S}} + b_{n_2,n_1} \cdot G_{\mathrm{M}} \\ & - l_{n_1,n_2} \cdot P_{L,n_1,n_2} - P_{r,\min} \,. \end{aligned} \tag{3.37}$$

with

$$b_{n_1,n_2}, l_{n_1,n_2} \in \{0, 1\} \,.$$

### 3.3.3 Numerical Evaluation

The solution of the optimization problem stated above yields those per-node beamforming directions which result in the maximum sum of node degrees in a multi-hop network. It would now be interesting to obtain statistical data on this maximum – and the path probability $p_{\mathrm{Path}}$ that comes along with it – for a large number of network topologies, and compare to multi-hop networks with omnidirectional antennas. Alas, the computational complexity of solving this optimization problem is too high to provide such data here. The application of the methodology is therefore limited to one topology. In the example, $N = 50$ nodes are placed on a square system area of size $1000\,\mathrm{m} \times 1000\,\mathrm{m}$. The nodes are placed inhomogeneously, i.e. the network is clustered. The optimization was implemented using CPLEX/Concert 9.1, running on an AMD Opteron 2218 workstation with $6\,$GByte RAM. It was interrupted after 12 hours runtime, with a residual worst-case optimality gap of 65.4%.

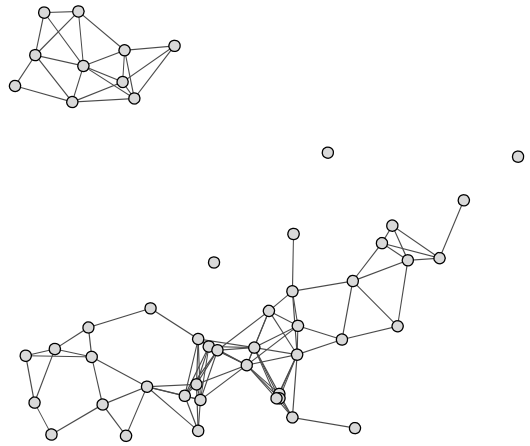
Although results are only provided for this one sample network, it is not just a visual validation of the solution. As we will see shortly, interesting observations can be made when comparing this optimal persistent beamforming configuration to per-packet beamforming.

Let us compare the following three cases:

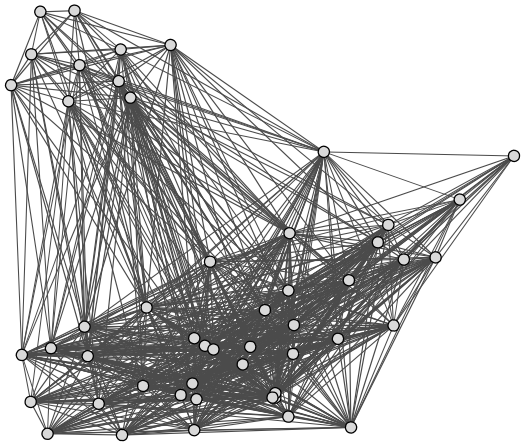
- Omnidirectional antennas ($G = 0\,\mathrm{dBi}$ for all nodes in all directions),
- per-packet beamforming, where it is assumed that all receivers and transmitters can provide the high antenna gain $G_{\mathrm{M}}$ for every single transmission taking place in the network,
- the optimal solution for persistent beamforming, maximizing the sum node degree according to Sec. 3.3.2.

The sample network and the resulting links for these three cases are illustrated in Fig. 3.7. The corresponding values for the sum node degree and the path probability are summarized in Tab. 3.1. Besides these three cases, Fig. 3.7 and Tab. 3.1 provide results for a distributed approach, Maximum Node Degree Beamforming (MNDB), which will be discussed later on.

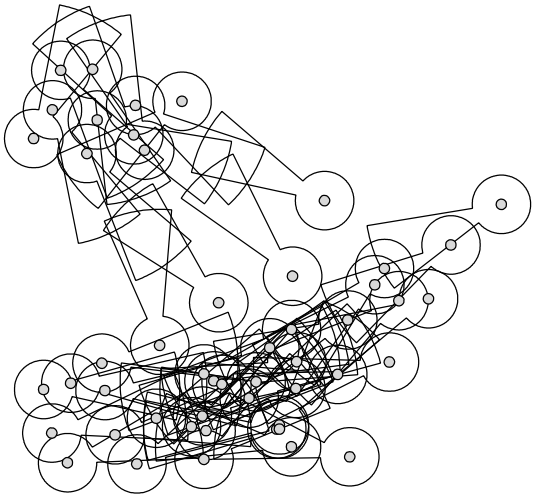
Looking at Tab. 3.1, the increase in the sum node degree over the omnidirectional case (232) is significant with optimal persistent beamforming (410), and quite drastic with per-packet beamforming (1732). For optimal persistent beamforming, it is interesting to note that the increase of the path probability is by far more significant than the



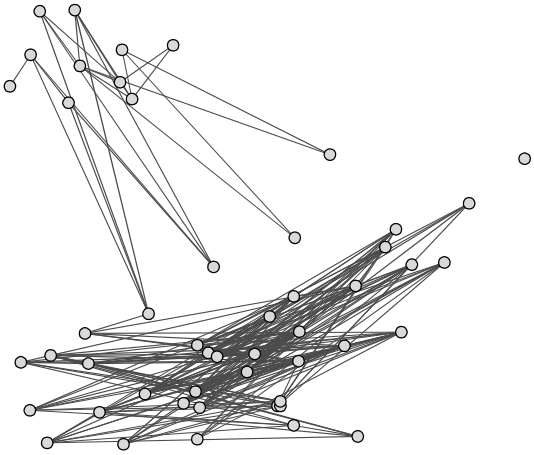
(a) Omnidirectional antennas.



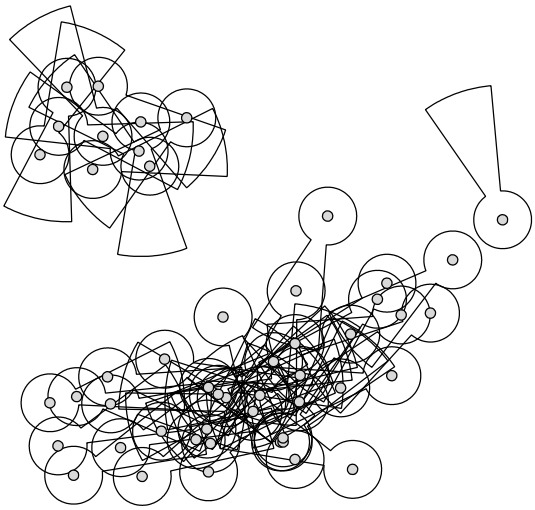
(b) Per-packet, fully directional.



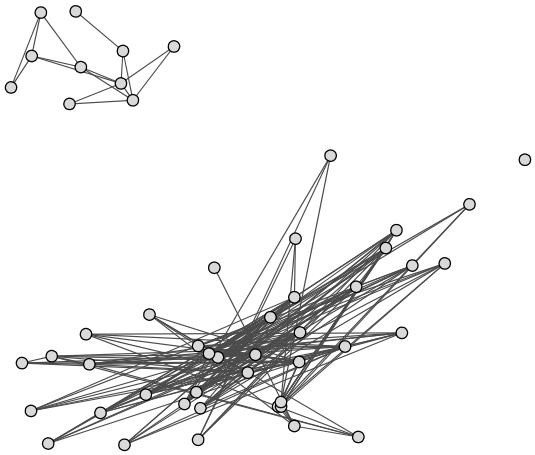
(c) Optimal persistent beamforming.



(d) Links resulting from optimization.



(e) Distributed MNDB.



(f) Links resulting from MNDB.

Figure 3.7: Connectivity in a sample network with $N = 50$ nodes.

■ Table 3.1: Connectivity measures evaluated for the sample network.

| | Sum of node degrees | Path probability |
|---|---|---|
| Omnidirectional | 232 | 0.580 |
| Distributed MNDB | 352 | 0.642 |
| Optimal persistent beamforming | 410 | 0.960 |
| Per-packet, fully directional | 1732 | 1 |

change in the sum node degree, although the latter is the objective function of the optimization. With only one node being disconnected (cf. Fig. 3.7(d)), persistent beamforming virtually reaches the path probability of per-packet beamforming.

We have here assumed that per-packet beamforming can establish fully directional links, i.e. links that can only occur if both the transmitter and the receiver point their main lobe toward each other. As will be discussed in chapter 5, it is not obvious that a beamforming-aware MAC protocol is able to establish fully directional links. It is an advantage of persistent beamforming that, even with very simple schemes, such links happen to occur. This will be illustrated later in connection with Fig. 3.11.

MNDB (described in Sec. 3.4.2) is a heuristic working on the same criterion as the above-presented optimal solution, i.e. the sum of node degrees. For the sample network, it yields a significant increase of the sum of node degrees, and a moderate improvement of the path probability as compared to omnidirectional antennas (cf. Tab. 3.1). While these are only qualitative observations drawn from one sample network, quantitative data for networks with varying node density will be provided in Sec. 3.4.3.

## 3.4 Distributed Beamforming Methods

Hitherto in this chapter, we have emphasized the fact that persistent beamforming can be deployed in multi-hop networks without adapting core networking and medium access functionalities. When assessing the practicality of beamforming, another important question is how much channel information and information about the antenna array setup is required. Non-persistent beamforming typically requires channel attenuation information for individual antenna elements (channel matrix of the Multiple Input Multiple Output (MIMO) system), or DOA information about neighboring nodes. This information has to be updated in the order of milliseconds, in case the beamforming is adjusted per packet.

In this section, we are interested in the following question:

> What connectivity gains can be achieved with beamforming antennas if the nodes have no channel information at all, and no or very limited information about the network topology?

### 3.4.1 Random Direction Beamforming

If a node has no information about beneficial beamforming directions, a naive and straightforward approach is to transmit in a random direction. Like omnidirectional transmission, beamforming in a random direction does not require any direction estimation among nodes. It is thus a very practical approach and has low complexity with respect to signal processing and signaling protocols.

Investigating beamforming in random directions is further interesting since it can be used to obtain performance bounds. This is to say that, if no neighbor location information is at hand, nodes may beamform randomly; once they obtain information about their neighbors, they can optimize the beamforming.

The authors of [BHM05] made a first step toward an analysis of network topology properties, where random direction beamforming was compared with omnidirectional transmission. It was found that, when using the same transmission power as with omnidirectional antennas, random direction beamforming significantly improves the level of connectivity among nodes. As a measure for connectivity, the path probability $p_{\mathrm{Path}}$ was used. It was further found that beamforming antennas with a small number of antenna elements are sufficient to increase the path probability significantly. Although this result gave a first insight on the gains in connectivity that can be achieved with beamforming, fundamental questions remained unanswered:

> What is the impact of directional transmission on the hop distance between two nodes?
>
> What hop distance can be expected by a source node for reaching a certain destination?
>
> How many hops are necessary to broadcast information to all nodes in the network?

An answer to these questions is by far not obvious. For example, one could argue that long links appearing with beamforming increase the connectivity, but may lead to "zigzag" routes and thus to increased hop distances. Resorting to beamforming for the purpose of flooding may be counterintuitive in the first place. Provided that the network is connected when using omnidirectional antennas, one might think that omnidirectional transmission is better suited for flooding information. In particular, if the flooding is started at a node located in the center of the network, it seems that omnidirectional transmission lets the flooded message advance more effectively in all directions than in the beamforming case.

Existing work on hop distances [BE03, VE05, MA05] does not give an answer to these questions since it focuses on omnidirectional antennas. We consider each of these questions in turn.

#### Description of the Algorithm

With Random Direction Beamforming (RDB), each node $n_i$ randomly chooses a main lobe direction $\gamma_{n_i} \in [0; 2\pi[$ relative to its antenna array direction, and instructs its

beamformer to maximize the antenna gain in this direction (Sec. 3.1.3). The main lobe direction in the system plane is given by taking into account $\gamma_{n_i}$ and the array orientation. Once $\gamma_{n_i}$ is chosen, the resulting pattern is fixed and used for transmission as well as for reception.

The formalized Algorithm 1 is carried out by all nodes in the network independently of each other.

**Algorithm 1**: Random Direction Beamforming

1 **initialize** $\gamma_{n_i} = \texttt{random()} \cdot 2\pi$ `/* random() → ∈ [0;1[ */`
2 Steer main lobe to $\gamma_{n_i}$

In a practical realization, each node may choose $\gamma_{n_i}$ when the device is turned on. We assume that $\gamma_{n_i}$ is fixed as long as the device is operative. Adjusting $\gamma_{n_i}$ is subject-matter of Sec. 3.4.2.

### Simulative Analysis

This section provides a simulative analysis of RDB with respect to several measures that are critical from a networking perspective. We consider two network scenarios with equal node density:

- A small network scenario with $N = 50$ nodes on a square system area of $100\,\mathrm{m} \times 100\,\mathrm{m}$, and
- a large scenario with $N = 5000$ nodes on a square system area of $1000\,\mathrm{m} \times 1000\,\mathrm{m}$.

In both types of networks, the nodes are placed homogeneously at random, as described in Sec. 2.3.1.

We are interested in graph theoretical measures related to the hop distances between nodes. The hop distance, i.e. the number of transmissions of a message between its source and destination, has a direct impact on the end-to-end delay of this message. Each node along its path causes various types of delay, such as coding and medium access delay. Furthermore, each hop consumes energy, which is a scarce commodity in most multi-hop networks. In mobile multi-hop networks, each additional hop of a route increases the probability of a route break during communication, which usually requires error detection and re-routing, both involving long delays caused by protocol mechanisms. Finally, the hop distance characteristics of a network topology have a direct impact on *broadcasting* and *flooding*. Broadcasting on a network level describes the message distribution from a source node to all other nodes in the network. The most commonly applied method for this is flooding. The number of hops that is necessary to flood the network, i.e. the number of retransmission steps of a message in a time-discrete manner, directly impacts the "costs" of flooding. These costs come in the form of the message overhead (*broadcast storm problem* [NTCS99]), the delay, resource consumption and temporal information inconsistency. The costs of flooding have thereby an immediate effect on the overall performance of mechanisms that

rely on flooding and broadcasting, such as routing protocols and data dissemination schemes.

To clearly characterize the impact of beamforming on hop distances, the following measures are investigated:

- The hop distance $h$ between two nodes, defined as the minimum number of edges forming a path in the graph between the two nodes, statistically described by the probability mass function $P(H = h)$ of the corresponding random variable $H$,
- the network diameter $d$, defined as the maximum of all hop distances in the network, statistically described by the probability mass function $P(D = d)$ of the corresponding random variable $D$, and
- the number of hops $f$ needed to flood the network, statistically described by the probability mass function $P(F = f)$ of the corresponding random variable $F$.

It is the goal of simulation efforts to obtain good estimates for the probability mass functions $P(H = h)$, $P(D = d)$, and $P(F = f)$.

For $P(H = h)$, we generate a random topology, measure the hop distances from each node to each other node, and construct a histogram $p_i(h)$. The histogram represents the hop distance distribution of the given topology. Unconnected node pairs are not counted since their hop distance is not defined. The process is repeated for a large number $\Omega$ of random topologies. The average

$$\frac{1}{\Omega}\sum_{i=1}^{\Omega} p_i(h) \quad \forall\, h \in \mathbb{N} \tag{3.38}$$

converges toward the probability mass function $P(H = h)$ for the respective type of topology. In the simulations, $\Omega$ is chosen large enough such that all $p_i$ have a confidence interval of 5% with a confidence level of 95%.

The methodology for obtaining $P(D = d)$ is straight-forward, since only one value can be obtained for each randomly generated topology. The generation of topologies is again repeated until the above-mentioned confidence is achieved.

For $P(F = f)$, each randomly generated topology with $N$ nodes provides $N$ values: The flooding process is started for one of the nodes, and the number of forwarding steps needed to completely flood the network (or the reachable sub-graph in case the network graph is disconnected) is counted. This process is repeated for each of the nodes. The probability mass function $P(F = f)$ is estimated by averaging over $N$ source nodes and $\Omega$ topologies, with $\Omega$ again being large enough such that the desired confidence is reached.

### Numerical Results and Discussion

Fig. 3.8 depicts the estimated $P(H = h)$ for the small and large network scenario, comparing omnidirectional, UCA4 (UCA with $m = 4$ antenna elements), and UCA10 antennas. The results show that the use of randomized beamforming lowers the

probability of large hop distances. In the small network scenario, hop distances $h \geq 5$ occur less likely with beamforming than with omnidirectional antennas. In turn, small hop distances $h = 2$, 3, and 4 occur with a higher probability. This decrease of hop distances becomes much more significant in the large network scenario. The results are emphasized by Tab. 3.2, which shows the expected value of $H$ in the two different scenarios. Using UCA4 antennas reduces $E\{H\}$ by 43% in the large network case, with UCA10 antennas the reduction amounts to 57%.

■ Table 3.2: Expected hop distance $E\{H\}$.

| | Small network $N = 50$ $100\,\mathrm{m} \times 100\,\mathrm{m}$ | Large network $N = 5000$ $1000\,\mathrm{m} \times 1000\,\mathrm{m}$ |
|---|---|---|
| Omnidirectional | 3.87 | 35.97 |
| UCA4 | 3.26 | 20.60 |
| UCA10 | 3.06 | 15.41 |

In summary, RDB significantly improves the statistics of the hop distance, especially for large networks.

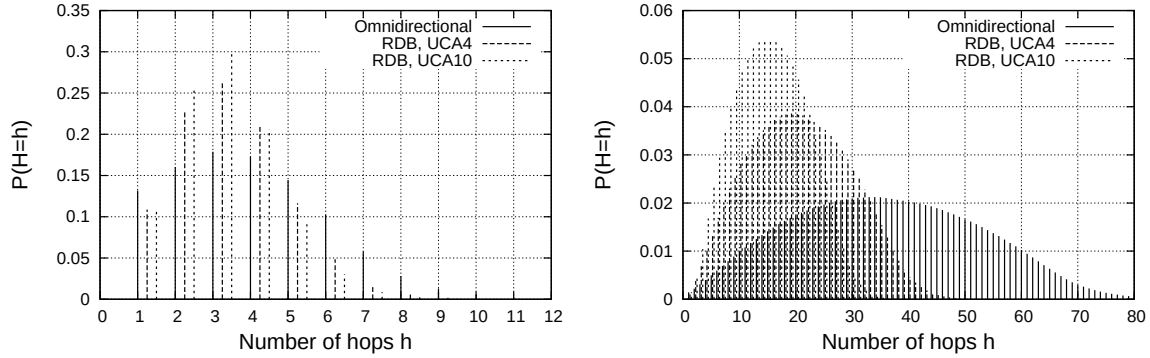
Fig. 3.9 shows the simulation results for the probability mass function $P(D = d)$. As expected from the results on $P(H = h)$, RDB reduces the network diameter. For example, the typical diameter of a large network with UCA10 antennas has less than half the diameter of the same network with omnidirectional antennas. Since the network diameter determines the maximum occurring delay in the transmission between two nodes, this result illustrates the benefits of randomized beamforming.

From the results on $P(D = d)$, it can be derived that RDB improves the worst case for flooding, i.e. the case where a node located at the border of the network starts flooding a message. As a generalization of this, we investigate the case where any node of the network may start the flooding process, i.e. we investigate the above-defined probability mass function $P(F = f)$. As Fig. 3.10 shows, randomized beamforming surprisingly reduces the number of hops needed to flood a network, in particular in large networks. This reduction in flooding hops would then carry over to faster message distribution and reduced signal processing in devices.

### 3.4.2 Iterative Distributed Approaches

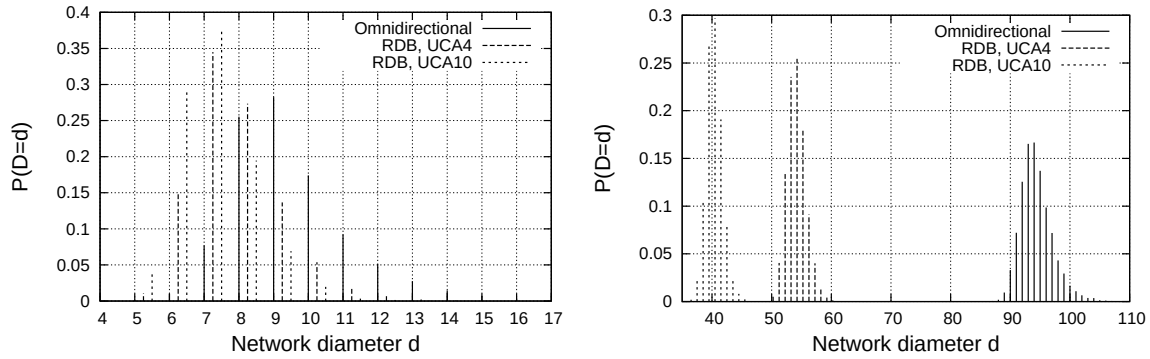
The RDB approach may be used as wake-up or fall-back strategy when operating a multi-hop network. It introduces no protocol overhead, no signal processing complexity for DOA estimation and beamforming adjustments, avoids convergence and stability issues in the interconnection of mobile devices, but still provides significant connectivity improvements.

Despite its advantages, beamforming in a random direction is not optimal. For instance, let us have a look at a clustered scenario with $N = 100$ nodes, depicted in Fig. 3.11. In Fig. 3.11(a), the nodes and links with omnidirectional antennas are shown. Fig. 3.11(b)



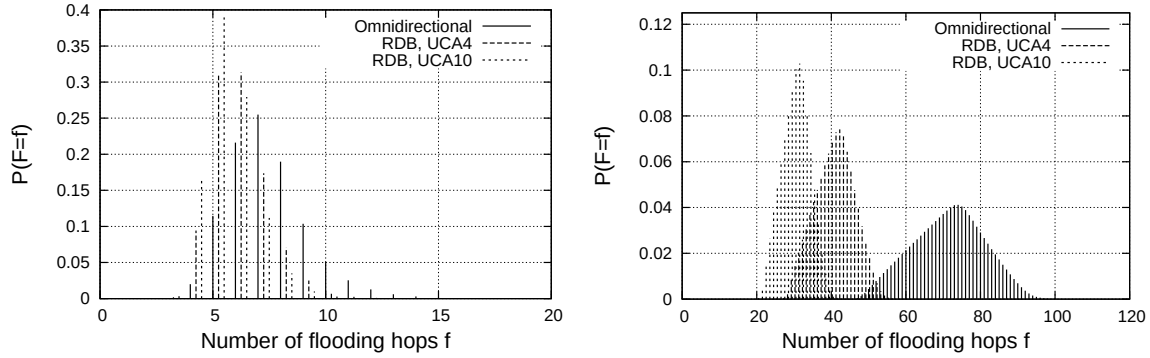

(a) Small network. (b) Large network.

Figure 3.8: Probability mass function $P(H = h)$ of the hop distance $h$ between two random nodes.



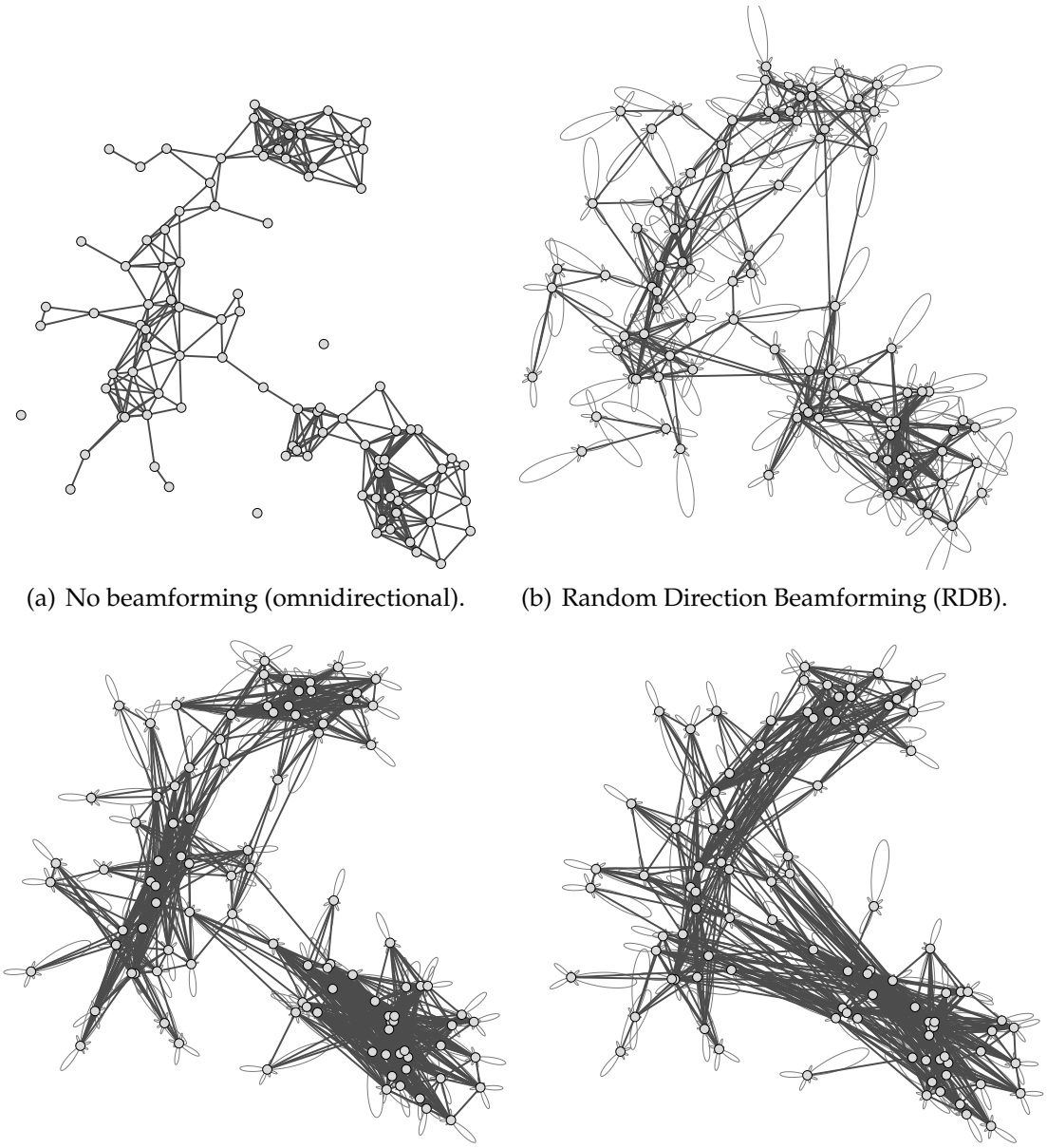

(a) Small network. (b) Large network.

Figure 3.9: Probability mass function $P(D = d)$ of the network diameter $d$.




(a) Small network. (b) Large network.

Figure 3.10: Probability mass function $P(F = f)$ of the number of hops $f$ needed to flood the network.



(a) No beamforming (omnidirectional). (b) Random Direction Beamforming (RDB).

(c) Maximum Node Degree Beamf. (MNDB). (d) Two-hop Node Degree Beamf. (TNDB).

Figure 3.11: Sample topologies resulting from the different beamforming approaches, for $N = 100$ nodes with inhomogeneous (clustered) node distribution. In general, beamforming provides better connectivity than omnidirectional antennas. Information about the neighborhood can be used to significantly improve the resulting topology as evidenced for MNDB (one-hop information) and TNDB (two-hop information).

shows how RDB improves the connectedness of the network by the occurence of long links. While no isolated nodes occur any more in this example, some of the nodes located on the edge of the network beamform away from the network core. They therefore run the risk of getting disconnected. This is clearly undesired.

We have seen that finding optimal beamforming directions is prohibitive in practice, in particular when realistic antenna patterns have to be considered. Therefore, let us now propose heuristic approaches that aim at providing stronger connectedness of multi-hop networks than RDB.

The intention is to develop beamforming strategies that are significantly improving over RDB in terms of connectivity and robustness of routing, but still require limited state information and are therefore feasible in self-organizing multi-hop networks.

The general idea behind the following strategies is to find suitable beamforming directions using an iterative approach: each node chooses a preliminary beamforming direction, waits for other nodes to adjust their beamforming as well, and then re-adjusts the preliminary beamforming. These iterations are continued until no improvement by re-adjusting the beamforming can be achieved any more.

In each iteration, the beamforming is adjusted on the basis of a connectivity criterion that is computed from aggregate neighborhood information. This neighborhood information is gathered by sweeps of the antenna main lobe. Sweeping the main lobe means that a node steers the main lobe in $n_K$ directions with an angular separation of $\frac{2\pi}{n_K}$. The most beneficial of these $n_K$ beamforming directions is chosen in an iteration, and re-adjusted in the next iteration.

Such iterative schemes involve convergence issues which will be discussed in Sec. 3.4.4.

### Maximum Node Degree Beamforming

The first iterative beamforming approach described here is referred to as *Maximum Node Degree Beamforming* (MNDB). Sec. 3.3 provided an optimal solution to beamforming based on the node degree criterion, assuming full topology knowledge and a simplified antenna model. MNDB is a distributed heuristic aiming at reaching this optimal solution as good as possible.

As opposed to the linear formulation in Sec. 3.3, which was bound to the keyhole antenna model, MNDB can be applied to any kind of antenna setup and beamforming method. In the following, we assume UCA antennas with $m = 10$ elements and gain maximization as described in Sec. 3.1.3.

MNDB is a fully distributed algorithm consisting of only local interaction between nodes. The steps to be carried out by each node $n_i$ can be formalized as in Algorithm 2. With MNDB, the connectivity criterion $\delta_c$ is the node degree, i.e. the number of *one-hop* neighbors.

As can be seen from Fig. 3.11(c), MNDB eliminates the problem of border nodes that beamform away from the network. On the other hand, we can see that MNDB tends to tighten node clusters. While the nodes point their beams such that they form clusters with strong connectivity within the clusters, there are few connections between different clusters.

**Algorithm 2**: Iterative beamforming

1 **initialize** $\gamma_{n_i} = \texttt{random()} \cdot 2\pi$
2 **initialize** $n_K$, $T$
3 **repeat**
4 **initialize** $\delta_{\max} = 0$
5 **initialize** $\gamma^*_{n_i} = \gamma_{n_i}$
/* main lobe sweeping */
6 **for** $k = 0$ **to** $n_K - 1$ **do**
7 Steer main lobe to $\gamma_{n_i} + k \cdot \frac{2\pi}{n_K}$
/* $\delta_c$ = # neighbors within 1 (MNDB) or 2 (TNDB) hops */
8 Maintain this direction for time $\frac{t_{\text{sweep}}}{n_K}$, while determining $\delta_k$
9 **if** $\delta_c > \delta_{\max}$ **then**
10 $\delta_{\max} = \delta_c$
11 $\gamma^*_{n_i} = \gamma_{n_i} + k \cdot \frac{2\pi}{n_K}$
12 **end**
13 **end**
14 Steer main lobe to $\gamma^*_{n_i}$
15 Wait for time $T$
16 **until** $\gamma_{n_i}$ converged

### Two-Hop Node Degree Beamforming

The second distributed approach proposed here is called *Two-hop Node Degree Beamforming* (TNDB). It aims at providing strong connectedness within node clusters *and* between clusters.

Like MNDB, TNDB is a fully distributed algorithm. The steps to be carried out by each node $n_i$ can be formalized the same way as with MNDB by Algorithm 2, except that, in line 8, the connectivity criterion $\delta_c$ is the number of one-hop neighbors *plus* the number of *two-hop* neighbors.

By maximizing the number of neighbors within two hops, nodes are more likely to adjust the main lobe so as to interconnect clusters of nodes. Intuitively, the following effect occurs with Two-hop Node Degree Beamforming (TNDB): If a node is at the border of a cluster with strong internal connectivity, it will remain connected to the cluster through its antenna side lobes, even if the main lobe does not point toward the cluster. As a consequence, the node will only be connected to a few neighbors of the own cluster, but these will in turn provide connectivity to the rest of the cluster. If, at the same time, the main lobe is set in the direction of a remote cluster, the node gains a large number of further two-hop neighbors in the remote cluster, while losing only few neighbors in the own cluster as compared to MNDB. This can be observed in Fig. 3.11(d), where connectivity between clusters is almost as strong as connectivity within clusters.

### Algorithm Details

The connectivity criterion for MNDB in each of the $n_K$ beamforming directions can be obtained either by passive monitoring of traffic, or by neighbor discovery using explicit packet transmissions (*beacons*). A third possibility is to use networking protocols that are anyway in place, e.g. protocols for neighbor discovery or routing. With any of these possibilities, a node has to analyze source addresses contained in the transmitted data, and perform bookkeeping of distinct node identifiers, e.g. MAC addresses. Passive monitoring has the advantage that no wireless resources are used for determining the $\delta_k$, while beacons can be sent frequently irrespective of normal data transmissions, thereby allowing for faster main lobe steering. For TNDB, one-hop neighborhood information must be exchanged between neighbors so as to provide two-hop information. Using explicit beacons is therefore the most practical solution for TNDB.

It is beneficial to keep the ratio of the sweep duration and the time between sweep iterations, $\frac{t_{\text{sweep}}}{T}$, as small as possible. Thereby, negative effects on the stability of routing protocols can be avoided. On the other hand, $T$ should be small enough to follow possible topology changes.

In the analysis of the algorithms, it is assumed that the nodes carry out the beamforming adjustment one after the other, and it is disregarded that two (in particular neighboring) nodes may adjust their beamforming simultaneously. An iteration is finished after all nodes $n_i$ have re-selected $\gamma_{n_i}$. In the next iteration, the nodes act in the same order again.

Such an ordered behavior can be achieved in practice by synchronizing the network in time intervals with duration $T + t_{\text{sweep}}$, with each node using its Identifier (ID) to determine a node-specific offset from the beginning of the interval. While $T$ should be chosen as small as possible to allow for quick convergence of the distributed scheme, it should be large enough to separate the beamforming adjustments of individual nodes in time.

Synchronization has the further advantage that the exchange of beacons can be limited to a short time period within $T + t_{\text{sweep}}$. Thereby, limited additional overhead is introduced even by frequent beaconing within that short time period, which in turn allows for a small $t_{\text{sweep}}$.

Looking at Algorithm 2, the beamforming iterations stop when all directions $\gamma_{n_i}$ have converged. Fortunately, most of the benefits of MNDB and TNDB are attained after only few iterations (cf. Sec. 3.4.3). So, in practice, a node $n_i$ may stop when $\gamma_{n_i}$ remained unchanged for a certain number of iterations, or simply after a predefined number of iterations.

Another technicality exists with respect to choosing an appropriate value for $n_K$. The larger $n_K$, the better the spatial resolution of the main lobe sweep. On the other hand, $t_{\text{sweep}}$ has to be increased with $n_K$ to have enough time to determine the connectivity criterion $\delta_k$ in each direction. The simulations below assume $n_K = 36$ resulting in $10^\circ$ main lobe increments, which may be a suitable value in a practical implementation.

### 3.4.3 Numerical Results and Comparison

The performance analysis of the distributed algorithms covers several aspects. First, we analyze the network connectivity by evaluating the path probability $p_{\mathrm{Path}}$. Second, the availability of node-disjoint paths is analyzed. Third, we study the convergence behavior of the algorithms for static networks. Finally, we look at mobile scenarios, and assess to what extent fixed beamforming can be used when nodes move.

All results in this section are obtained by averaging over $\Omega = 50$ topologies. For each topology, node-disjoint paths are determined for all possible node pairs. Thus, in the case of a scenario with $N = 100$ nodes, the results are based on 247,500 node pairs.

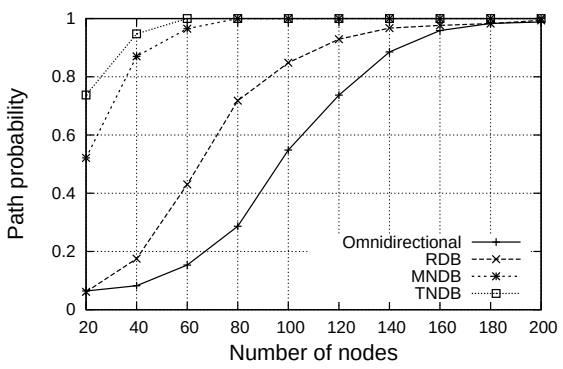
The analysis is carried out both for scenarios with homogeneous node distributions and for clustered topologies. The impact of the node density is studied by varying the number of nodes on a fixed-size system area of $1000\,\mathrm{m} \times 1000\,\mathrm{m}$.

Fig. 3.12 summarizes the results for connectivity and availability of node-disjoint paths. The simple RDB beamforming scheme leads to a better connected network when compared to omnidirectional antennas (Fig. 3.12(a)), as already observed in previous work [BHM05]. In clustered scenarios, the benefit of RDB lies in the fact that connections between clusters occur even without coordination between nodes. This leads to a significant improvement even when the number of nodes in the system area is already high. As expected, MNDB outperforms RDB in terms of connectivity, most importantly when the node density is low. This is mostly because nodes close to the system area border do not beamform away from the network, as is possible in the case of RDB. However, with inhomogeneous node distributions, MNDB tends to form clustered topologies. As a result, the path probability does not improve over the RDB case when the network comprises many nodes (Fig. 3.12(c)). Comparing Fig. 3.12(c) to Fig. 3.12(a), using omnidirectional antennas and RDB, the network is better connected with a clustered distribution than with a homogeneous one. This is due to the fact that, with a low number of nodes homogeneously distributed, a node would rarely find neighbors within its communication range. However, in a clustered topology, a node has a higher chance to connect to neighbors, and occasionally to its chosen destination node. This effect is less pronounced with MNDB or TNDB, since they already show much better connectivity even with low node densities. The best connectivity in all cases can be achieved by exploiting two-hop information using TNDB. It is most advantageous in increasing the path probability in the clustered case.

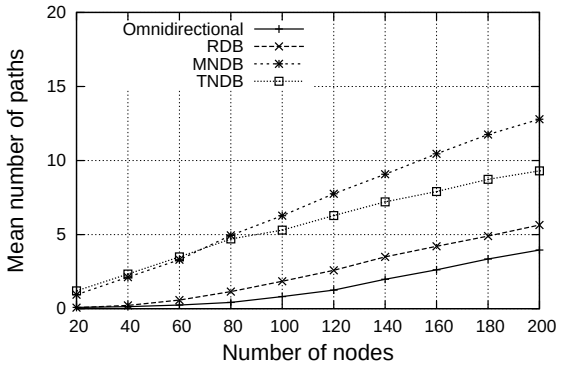
The average number of available node-disjoint paths is significantly higher with MNDB than with RDB (Fig. 3.12(b) and Fig. 3.12(d)). With respect to the number of node-disjoint paths, MNDB is the preferred beamforming scheme when looking at average values. TNDB performs similarly when the node density is low.

Looking at both performance measures, TNDB can be regarded as the best-performing beamforming scheme among the different distributed algorithms.

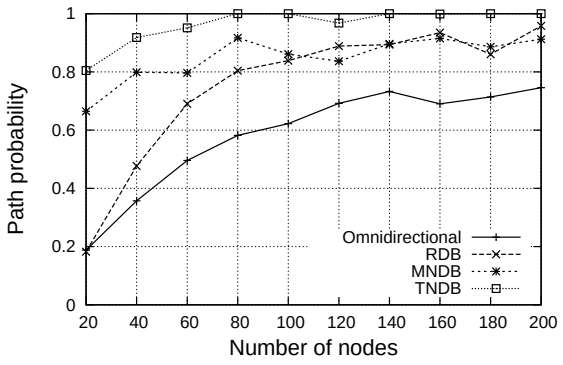
A natural improvement over iterative beam sweeping with the two-hop node degree criterion would be, for a network with $N$ nodes, a $(N-1)$-hop node degree criterion. Such a method would be a distributed heuristic for maximizing the path probability



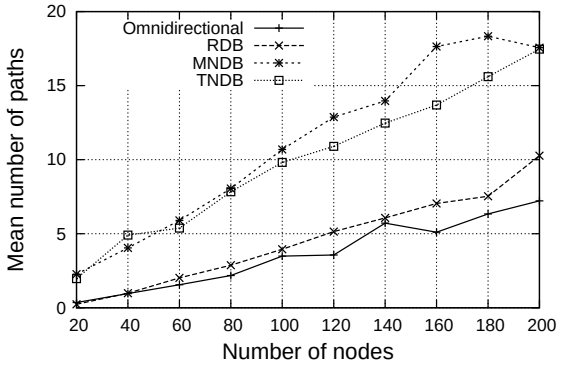

(a) Path probability with homogeneous node distribution.




(b) Mean number of paths with homogeneous node distribution.




(c) Path probability with clustered node distribution.




(d) Mean number of paths with clustered node distribution.

Figure 3.12: Estimated means for the path probability and the number of node-disjoint paths between two randomly chosen nodes on $1000\,\mathrm{m} \times 1000\,\mathrm{m}$.

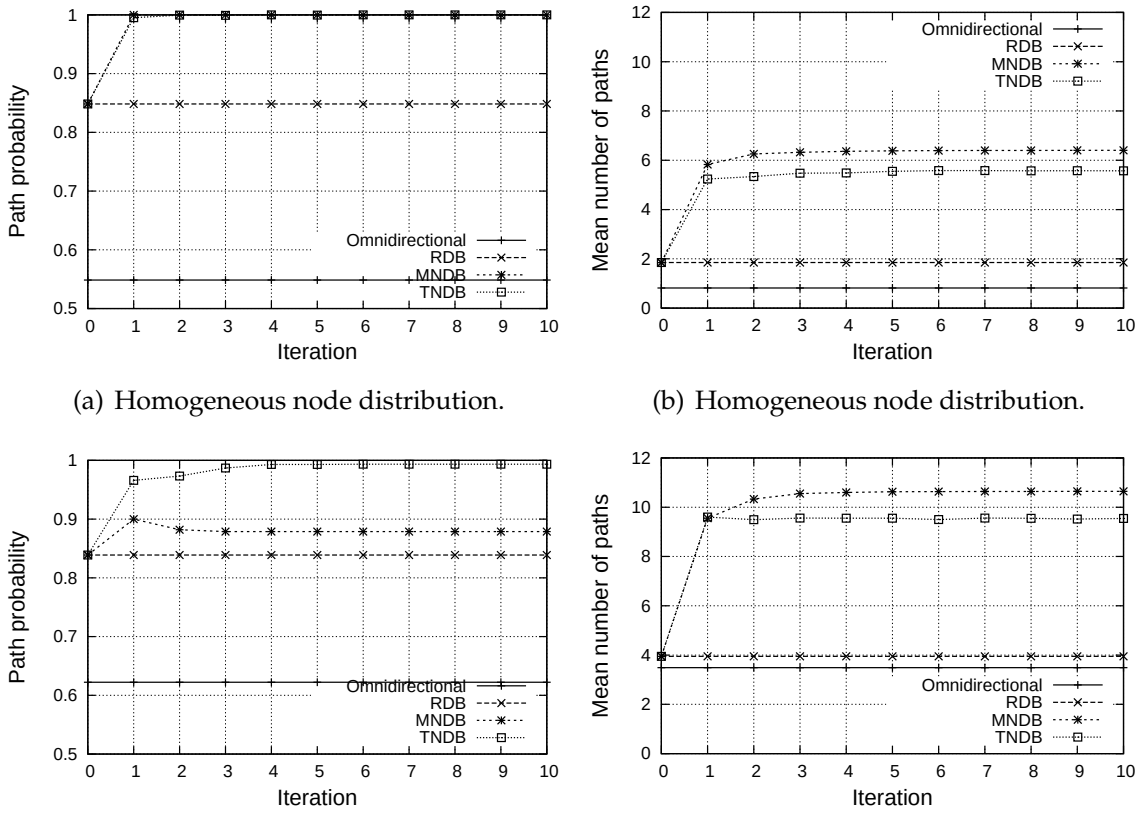
$p_{\mathrm{Path}}$. The inherent overhead for determining the $(N-1)$-hop node degree for each beamforming direction is, however, prohibitive in large networks, and it is questionable whether the performance gap to TNDB would justify the induced message complexity even in small networks.

### 3.4.4 Convergence

A critical aspect of both MNDB and TNDB is convergence. Weak convergence means that the network requires a large number of iterations before beamforming directions $\gamma_{n_i}$ have been found that improve the network topology significantly.

In order to assess the convergence behavior, let us analyze the above-mentioned performance measures after each iteration. Average values are gathered over $\Omega = 50$ topologies with $N = 100$ nodes. The results are shown in Fig. 3.13.

Both MNDB and TNDB show quick convergence, and the performance measures are stable after only few iterations. Moreover, the benefits of these iterative schemes would be largely achieved even if limited to only one iteration. Interestingly, for MNDB




(a) Homogeneous node distribution.

(b) Homogeneous node distribution.

(c) Clustered node distribution.

(d) Clustered node distribution.

Figure 3.13: Path probability and number of node disjoint paths for $N = 100$ nodes on $1000\,\mathrm{m} \times 1000\,\mathrm{m}$ with different location distributions.

with inhomogeneous (clustered) node distributions, further iterations after the first one slightly reduce the path probability and the average number of node disjoint paths. This is due to a tighter clustering after continued re-adjustments of the main lobes toward locations with high node density. In contrast, TNDB profits from further iterations.

As a means for preventing divergence in a practical implementation, threshold parameters for re-adjusting $\gamma_{n_i}$ may be useful: Then, nodes only change $\gamma_{n_i}$ if this would result in an improvement of the connectivity measure $\delta_k$ exceeding some preset threshold, and/or when $\delta_k$ fell below a preset threshold. This would still require regular main lobe sweeps, but the main lobe direction would be reset to the previous direction if the node degree does not improve significantly. The dynamics introduced by main lobe adjustments are thereby limited, and the need for re-routing in the network thus reduced. In an extreme case, nodes may only start over with the iterative algorithm when disconnected from the rest of the network ($\delta = 0$).

### 3.4.5 Applicability to Mobile Scenarios

The previous sections were concerned with static scenarios. A major caveat with persistent beamforming is the question whether such approaches can perform well in mobile scenarios. Intuition suggests that a directional antenna gain leads to failure-prone links when nodes move independently of each other.

This section quantifies the life time of routes, comparing omnidirectional with directional antennas. The same modeling assumptions as in previous sections are assumed. Routing protocols are not regarded in the analysis. It is assumed that all node-disjoint paths between the node pair under consideration have been determined before nodes start to move. For assessing the route life time, all possible node pairs in the scenario are considered, and probability distributions for the route life time are estimated by means of simulations. In the case of directional antennas, MNDB is deployed in this section. Routes break due to changes in attenuation, which is caused by node mobility. The random direction mobility model with bounce-back border behavior is used, the speed of the mobile terminals is $5\,\mathrm{km/h}$. The mobility model is rotation-aware in the sense that the antenna array rotates in systems dimensions as the motion direction of a node changes.

Fig. 3.14 shows the percentage of routes remaining after a given elapsed time as Complementary Cumulative Distribution Function (CCDF). Initially, the CCDF for MNDB is above the one for omnidirectional antennas since MNDB can provide a higher number of paths. Then, areal shearing effects due to the directional antenna characteristics lead to rapid path breaks. Consequently, the distribution of the route life time downs even though the speed of the mobile terminals is relatively low. After about $11\,\mathrm{s}$, a lower percentage of established routes remains in the beamforming case as compared to using omnidirectional antennas.

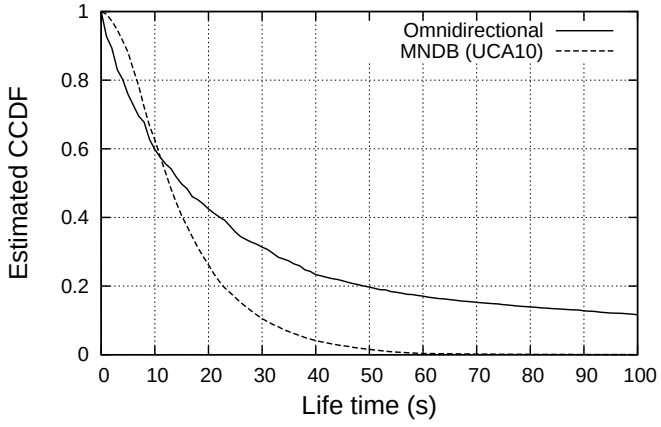



Figure 3.14: Route life time with random direction mobility model for $N = 100$ nodes on $1000\,\mathrm{m} \times 1000\,\mathrm{m}$.

Let us here conclude that persistent beamforming should only be applied in scenarios with low or moderate mobility. As the mobility increases, the beamforming scheme

(MNDB in this case) should be re-performed regularly, even when the scheme has already converged before. With high mobility, omnidirectional antenna patterns lead to more robust routes.

Such a need for adapting networking functionality to the multi-hop network environment is also apparent in routing. As the level of mobility increases, route breaks become more frequent, and re-routing must be performed often. With high mobility, the overhead of routing can become excessive, and simple flooding may be more efficient. Then, the RDB, MNDB and TNDB schemes can in fact again be helpful.

## 3.5 Partition-Based Beamforming

The previous sections showed the importance of connecting clusters of nodes so as to improve the path probability $p_{\mathrm{Path}}$. Persistent beamforming using a two-hop node degree criterion can already lead to better interconnection of clusters, but it is suboptimal. On the other hand, we discussed the optimal solution maximizing the path probability. It inherently improves connectivity, but it is prohibitive in distributed multi-hop networks in practice. This section is devoted to the question of how to explicitly interconnect clusters of nodes by beamforming. We are interested in a completely distributed approach.

Fig. 3.15 shows a sample network with links as obtained by omnidirectional antennas. Obviously, the network graph is partitioned into two subgraphs.

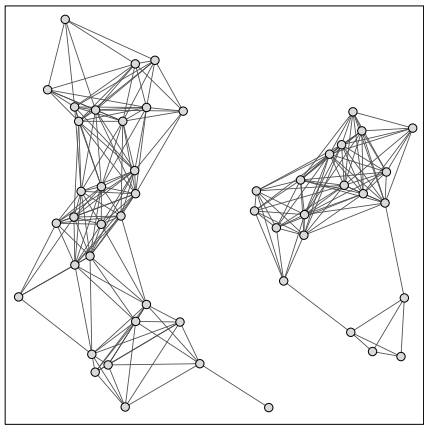


Figure 3.15: A clustered network with $N = 50$ nodes.

If the nodes had information about these two network partitions, nodes at the border of the partitions could beamform toward each other, thereby establishing links between them.

A solution to this could be a similar approach as discussed before: Nodes perform a $360^\circ$ sweep of the main lobe, detect angular directions with disconnected clusters of nodes, and fix the main lobe in these directions. While side lobes could still provide connectivity to the own cluster, long links toward other network partitions could significantly improve $p_{\mathrm{Path}}$.

This section does not go into details of main lobe sweeping and beamforming. Instead, we are interested in much more substantial issues that must be solved before such beamforming can by applied:

> How do nodes detect that the network is disconnected?
>
> How do nodes distinguish nodes belonging to the own partition from nodes belonging to other partitions?
>
> Can partitions be identified in a completely distributed fashion, without knowledge of the entire network graph?
>
> And finally, can these questions be addressed if the network is *not disconnected*, but still shows clusters of nodes with weak interconnection?

The latter is the case in both subgraphs of Fig. 3.15. Within the two subgraphs, there are clusters of nodes with weak interconnection, and in the case of mobility and shadowing these interconnections may break and the network thus get disconnected even further. Beamforming should prevent such disconnections.

Up to now we have used the terms *cluster* and *partition* interchangeably. In the following we will only use *partition* for the sake of clarity. The reason for this is that *clustering* has a quite different meaning in the context of ad hoc networking. There, it is used to describe the introduction of a network hierarchy into an otherwise "flat" network. By the selection of *cluster heads* and the affiliation of the remaining nodes with exactly one cluster head, multi-hop networks can be structured. This structure may then be used to improve the scaling of all kinds of network functionalities, in particular routing. In contrast to this, the present section considers the detection of "clouds" of nodes in space.

Coming back to our questions stated above, it is most of the times not obvious what the partitioning of a multi-hop network should finally look like. Intuitively, disconnected subgraphs of the network graph should belong to different partitions. Partitions may be connected, however, and we then would still want to identify agglomerations of nodes as partitions. This corresponds to the last question stated above.

In essence, there are no "correct" or "wrong" partitionings. We can merely note that "meaningful" partitionings exist which we finally want to achieve. The situation is complicated by the fact that there are no reference nodes in the network with known partition membership. It is even unknown how many partitions exist. For instance, one could partition the network of Fig. 3.15 into two partitions, while a partitioning into e.g. four partitions (cf. Fig. 3.19(b)) is meaningful as well.

The partitioning problem finds a close equivalent in pattern classification. There, it is referred to as *unsupervised learning* or *unsupervised clustering*. By looking at the domain of pattern classification, we can see that this problem is hard to approach even when there is global knowledge about all node positions (*features* in pattern classification), and when there exists a central entity (*classifier*) which defines partition borders in a centralized fashion.

Applying the idea of unsupervised learning to the partitioning of multi-hop networks, we see that this is even more involved. First, there is *per se* no knowledge about

node locations which could be used for partitioning. Second, the clustering has to be determined in a completely distributed fashion, without a central entity and, preferably, only local neighbor-to-neighbor interaction.

As a solution to this complex problem, this section proposes to use *emergent behavior* for the partitioning of multi-hop networks.

Before we clarify the notion of *emergence*, a centralized optimal solution to the partitioning of multi-hop networks is formulated. The methodology is as follows (Fig. 3.16). First, a mesh distance measure and a criterion function for partitioning are defined. Based on these definitions, the partitioning problem can be formulated and solved using an Integer Linear Program (ILP). The criterion function can further be used to evaluate the partitioning results obtained from emergent approaches. Such approaches consist of locally performed algorithms. They are hard to investigate analytically. We therefore simulate such algorithms, and compare the results to the optimal solution. As will be discussed later, emergent algorithms can hardly be derived rigorously for a given problem. We therefore assist the construction of algorithms by randomizing their parameters, and compare the performance of the parameterized algorithms by means of simulation. This will become clearer in Sec. 3.5.2.

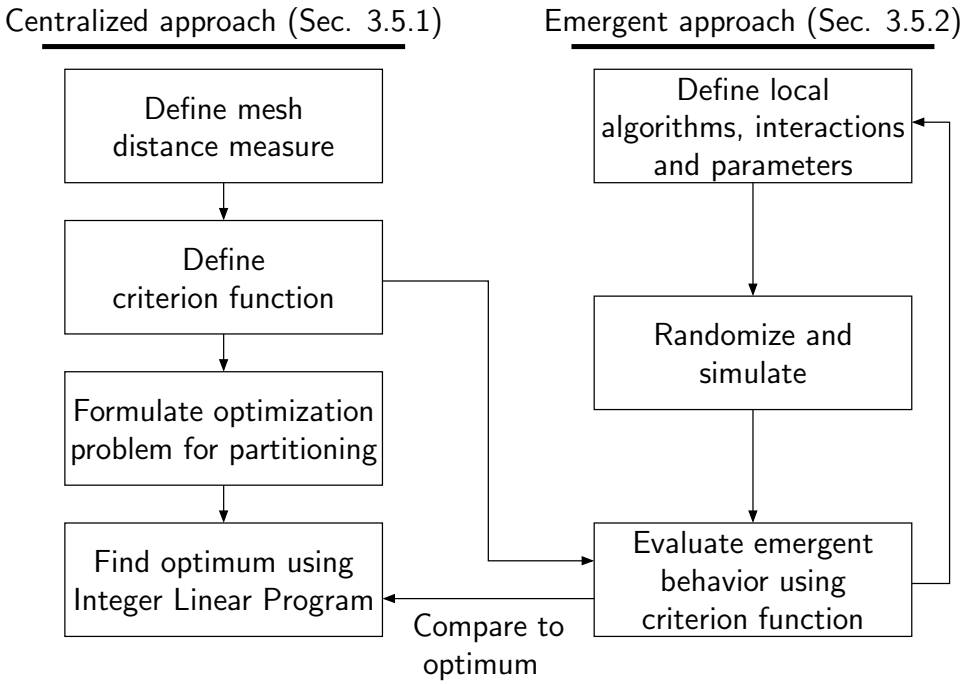



Figure 3.16: Methodology of Sec. 3.5.

### 3.5.1 Centralized Partitioning

The problem of partitioning a multi-hop network has a close counterpart in pattern classification ([DHS01]. There, *clustering* seeks to find similarities between data samples by identifying sample clusters (partitions, in our context) in the feature space. A typical

approach is to first find partition centers, and then assign each sample to one of the centers while minimizing a distance measure between the sample and the center. For certain problems, minimizing the distances between samples and the partition center is equivalent to minimizing the sum distance between the samples of a partition [DHS01]. The latter is the object of this section.

In order to be able to apply pattern classification techniques to wireless networks, we could consider the coordinates of the geographical location of mobile devices as features. But we consider a connectivity-related measure rather than geographical locations, since link states are of higher relevance from a networking perspective, in particular in the presence of shadow fading. Even without shadow fading, meaningless partitionings can result from position-based distance measures when assigning somewhat closely located, but disconnected nodes to the same partition.

### Mesh Distance Measure

An important measure in multi-hop networks is the *hop distance*, defined as the number of edges of the shortest path between two nodes. Generally, we can say that the smaller the hop distance, the higher the performance and reliability of the end-to-end communication between two nodes. Another aspect of communication reliability is the availability of *node-disjoint paths*. In case the shortest path (or, in general, the path currently used for communication) breaks, node-disjoint paths can serve as backup paths.

We define a distance measure $s_{n_i,n_j}$ that reflects how well two nodes are connected in a mesh topology. It is determined both by the hop distance and the number of node-disjoint paths. Since the shortest path is usually used for communication, its length has a higher relevance than the length of disjoint paths. For the set of $N$ nodes, $\mathcal{N}$, we therefore define

$$s_{n_i,n_j} = \begin{cases} \frac{d_{h,n_i,n_j}}{k_{n_i,n_j}} & \forall\ n_i, n_j \in \mathcal{N}, \\ 0 & \text{for} \quad n_i = n_j, \end{cases} \tag{3.39}$$

where $s_{n_i,n_j}$ is called the *mesh distance* between node $n_i$ and $n_j$, $d_{h,n_i,n_j}$ is their hop distance, and $k_{n_i,n_j}$ is the number of node disjoint paths between them. The measure $s_{n_i,n_j}$ is small if two nodes have many short paths connecting them, and it is large for distant nodes with few node disjoint paths. It thus captures aspects relevant to multi-hop networking. In numerical computations, $s_{n_i,n_j}$ is set to $N^2$ in case there is no path between $n_i$ and $n_j$. Fig. 3.17 illustrates four cases. In each case, two nodes $n_1$ and $n_2$ (drawn as boxes) have a mesh distance $s_{n_1,n_2} = 1$.

### Criterion Function

The distance measure defined in the previous section is now used to define a criterion function for a given partitioning $\Pi_p$. A network partitioning $\Pi_p$ partitions a network

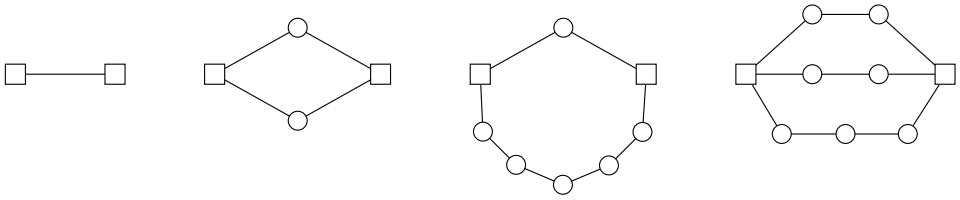


Figure 3.17: Four examples with a mesh distance equal to 1.

comprised of a set of nodes $\mathcal{N}$ into $p$ partitions, $\mathcal{P}_1, \mathcal{P}_2 \ldots \mathcal{P}_p$, with $\mathcal{P}_1 \cup \mathcal{P}_2 \ldots \cup \mathcal{P}_p = \mathcal{N}$. The criterion function is a measure for the quality of $\Pi_p$. The optimum partitioning $\Pi_{p,\mathrm{opt}}$ under the mesh distance measure is the one that leads to minimal mesh distances between nodes belonging to the same partitioning, given $p$. We therefore define the criterion function $J_e$,

$$J_e = \sum_{i=1}^{p} \sum_{(n_j,n_k)\in\mathcal{P}_i{}^2,\, n_j\neq n_k} {s_{n_j,n_k}}^2 \,, \tag{3.40}$$

which is the accumulated square intra-partition distance under the mesh distance measure, summed up over all partitions $\mathcal{P}_i$.

The task is now to find $\Pi_{p,\mathrm{opt}}$ by minimizing the criterion function $J_e$,

$$\min\ J_e \,. \tag{3.41}$$

Later, we will extend this criterion function so as to find an optimal number of partitions, $p_{\mathrm{opt}}$, *en passant* evaluating $J_e$.

### Optimal Partitioning

When partitioning a network with $N$ nodes into $p$ partitions, there are roughly $p^N/p!$ different valid partitionings $\Pi_p$ of the network. This makes an exhaustive search for the partitioning extremizing the criterion function (3.40) computationally complex.

For finding $\Pi_{p,\mathrm{opt}}$ we formulate the problem as (binary) Integer Linear Program (ILP). The resulting model is solved for the sample network of Fig. 3.15 with $N = 50$ nodes using the CPLEX/Concert 9.1 software package.

First, we precompute the distance matrix $\mathbf{S}$,

$$\mathbf{S} = \begin{pmatrix} s_{n_1,n_1} & \cdots & s_{n_1,n_N} \\ \vdots & \ddots & \\ s_{n_N,n_1} & \cdots & s_{n_N,n_N} \end{pmatrix} \,. \tag{3.42}$$

The criterion function for $\Pi_p$ can be calculated by

$$J_e = \sum_{i=1}^{p} J_{e\,(i)}, \tag{3.43}$$

with the per-partition cost functions

$$J_{e\,(i)} = \mathbf{x}_{(i)}^{\mathrm{T}}\ \mathbf{S}\ \mathbf{x}_{(i)}\ . \tag{3.44}$$

Thereby, we use partition membership indicator variables

$$\mathbf{x}_{(i)} \in [0,1]^N\ ,\ i \in [1,\ldots p]\ . \tag{3.45}$$

This means that *iff* $\mathbf{x}_{(i)}(j)$ is set to 1, then node $n_j$ is assigned to partition $i$.
A partitioning is valid *iff* it assigns each node to exactly one of $p$ partitions, resulting in the constraint

$$\sum_{i=1}^{p} \mathbf{x}_{(i)} = \mathbf{1}\,, \tag{3.46}$$

where $\mathbf{1}$ is a vector of ones with dimension $N$.
The expression $\mathbf{x}_{(i)}^{\mathrm{T}}\ \mathbf{S}\ \mathbf{x}_{(i)}$ selects those values from $\mathbf{S}$ which contribute to the per-partition costs of the partition under consideration. It does not have a linear form as required by the ILP, since it is a quadratic expression of $\mathbf{x}_{(i)}$. But we are not computing squares in the first place, but only want to select proper values from $\mathbf{S}$. Therefore, we can express (3.44) in a linear form by concatenating $N$ duplicates of $\mathbf{x}_{(i)}$, and replacing certain duplicates with zero entries (cf. appendix B), thereby generating new solution variables $\overline{\mathbf{x}}_{(i)}$ of dimension $N^2$. In addition, we concatenate the rows of $\mathbf{S}$ into a vector $\mathbf{c}$ of dimension $N^2$. The criterion function is then

$$J_e = \sum_{i=1}^{p} \mathbf{c}^{\mathrm{T}} \cdot \overline{\mathbf{x}}_{(i)}. \tag{3.47}$$

In order to have the $\overline{\mathbf{x}}_{(i)}$ be equivalent to the $\mathbf{x}_{(i)}$, we impose constraints on $\overline{\mathbf{x}}_{(i)}$,

$$\mathbf{A}' \cdot \overline{\mathbf{x}}_{(i)} = \mathbf{0}\ . \tag{3.48a}$$

$$\mathbf{A}'' \cdot \overline{\mathbf{x}}_{(i)} \leq \mathbf{1}\ . \tag{3.48b}$$

The constraint (3.46) can be reformulated for the $\overline{\mathbf{x}}_{(i)}$ as

$$\mathbf{A} \cdot \overline{\mathbf{x}} = \mathbf{1}\,. \tag{3.48c}$$

The matrices $\mathbf{A}$, $\mathbf{A}'$ and $\mathbf{A}''$ are detailed in appendix B. $\overline{\mathbf{x}}$ is the vector containing the concatenated membership indicator variables of all $p$ partitions.
The ILP model is fully described by (3.47), (3.48a), (3.48b), and (3.48c). The resulting $J_e$ obtained from CPLEX is shown in Table 3.3 for different values of $p$. The run time is the time it took an AMD Opteron 2218 workstation (6 GByte RAM) to complete the optimization. For those optimizations where the solver could not decide on the

global optimum within a preset time, the last column of the table shows the remaining worst-case criterion function gap between the continuous and the discrete solution. For the completed optimizations, the optimal solutions was always found well before no gap remained. A visualization of the resulting partitionings $\Pi_{p,\mathrm{opt}}$ is provided in Fig. 3.19 for $p \in [3, \ldots 8]$.

### Decision about the Number of Partitions

By defining the number of partitions, $p$, we add information to the optimization problem that is not available *per se*. If the number of partitions is simply left to the optimizer solving (3.41), it is clear that it will result in the partitioning $\Pi_N$, since then $J_e = 0$. Here, we forbid this trivial solution. We incorporate the issue of finding a reasonable $p_{\mathrm{opt}}$ into the optimization problem by preferring partitionings with a rather small number of partitions. This can be accomplished by defining the optimization problem

$$\min\ p^4 \cdot J_e, \quad p \in \mathbb{N}^+. \tag{3.49}$$

Solving this optimization problem yields an optimal number of partitions, $p_{\mathrm{opt}}$, along with the optimal node partitioning minimizing $J_e$ given $p_{\mathrm{opt}}$, i.e., $\Pi_{(p_{\mathrm{opt}})}$.

Separating a multi-hop network into a higher number of partitions has to result in a significant reduction of $J_e$ in order to make it a preferable solution in the optimization. The consideration of $p$ to the power of $4$ is motivated by the fact that we are considering network partitioning in a two-dimensional space, and a criterion function containing the square of the mesh distances. The values for $p^4 \cdot J_e$ in our sample network are also shown in Table 3.3. As can be seen, (3.49) results in $p_{\mathrm{opt}} = 6$, which is a choice that a human analyzing the topology of the sample network might make as well. Fig. 3.18 depicts the values for $p^4 \cdot J_e$, showing that other reasonable choices, $p = 2$ and $p = 7$, are also preferred solutions of the centralized partitioning approach.

### Alternative Approaches

The centralized solution presented above may appear as being somewhat arbitrary, since – although the methodology is otherwise rigorous – both the mesh distance and the criterion function have been defined somewhat arbitrarily. In particular, one might trade off the hop distance and the number of node disjoint paths differently. Also, instead of the hop distance, one might rather choose a distance measure that takes into account outage capacities or expected path life times. Choosing the accumulated square error criterion upon the distance measure is a similarly arbitrary choice.

In any case, problem-specific criterion definitions are a typical aspect of pattern classification. The above definitions reflect relevant aspects of multi-hop networking, and result in very reasonable network partitions. This makes our centralized approach conceptually consistent with paradigms of wireless multi-hop networks.

But let us go one step back and briefly review alternative approaches to graph partitioning. In [CR93], the authors study the NP-hard problem of partitioning a graph

■ Table 3.3: Optimization results.

| $p$ | Criterion $J_e$ | $p^4 \cdot J_e$ | Run time | Gap |
|---|---|---|---|---|
| 1 | invalid | invalid | – | – |
| 2 | 1478.16 | 23650.6 | $< 1$ s | – |
| 3 | 352.79 | 28576.3 | $< 1$ s | – |
| 4 | 128.35 | 32857.6 | 00:04:16 | – |
| 5 | 45.07 | 28170.1 | 02:02:15 | – |
| 6 | 17.16 | 22238.6 | 00:00:15 | – |
| 7 | 9.51 | 22826.9 | 00:01:40 | – |
| 8 | 6.37 | 26089.1 | 19:24:00 | – |
| 9 | 4.73 | 31036.7 | 03:23:50 | – |
| 10 | 3.75 | 37451.0 | 68:19:48 | 35.2 % |
| 11 | 3.24 | 47363.6 | 12:00:00 | 100.0 % |
| 12 | 2.74 | 56817.7 | 12:00:00 | 100.0 % |
| ⋮ | | | ⋮ | ⋮ |
| 45 | 0.06 | 259490.4 | 12:00:00 | 100.0 % |
| 46 | 0.04 | 190052.8 | 12:00:00 | 100.0 % |
| 47 | 0.03 | 142960.0 | 12:00:00 | 69.7 % |
| 48 | 0.02 | 94372.0 | 5 days | 50.0 % |
| 49 | 0.01 | 51242.6 | 00:21:00 | – |
| 50 | 0 | 0 | – | – |

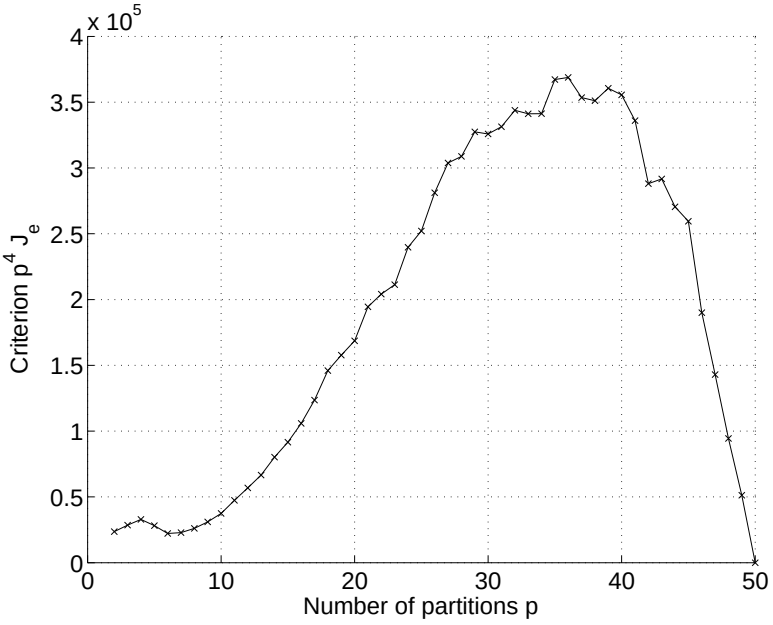
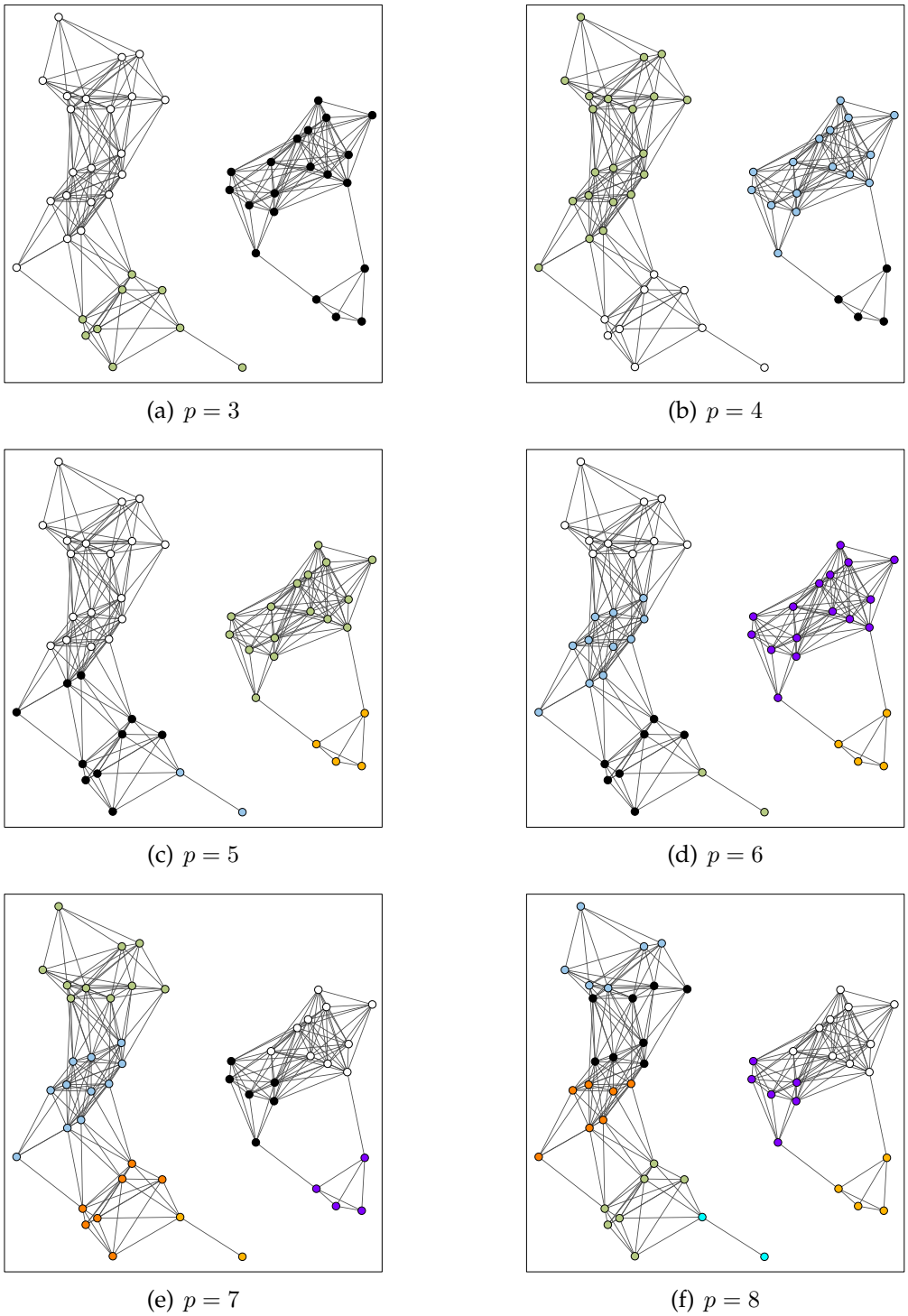



Figure 3.18: Selecting $p_{\mathrm{opt}}$ using a weighted criterion function.



(a) $p = 3$ (b) $p = 4$

(c) $p = 5$ (d) $p = 6$

(e) $p = 7$ (f) $p = 8$

Figure 3.19: Optimal partitionings $\Pi_{p,\mathrm{opt}}$ under the mesh distance measure. For $p = 2$ (not shown), the optimization identifies the two disconnected subgraphs as partitions.

into a predefined number of node sub-sets. The sought-after solution is the one that minimizes the edges cut of the partitioning. The authors describe several forms of this problem and give integer programming formulations for a solution. A similar problem, called *clique partitioning problem*, is subject-matter of [GW89]. Here, the authors look at complete graphs, i.e. graphs where every pair of nodes is connected by an edge. For our partitioning problem, however, problem formulations based on the notion of a minimum cut would lead to quite undesired solutions. In [SKS07], the partitioning problem is defined in the context of communication networks. The authors aim at minimizing the physical distance between nodes belonging to the same partition, and apply different optimization techniques to solve the problem.

From these approaches and the one presented in this section, it should be apparent that the problem of partitioning (wireless) networks is non-trivial, even with full location and link state information, and even when networks are small. Finding optimal solutions is computationally extremely complex.

In the following section, we want to achieve a reasonable partitioning without any sort of information about locations or link state. To make it applicable, a simple, light-weight solution is desired. Light-weight in terms of communication overhead, light-weight in terms of maintained state, and light-weight in terms of computational complexity.

### 3.5.2 Emergent Partitioning

#### Emergent Behavior

Emergent systems are distributed, self-organizing systems in nature, economy, or technology [Joh02]. As in any self-organizing system, local interactions between system entities (mobile devices, in our context) play an important role [PB05]. But what separates emergent self-organizing systems from non-emergent self-organizing systems?

In emergent systems, we clearly distinguish between local operations and interactions on the one hand, and the global (*emergent*) behavior on the other hand, resulting from local operations. The following concepts are of particular importance:

- The local operation is *much* more simplistic than the global behavior.
- Interplay of convergence and divergence.
- Lack of self-awareness: no monitoring and no explicit control loops.
- Exploitation of implicit information.

The first concept is what makes emergent systems so appealing, and it clearly separates emergent systems from traditional self-organizing systems. For instance, ad hoc link-state routing protocols are self-organizing, and necessary information is exchanged on a peer-to-peer basis by local interaction. But the local operation closely follows centralized routing, and the locally performed shortest path computation exactly

reflects the complexity of the routing problem. We would therefore not consider ad hoc link-state routing to be an emergent process.

Emergent systems can be further characterized by describing limitations in terms of design and analysis. Not all of them necessarily apply for an emergent system, but several of them typically do:

- Functionality requires a critical number of entities.
- The global functionality and system properties cannot be inferred from one entity.
- A suitable local interaction cannot be derived rigorously from the desired global behavior. Not only the functionality is emergent, but possibly also its development.
- Slight changes in local operation and local interaction may yield a completely different overall behavior.
- Desiring a slightly different global behavior may require that the local interaction is changed completely.

In summary, an emergent system is not an intelligent system: it is rather simplistic, ignorant. The entities are not "solving" an explicitly given task, they just perform. Emergent systems have several disadvantageous properties in terms of design and analysis. Due to their simplicity and robustness, emergence may still become an increasingly considered paradigm when tackling complex tasks.

### Design Process

We could image all kinds of local interactions that might lead to the desired partitioning of multi-hop networks. As claimed above, a rigorous derivation of well-functioning or even optimal interactions is not possible. The following describes an intuitive approach. We will see that it leads to the desired global behavior in various scenarios.

Let us render the notion of partition membership by state information maintained in the nodes. This state is in the following a three-dimensional vector $\xi \in [0;1]^3 \subset \mathbb{R}^3$. We will decide on partition membership by making a threshold decision $\xi_i \geq 0.5$ on each of the entries of $\xi$. This means that, by rounding the *soft* values $\xi_i$ to binary values, we can distinguish eight partition IDs.

Nodes interact by exchanging this state $\xi$. The hope is that the state of nodes belonging to the same partition converges, and diverges between nodes of different partitions.

Nodes broadcast their state every $T = 1\,\mathrm{s}$ to their one-hop neighbors. In the beginning, the elements of $\xi$ are random in $[0;1]$. Whenever a node receives a state from any neighbor, it merges the received state $\xi_r$ with its own state $\xi$. This represents a convergent force.

In a connected network, the convergent force would lead to equal state information in all nodes. To avoid this obviously undesired partitioning into $p = 1$ partition, a divergent force is necessary. The divergent force is based on auxiliary state information $\widetilde{\xi} \in [0;1]^3 \subset \mathbb{R}^3$. The elements of $\widetilde{\xi}$ are random in $[0;1]$ in the beginning. The local behavior consists of two basic rules:

- When receiving a state information packet, $\widetilde{\xi}_r$ is merged with $\widetilde{\xi}$ the same way as $\xi_r$ is merged with $\xi$.
- Whenever a received state $\xi_r$ is very similar to $\xi$, $\widetilde{\xi}$ is substituted for $\xi$, and $\widetilde{\xi}$ is randomized again.

This approach is inspired by genetics. As nodes interact, their state (*genes*) is mixed. Partitions of nodes with strong interconnection converge toward a common state by frequent interaction, and thereby *evolve* differently from other partitions. In the metaphor of genetics, using $\widetilde{\xi}$ resembles *mutations*. This avoids that weakly connected partitions converge toward the same state in the long run.

Various functions for merging $\xi$ with $\xi_r$ ($\widetilde{\xi}$ and $\widetilde{\xi}_r$, respectively) have been implemented and parameterized. Realizations included simple averaging by calculating $0.5 \cdot (\xi + \xi_r)$, but also various other linear and non-linear expressions of $\xi$ and $\xi_r$. Each of these realizations has been applied to sample multi-hop network scenarios, both with and without mobility. From ten thousands of realizations – randomly constructed in their mathematical details – we picked those that lead to a low $J_e$ (according to (3.40)) as compared to the other realizations. This methodology has already been adumbrated by Fig. 3.16.

A number of realizations was examined by a graphical output of the network, with partition memberships indicated by node colors. Interestingly, quite different realizations lead to similar behavior. Furthermore, somewhat similar realizations showed completely different behavior. This confirms the above statements on deficiencies of emergent behavior with respect to rigorous design. The finally selected realization is detailed in the following.

### Local Rules for Interaction

In the basic algorithm, each node performs parameter initialization at startup, periodically transmits state information $(\xi, \widetilde{\xi})$ using a *state information packet* with period $T = 1\,\mathrm{s}$, and merges its state information with each received state information packet.

Simulations showed that the basic algorithm works well with mobile nodes. In static cases, its divergent forces are too strong. The basic algorithm is therefore complemented by functionality for detecting mobility. The swapping of the state vector $\xi$ with $\widetilde{\xi}$ is made dependent on whether or not mobility is detected.

The rules for detecting mobility should be kept simple. This is achieved by simply detecting changes in the number of packets received during consecutive time periods $T$. Since nodes broadcast state information periodically, this number should not change in static environments. The two variables $v$ and $\widetilde{v}$ are used for this purpose, and mobility is indicated by the Boolean variable $M$. It was found useful to communicate the detection of mobility to nodes in the vicinity. This is done using a time-to-live value $h_{\mathrm{TTL}}$ in the broadcasted packets.

The complete algorithm is described by Algorithm 3–5. The `random` function used therein returns a random number in $[0; 1]$. The lines 19–27 of Algorithm 5 implement the

merging of the node's own state with the state contained in a received state information packet. The lines 5–18 implement the swapping of $\xi$ and $\widetilde{\xi}$ (*mutation*).

As discussed earlier, ten thousands of randomly generated algorithms have been simulated. The one described in Algorithm 3–5 with parameters $c_1 = 1$, $c_2 = 20$, and $c_3 = 7.78719 \cdot 10^{-14}$ showed the smallest criterion $J_e$, summing up $J_e$ over several scenarios and evaluation time instants (in case of mobile scenarios).

**Algorithm 3**: Initialization

1 **initialize** $\xi_i = \texttt{random()}, \quad i = 1, 2, 3$
2 **initialize** $\widetilde{\xi}_i = \texttt{random()}, \quad i = 1, 2, 3$
3 **initialize** $v = 0,\ \widetilde{v} = 0,\ M = \text{false},\ h_{\text{TTL}} = 0$

**Algorithm 4**: Periodic state broadcast

1 **initialize** $M = \text{false}$
2 **if** $|v - \widetilde{v}| > c_1$ **then**
3 $\quad h_{\text{TTL}} = c_2$
4 $\quad M = \text{true}$
5 **else**
6 $\quad h_{\text{TTL}} = \max(h_{\text{TTL}} - 1,\ 0)$
7 **end**
8 $\widetilde{v} = v$
9 **initialize** $v = 0$
10 Transmit packet containing $\xi$, $\widetilde{\xi}$, and $h_{\text{TTL}}$
11 Reset broadcast timer with time $T$

### Behavior in Sample Networks

Let us test the emergent scheme in the static network of Fig. 3.15. This is possible by using an application layer functionality for partitioning that was developed for the simulator (cf. appendix C). In the simulation, it takes 22 s for the scheme to converge to a partitioning, which is then stable. With $T = 1\,\text{s}$, this means that each node has broadcasted 22 packets containing six floating point values and one integer value before a stable partitioning emerged. The partitioning is shown in Fig. 3.20. It comprises $p = 3$ partitions. It is not equal to the optimal partitioning $\Pi_3$ shown in Fig. 3.19(a), but we can still consider the result to be a desired emergent behavior.

It is now interesting to see what the emergent behavior looks like in mobile networks. Let us confront the emergent scheme with two slightly different scenarios.

In a first scenario, $N = 200$ nodes are placed on a square network area of $1250\,\text{m} \times 1250\,\text{m}$. Nodes move with a speed of $10\,\text{km/h}$ according to the random direction mobility model with bounce-back border behavior. This is challenging for partitioning because nodes

**Algorithm 5**: Merging state upon packet reception

Receive packet containing $\xi_r$, $\widetilde{\xi}_r$, and $h_{\mathrm{TTL},r}$
$v = v + 1$
$h_{\mathrm{TTL}} = \max(h_{\mathrm{TTL}},\ h_{\mathrm{TTL},r} - 1)$
$M = (M\ \vee\ (h_{\mathrm{TTL}} > 0))$
**if** $\sum_{i=1}^{3} |\xi_i - \xi_{r,i}|\ <\ c_3$ **then**
  **if** $M$ **then**
    $\xi = \widetilde{\xi}$
  **else**
    **for** $i = 1, 2, 3$ **do**
      **if** $(\xi_i \geq 0.5\ \wedge\ \widetilde{\xi}_i \geq 0.5)\ \vee\ (\xi_i < 0.5\ \wedge\ \widetilde{\xi}_i < 0.5)$ **then**
        $\xi_i = \widetilde{\xi}_i$
      **else**
        $\xi_i = 1 - \widetilde{\xi}_i$
      **end**
    **end**
  **end**
  **initialize** $\widetilde{\xi}_i = \texttt{random()}, \quad i = 1, 2, 3$
**end**
**for** $(y, y_r) = (\xi_1, \xi_{r,1}),\ (\xi_2, \xi_{r,2}),\ (\xi_3, \xi_{r,3}),\ (\widetilde{\xi}_1, \widetilde{\xi}_{r,1}),\ (\widetilde{\xi}_2, \widetilde{\xi}_{r,2}),\ (\widetilde{\xi}_3, \widetilde{\xi}_{r,3})$ **do**
  **if** $y \geq 0.5\ \wedge\ y_r \geq 0.5$ **then**
    $y = 1 - 2\,(1 - y)^2\,(1 - y_r)^2$
  **else if** $y < 0.5\ \wedge\ y_r < 0.5$ **then**
    $y = 2\,y^2\,y_r^2$
  **else**
    $y = (y + y_r)\,/\,2$
  **end**
**end**

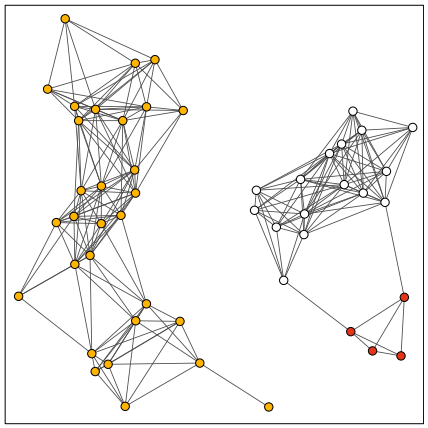


Figure 3.20: Emergent partitioning of the sample network.

do not move in cliques, but completely independently of each other. Furthermore, the chosen mobility model leads to a homogeneous node distribution [Bet03a], where partitions do not clearly stand out against each other. The global behavior over $90\,\mathrm{s}$, illustrated in Fig. 3.21, shows that the rules for local interaction work quite well in this mobile network.

In a second scenario, $N = 200$ nodes are placed on $1000\,\mathrm{m}\times1000\,\mathrm{m}$, and move according to the random direction mobility model with bounce-back border behavior and a speed of $5\,\mathrm{km/h}$. The node density is slightly higher than in the first scenario. This leads to a network that is well connected at almost all times. Identifying weakly connected partitions in this network is challenging, much more than e.g. identifying completely disconnected subgraphs of a network. As we can see from the results in Fig. 3.22, the emergent behavior is still such that partitions are identified in a reasonable manner.

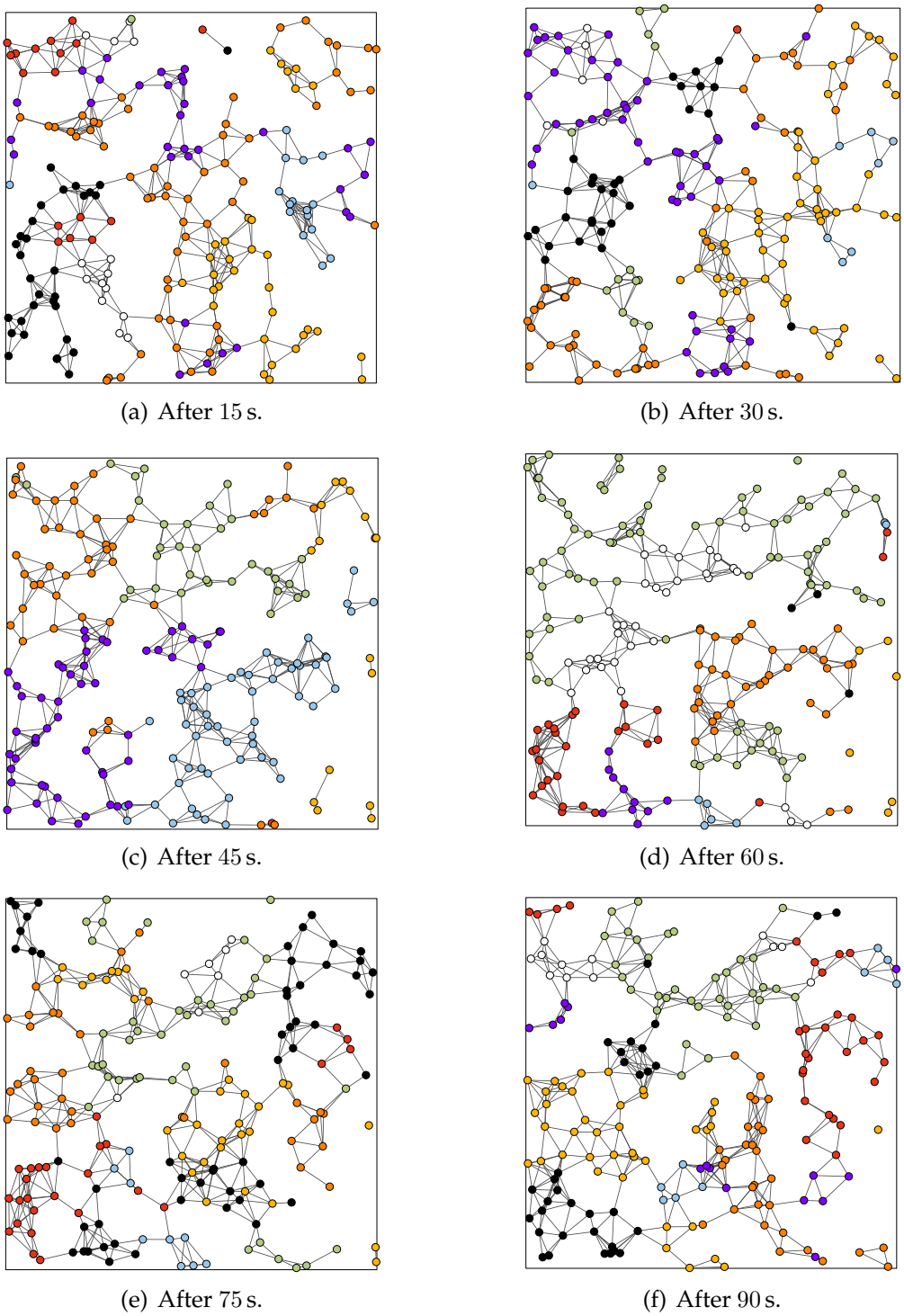
### Complexity Aspects

The presented emergent algorithm can easily be implemented on today's wireless devices – even simple wireless sensor architectures. This has been verified by implementing Algorithm 3–5 in nesC on Crossbow MICAz wireless sensor modules running the TinyOS operating system. With these low-performance modules (128 kB program memory, $8\,\mathrm{MHz}$ processor), three LEDs can be used to display the state vector $\xi$ after the threshold decision $\xi_i \geq 0.5$. With $N = 30$ available hardware modules, the same general behavior was observed as in the simulations.

The emergent partitionings distinguish eight partition IDs, since we chose $\xi$ and $\widetilde{\xi}$ to be three-dimensional vectors. When computing the optimal solution $\Pi_8$ for eight partitions using the centralized approach of Sec. 3.5.1, it took an up-to-date workstation more than 19 hours to accomplish this for a network with $N = 50$ nodes (cf. Tab. 3.3). For the network of Fig. 3.21 and Fig. 3.22 with $N = 200$ nodes, the runtime would be significantly longer. This – in comparison to the MICAz implementation – should well illustrate the complexity advantage of the emergent approach.

Besides the computational complexity, the message complexity is important, too. For instance, in order to achieve the partitionings of Fig. 3.21(a) and Fig. 3.22(a) after $t = 15\,\mathrm{s}$ with $N = 200$ nodes, $15 \cdot 200 = 3000$ packets containing six floating point values and one integer value have to be transmitted.

In a centralized scheme, where node IDs, neighbor lists and possibly node locations have to be gathered at a central entity, and the partitioning results have to be reported back to the individual nodes, we would roughly expect a similar message complexity. However, these communication efforts have to be repeated for each re-partitioning. This is in contrast to the emergent scheme, where the partitioning evolves continuously over time. If the partitioning has to be computed often in mobile scenarios, a centralized optimization may therefore be prohibitive not only due to the computational complexity, but also for message complexity reasons.

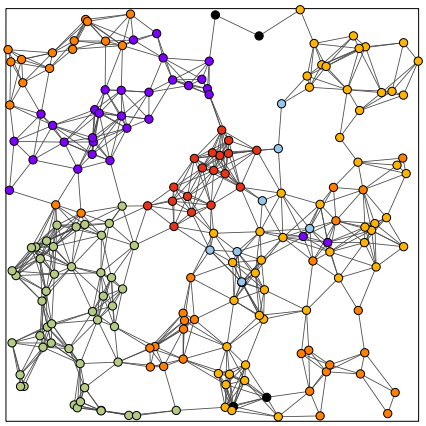


(a) After 15 s. (b) After 30 s.

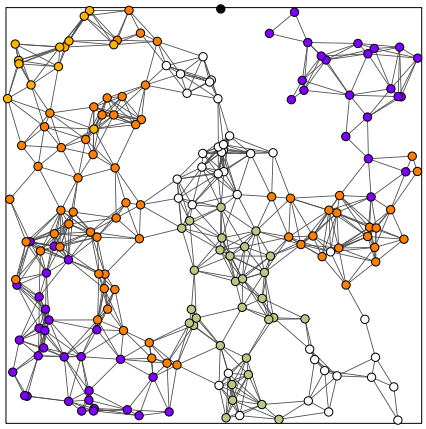
(c) After 45 s. (d) After 60 s.

(e) After 75 s. (f) After 90 s.

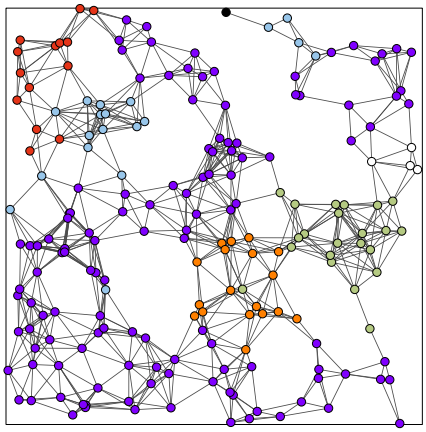
Figure 3.21: Emergent partitioning in a mobile network of size $1250\,\mathrm{m} \times 1250\,\mathrm{m}$ with $N = 200$ nodes.



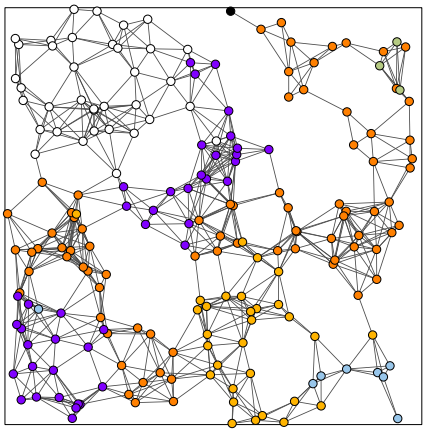
(a) After 15 s.



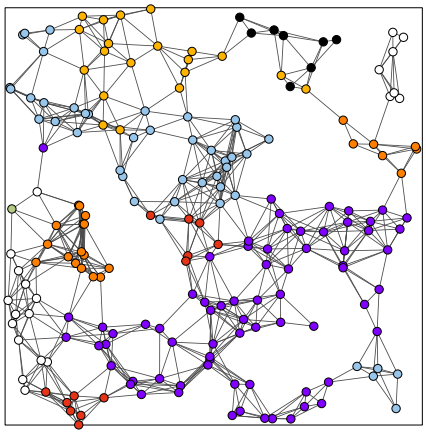
(b) After 30 s.



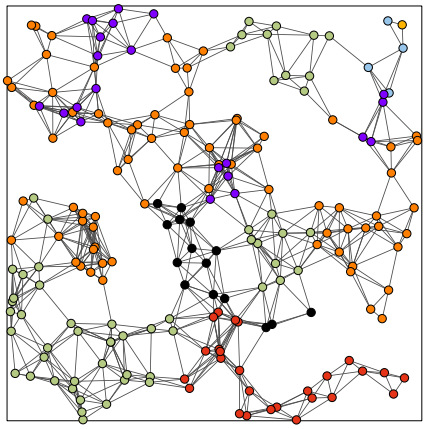
(c) After 45 s.



(d) After 60 s.



(e) After 75 s.



(f) After 90 s.

Figure 3.22: Emergent partitioning in a mobile network of size $1000\,\mathrm{m} \times 1000\,\mathrm{m}$ with $N = 200$ nodes.

### Application to Beamforming

Once partitions have formed, the application of beamforming is straight-forward. For instance, a node can sweep a main lobe in different directions, and exchange hello packets containing the state information $\xi$ in each direction. Then, it can fix the main lobe into the direction where the highest number of nodes belonging to other partitions has been found. Alternatively, a node can choose the direction in which the highest number of different partition IDs has been discovered. Advanced beamforming schemes may form multiple beams, each to a node with a partition ID different from the own partition ID. In a practical realization, further implementation details are with respect to how often the adjustment of beamforming directions are repeated, and whether state information packets are broadcasted omnidirectionally or using beamforming.

The results achieved by emergent partitioning may not be perfect (which we should anyway not aim for in self-organizing systems [PB05]), but appear to be sufficient for the purpose of beamforming. Yet, this section does not intend to provide a numerical analysis of potential connectivity improvements by partition-based beamforming.

## 3.6 Routing-Based Interference Reduction

In this section we explore a persistent beamforming approach intending to reduce interference and increase the spatial reuse of radio resources. The eventual goal is to improve the medium access opportunities of nodes by physical layer decoupling, and thereby increase the throughput performance of the overall network.

The algorithm described in the following can be used without prior beamforming, but also "on top" of the RDB, MNDB, and TNDB algorithms. The algorithm adapts the beamforming pattern to the communication pattern in the neighborhood of a node. This communication pattern is inferred from network layer information, in particular the forwarding table of the routing protocol, and from overheard traffic.

By looking at the local packet forwarding information, a node distinguishes between "desired" neighbors with which it communicates, and "undesired" neighbors. Adaptive nulling is then used to suppress interference from the directions of undesired nodes.

The following questions are addressed in the following:

> Is there room for placing antenna gain nulls in a persistent manner in multi-hop networks, where a node may communicate with neighbors in many different angular directions?
>
> If so, what improvements in terms of interference and signal quality can be achieved?
>
> How do such improvements carry over to the end-to-end throughput performance of the network?
>
> Are there negative effects on the connectivity of a multi-hop network when adaptive nulling is applied in a persistent manner?

### 3.6.1 A Greedy Algorithm for Adaptive Nulling

For a given node, a neighbor is called a *desired* node if the node either transmits to or receives from that neighbor in connection with any of the data flows in the network. All nodes that are not desired nodes are called *undesired*.

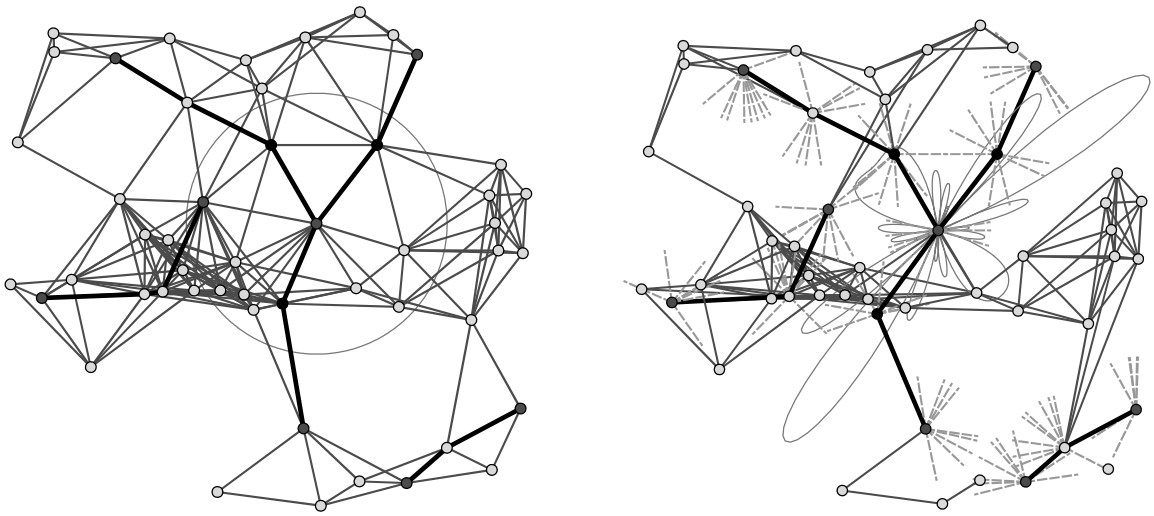
For an individual packet transmission, even a generally desired node is a potential interferer. This is a disadvantage of persistent beamforming, where the antenna patterns are not adapted per packet. On the other hand, the persistent approach to adaptive nulling has the advantage that it can be applied without affecting the MAC layer functionality.

A suitable beamforming method to be used with information about desired and undesired nodes would be one that reduces the antenna gain in the direction of the undesired nodes a much as possible, *while* increasing the antenna gain in the direction of the desired nodes as much as possible, at least above a level that is necessary to maintain the links with the desired nodes.

Such a scheme would raise several questions with respect to design and optimality. The answer to such questions depends on whether the interference management in the network completely relies on interference suppression by adaptive nulling, or whether a MAC protocol is in place on top. The first extreme would clearly aim for high SINR values, and the hope would be that beamforming alone can provide a sufficient SINR for all of the transmissions taking place. The second would rather aim for a small number of interference sources remaining after nulling. This is because each interferer will gain its system bandwidth share in the MAC contention between nodes. It is therefore better for a node to share the wireless medium with one strong interferer (e.g. in time), than getting repeatedly blocked by a plurality of weak interferers. This holds in particular for MAC protocols comprising carrier sensing, where channel access attempts are only made when the channel is sensed free. This is the case for the vast majority of MAC protocols.

In the following, we deploy a heuristic scheme that combines the nulling method of Sec. 3.1.4 (physical layer) with a greedy algorithm based on network layer information. Its intention is to completely suppress the interference of individual interference sources. As a secondary goal, it aims at high SINR values by preferring the suppression of strong interferers over the suppression of weak interferers.

To this end, each node in the network annotates its neighborhood table with information from the physical layer about signal strengths. A node starts with establishing a list of undesired and desired neighbors along with DOA estimation. It then nulls the undesired node with the highest interference power using one Degree of Freedom (DOF) of the antenna array. The remaining DOFs of the array are used to preserve the original antenna pattern as good as possible. The changes may still be significant enough such that links to communication partners are lost, and the corresponding multi-hop routes break. Therefore, the node then checks whether the link budget of all desired neighbors is still satisfactory. If the link budget falls below a given threshold, the node removes the null. Irrespective of whether or not the null is maintained, the node subsequently places a null toward the next highest interferer, and the same procedure is repeated.



(a) Links before nulling.

(b) Links after nulling.

Figure 3.23: Sample scenario with four data flows (corresponding routes indicated by thick lines). The beamforming pattern is exemplified for the node in the center. The nodes on the four routes perform adaptive nulling (dashed lines indicate null directions).

The algorithm ends after nulling all undesired neighbors, or upon reaching the nulling limitations at the node, which depend on the number of antenna elements $m$.

A sample nulling scenario with four multi-hop data flows is shown in Fig. 3.23. While all active nodes – i.e. nodes that act as source, sink, or forwarding node of a flow – carry out the nulling procedure, only the pattern of one node is illustrated for the sake of clarity. After nulling as many undesired nodes as possible, the links toward the desired nodes are still present.

Several factors advocate a re-computation of null directions:

- Changes in the network topology: when nodes move out of a node's neighborhood, and when nodes move in,
- changes in communication patterns: when data flows or connections end and communicating nodes become potential interferers (and vice versa),
- changes in channel conditions: when shadowing and changes in path loss significantly change the signal strength of communicating nodes and interferers.

The frequency of updating null directions depends on the application scenario and must take the above factors into consideration. To allow for an establishment of new routes between nodes where a null exists, it is possible to either periodically send and receive control packets in an omnidirectional manner, or let the nulls time out after a certain period. For now, we consider the case of given traffic relationships between nodes that do not change for the duration for which the beamforming nulls are fixed.

### 3.6.2 Simulative Analysis

For simulations experiments we generate random scenarios in which $N = 50$ nodes are homogeneously placed in an area of $500\,\mathrm{m} \times 500\,\mathrm{m}$. For each scenario multi-hop traffic is generated. All nodes that are generating traffic randomly choose one of the 49 remaining nodes as sink, and determine the shortest path to this sink in terms of the number of hops. We will look at two different setups. In the first setup, 10 nodes out of the 50 nodes are data sources. In the second, all 50 nodes are active and generate traffic, resulting in a scenario with highly meshed traffic relationships. For all results in this section, statistics are averaged over $\Omega = 1000$ random networks, and all contemplable nodes in these networks.

The antenna configuration used in the following is a ULA with $m = 10$ antenna elements. The symmetric characteristic of ULA imposes limits on how arbitrarily the antenna pattern can be shaped. It is therefore interesting to analyze whether traffic relationships occurring in multi-hop networks allow for using such antenna configurations.

#### Analysis of Interference Reduction

With our nulling approach, not all contiguous interference sources can be nulled in arbitrary traffic relationship constellations. Possible gains in spatial reuse are thus not only depending on the nulling capabilities of the adaptive antennas, but also on topology and traffic scenarios. Several questions arise:

> How many desired and undesired neighbors does a node typically have?
>
> Does the presence of desired nodes inhibit null placement, in particular since the limited number of antenna elements of the antenna array do not allow for arbitrary antenna gain shaping?
>
> Can the proposed nulling approach in connection with the chosen beamforming method effectively suppress interference?

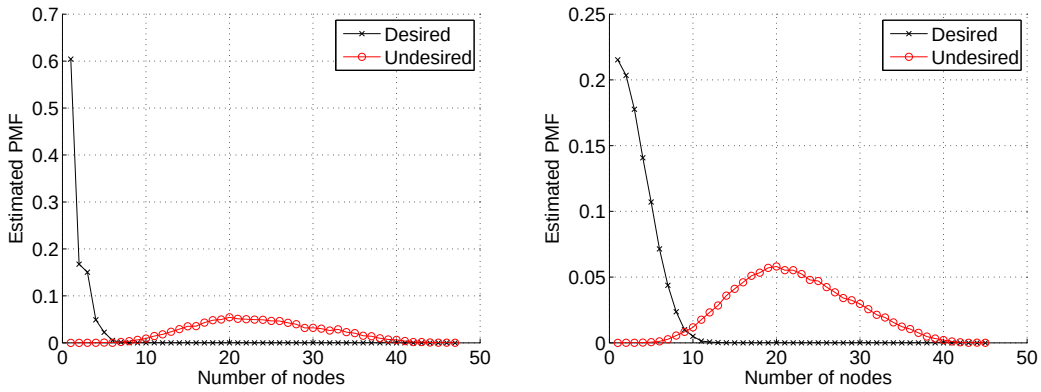
To give an answer to these questions, extensive simulations where performed. Let us begin with an analysis of how many neighbors are desired nodes, and how many neighbors a node wishes to null. The corresponding probability mass functions are shown in Fig. 3.24. They indicate that the number of desired nodes is typically limited to few neighbors, even in the case of 50 data flows. The number of undesired nodes is comparatively large.

It is now interesting the analyze how many nulls a node actually forms under the constraints of preserving desired links and limited degrees of freedom of the antenna array. The probability mass function of the number of placed nulls is shown in Fig. 3.25. Many nodes fully use the DOFs of the ULA10 for nulling. Interestingly, this also holds in the case with 50 data flows, where we would expect that the traffic relationships lead to many desired links and few undesired angular directions.

Each antenna gain null is placed explicitly toward an undesired node. However, we expect that each placed null will on average suppress more than one interferer lying in

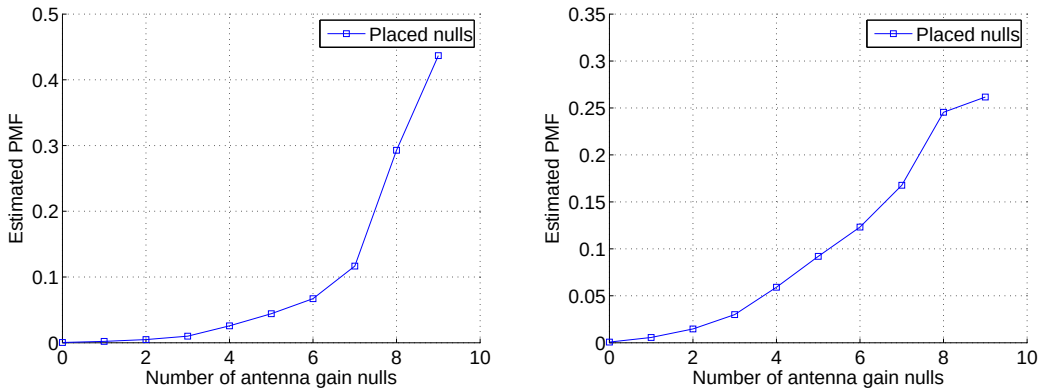
the angular direction of the null. Fig. 3.26 shows the probability mass function of the number of nodes in sensing range before and after nulling. As can be seen, this number is reduced dramatically. From these results we expect a significant reduction of MAC layer blocking by explicit control messages or by carrier sensing, and a corresponding increase in the spatial reuse of radio resources.

We are also interested in the SIR improvement provided by nulling. As can be seen in Fig. 3.27, the worst case SIR benefits significantly from the greedy nulling approach. Worst case means that, when computing the SIR of a transmission, it is assumed that all active nodes simultaneously generate interference, as it would be the case in the absence of any MAC protocol.



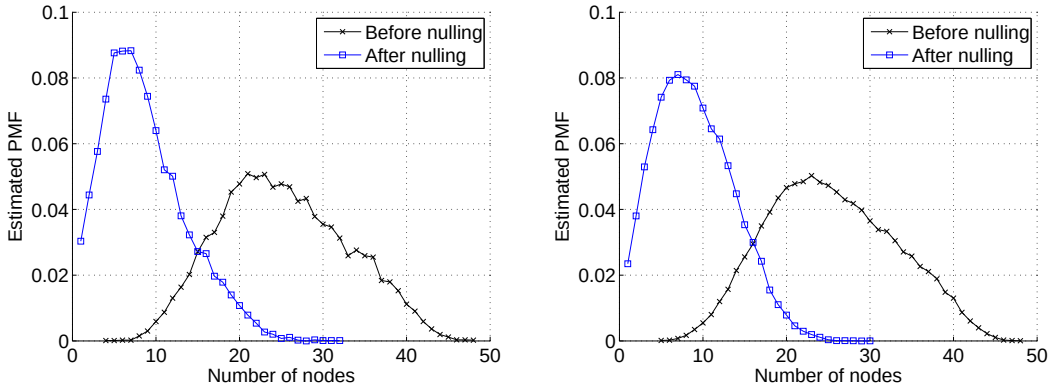

(a) 10 data flows. (b) 50 data flows.

Figure 3.24: Probability mass function of the number of desired and undesired nodes.



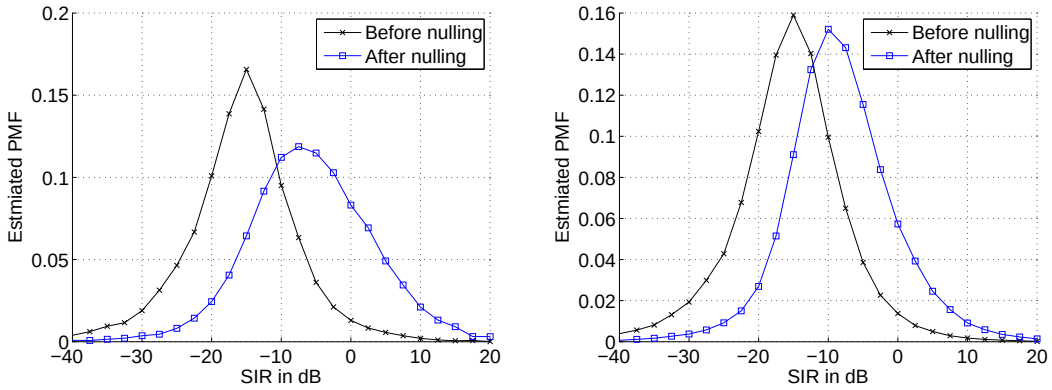

(a) 10 data flows. (b) 50 data flows.

Figure 3.25: Probability mass function of the number of placed nulls.




(a) 10 data flows. (b) 50 data flows.

Figure 3.26: Probability mass function of the number of nodes in sensing range.




(a) 10 data flows. (b) 50 data flows.

Figure 3.27: Probability mass function of the worst-case signal-to-noise ratio.

### Throughput Performance Analysis

After we have estimated possible benefits of adaptive nulling in a persistent approach, it is now interesting to study how these improvements carry over to the end-to-end throughput performance.

For this purpose, we consider the same network scenario with $N = 50$ nodes as before, and vary the number of data flows. Any MAC protocol can be used in connection with persistent beamforming. Here, we use the slotted ALOHA [Abr70, Rob75] protocol.

In the beginning of the simulation, the nodes are randomly placed in the system area. According to the number of data flows, a number of source nodes is randomly selected. Then, shortest paths between these sources and the respective destinations are computed. When persistent nulling is used, then desired and undesired neighbors are determined based on these routes, and nulls are placed accordingly. Finally, the data generation is started.

As performance measure, the end-to-end throughput in terms of successfully transmitted data per flow is computed. The results are shown in Fig. 3.28 for an average packet inter arrival time of 0.5 $s$. Considerable throughput improvements can be observed. The number of sustainable traffic flows is roughly doubled, from about 10 flows with omnidirectional antennas, to about 20 flows with persistent nulling.

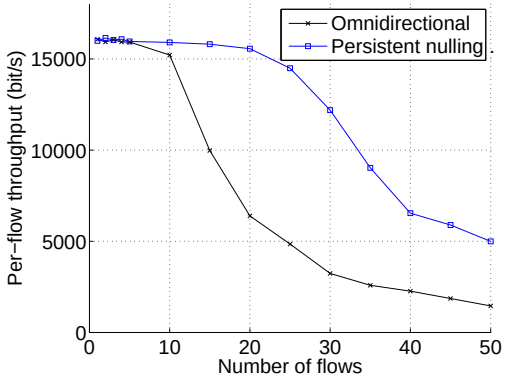



Figure 3.28: End-to-end throughput in the multi-hop scenario.

### 3.6.3 Impact of Nulling on Connectivity

The previous sections showed how placing antenna gain nulls can drastically reduce interference. Thereby, the nulling approach tries to ensure that desired links persist.

Antenna nulls may still prohibit desired links in case of movement or rotation of mobile nodes, changes in communication relationships, and shadow fading.

The involved loss of desired links might induce losses in terms of connectivity, and result in an increase of path lengths. Means to counteract these shortcomings could be:

- Updating antenna nulls by repeating the process of neighborhood exploration and DOA estimation,
- regular or tentative removal of some or all antenna nulls, as a conservative means to avoid long-term disconnection and routing detours,
- analysis of routing protocol control packets: with certain routing protocols, a node may infer that a (control) packet has traversed one of its nulled neighbors; removing the corresponding antenna null can eliminate the obviously existing route detour.

In the following, we will go one step back by asking the following question:

> With our nulling approach, do connectivity and path lengths deteriorate at all in the considered network scenario? If so, how much?

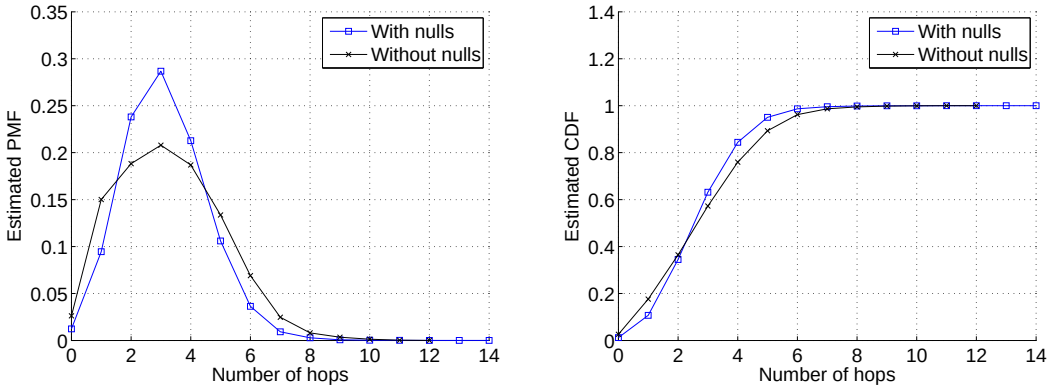
We answer this question by dynamically changing traffic relationships in the network. In particular, upon adapting antenna nulls to desired and undesired nodes in the neighborhood, we let each source reselect its chosen sink. While the prevailing nulls have

been placed in response to the initially chosen sinks and the resulting communication paths, the new sinks now have to be reached via the links that remained after this null placement.

Upon reselecting sinks and performing routing to these sinks, we numerically analyze the resulting path length distribution. We then compare the results to the path length distribution we would have obtained if no nulls would have been placed at all.

In the same network setup as considered above, all $N = 50$ nodes set up a data flow toward a randomly chosen sink, separately. Results are again averages over $\Omega = 1000$ random topologies.

The results for the path length distribution for the reselected sinks are shown in Fig. 3.29. On the horizontal axis, a number of hops equal to zero means that the sink cannot be reached (disconnected). A number of hops equal to one means that the sink is a one-hop neighbor. The curve labeled "Without nulls" represents the path length distribution that would be obtained if the antenna nulls placed upon selecting the initial sinks would be removed before routing to the reselected sinks.




(a) Estimated probability mass function.

(b) Estimated cumulative distribution function.

Figure 3.29: Path length distribution from source to reselected sink.

Surprisingly, it can be observed that nulling does not deteriorate the path length distribution. Less long paths exist when the nulls based on the initially chosen sinks/routes are maintained. This means that, although the antenna nulls have been placed for different traffic relationships, the antenna gain patterns let the new sinks often be reached in a smaller number of hops.

This result can be explained by recalling the results on RDB. Antenna gains exceeding a factor of one can lead to very long links in the network. These long links can provide shortcuts toward the sink which are not available with omnidirectional antennas. The tail of the hop distance distribution gets smaller, although the average node degree slightly decreases. Thus, similar effects as with RDB occur with adaptive nulling, at least when using the nulling method of Sec. 3.1.4.

The results shown here indicate that for many scenarios the proposed nulling scheme

does not inhibit a well-connected multi-hop network, even when nulls are not up-to-date with respect to current communication paths.

## 3.7 Related Work

The idea of Random Direction Beamforming (RDB), first proposed in [BHM05], has been adopted by two research groups [Kos06, ZJDS07, ZDJ07]. Both intend to provide closed-form results on the effect of RDB on the connectivity of multi-hop networks.

The work in [Kos06] takes on an intermediate antenna model, in between an electronically steering UCA antenna and an idealistic keyhole model. This is done by fixing the beamforming direction, and assuming that the resulting antenna pattern simply rotates as the main lobe is steered. The model allows for a curve fitting of the antenna gain distribution of the array. This distribution is then combined with the distribution of node distances (resulting from their random and homogeneous distribution in space), yielding a node degree distribution. With known theorems [Bet04c], the probability of $k$-connectivity is then derived from this node distribution. The author concludes that whether or not beamforming improves the probability of $k$-connectivity depends on the transmission power relative to the receive power threshold. There is a slight discrepancy between the analytical and the simulation-based results. This may be due to the assumption in [Kos06] that networks with RDB resemble pure random graphs with respect to the existence of links. This is, however, not the case. They are neither strictly random, nor purely geometric. This questions the application of graph-theoretical theorems which map between node degrees and ($k$-) connectivity, making an analytical study of RDB difficult.

Connectivity-related aspects of beamforming are also considered in [ZJDS07] and [ZDJ07]. The authors' "greedy" algorithm is somewhat similar to MNDB. It is found that this approach is superior to omnidirectional antennas and RDB in terms of the path probability and the probability of node isolation. The authors further analyzed the impact of the path loss exponent on RDB. They report that, for a path loss exponent below $3$, RDB improves the local connectivity (node degree), while it is degraded for values above $3$. Throughout the present work, a path loss exponent $\alpha = 3$ is assumed.

## 3.8 Summary

This chapter was devoted to using beamforming in a persistent manner in multi-hop networks.

The first contribution of this chapter was a linear formulation for optimizing the steering angles of beamforming antennas, such that the sum of the node degrees of the nodes is maximized. This leads to a significantly better connected network. On the downside, it requires a central entity with full network knowledge, and is computationally complex. Further research is necessary to develop a more efficient optimization methodology, e.g. using column generation [Dan63]. Once optimizing the

node degree can be accomplished more efficiently, the next step would be to optimize the path probability $p_{\mathrm{Path}}$, which is an even more complex task.

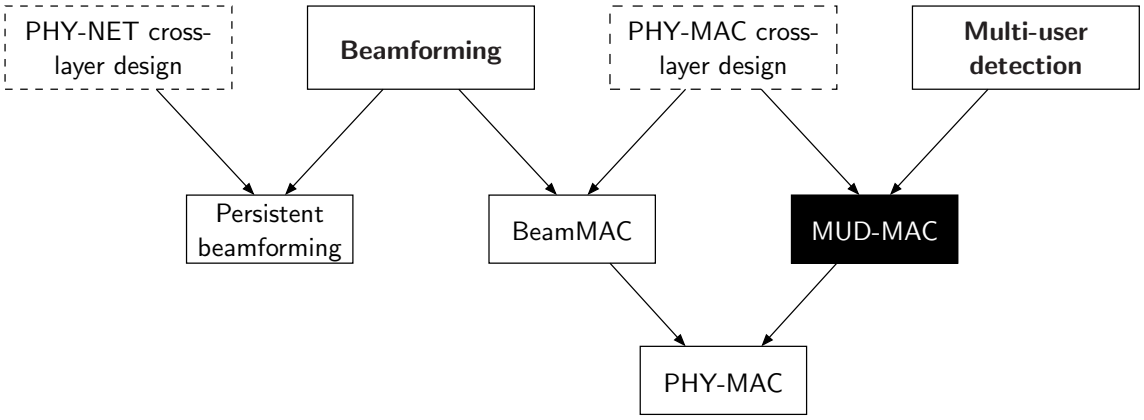
The following contribution was the proposal of schemes for practical realizations. We started with a thorough simulation-based analysis of RDB, which was first proposed in [BHM05]. Then, low-complexity beamforming schemes improving over RDB were proposed and analyzed. These distributed schemes showed quick convergence to stable beamforming directions, and drastic improvements, as compared to omnidirectional antennas, of the path probability $p_{\mathrm{Path}}$ in both homogeneous and clustered node distributions.

As a prerequisite for beamforming based on disconnected or weakly connected subnetworks, we discussed the problem of partitioning multi-hop networks. For baseline comparisons, we first developed a methodology for centralized partitioning, optimized under a mesh distance criterion. Then, we studied emergent behavior as a candidate design paradigm for a distributed approach. Since emergence is a completely open research area, we found it useful to first identify key properties and challenges of emergent behavior. Based on this discussion, an algorithm for emergent partitioning was developed. From the results obtained for different network scenarios, it seems that simple emergent schemes could indeed be used for partition-based beamforming.

Subsequently, this chapter contributed by a persistent beamforming approach for interference suppression. By identifying desired and undesired neighbors using network layer information, nodes contending for medium access can be effectively decoupled using adaptive beamforming. This, in turn, was shown to lead to significant throughput improvements.

Further work on persistent beamforming is necessary to analyze the impact of mobility. As results on route life times of this chapter indicate, beamforming can make multi-hop networks prone to link breaks if nodes move. It would be interesting to investigate this further in different scenarios. For instance, in car-to-car applications, the mobility patterns may be less rugged for persistent beamforming than the random direction mobility model used in this chapter. It would further be interesting to assess the different schemes in multi-path environments, and in scenarios with hilly terrain and non-ideal antenna array orientations. The attempt to formulating closed-form expressions for the impact of beamforming on topology properties of [Kos06] should be explored further. An interesting result of such theoretical work could be optimal beamforming patterns maximizing the path probability in a network with random node distribution. Finally, the work on connectivity improvements on the one hand and throughput improvements on the other hand could be addressed in an integrated approach. Multi-hop networks are usually not connectivity-limited and interference-limited at the same time. With low node density, they are rather connectivity-limited, and interference-limited with high node density. A cross-layer functionality could therefore utilize adaptive beamforming depending on the currently predominant limitation of the network.

# 4 Medium Access and Multi-User Detection




Self-organizing multi-hop networks have no central entity assigning radio resources or organizing medium access. The mutual interference between the devices, also called Multiple Access Interference (MAI), can occur in a quite inordinate way, and is a limiting factor for the transmission capacity and the overall performance of the system.

Distributed MAC protocols are used to restrict this interference. They do so by coordinating exclusive medium access between nodes, typically using time division. In this process, fair medium sharing and efficient use of resources are the most important goals.

For a node receiving data in a multi-hop network, the number of strong interferers is usually small. The number of interferers within sensing range can however be large, with very diverse interference power levels. With most of the MAC protocols, the underlying assumption about data collisions is quite strict. They typically comprise physical carrier sensing, and channel access attempts are only made if the wireless medium is sensed interference-free. This mechanism avoids even small interference levels, resulting in a very pessimistic blocking of nodes. This, in turn, limits the system performance.

The work presented in this chapter is motivated by the idea that interference does not necessarily have to be handled by the MAC layer alone [TNV04]. Instead, we aim at using multi-user detection signal processing on the physical layer to counteract interference. The goal is to achieve higher data throughput in the network by an increased spatial reuse of radio resources.

As we will see, strong physical layer capabilities alone are not sufficient. Instead, the data transmissions in a multi-hop network have to be arranged in a way that is accessible to multi-user detection. Also, without any medium access control, the use of

radio resources may become too aggressive, and the resulting MAI of the network may degrade performance.

The contribution of this chapter is a cross-layer design regarding the MAC layer and the physical layer. It rises to the challenge of "loading" networks with interference in a controlled manner. This is a difficult task when keeping in mind that multi-hop networks are *per se* unordered, distributed systems.

We are particular interested in MUD in spread spectrum systems. Sec. 4.1 summarizes signal processing aspects in this regard. In a cross-layer design with multi-user detection, the primary task of the MAC layer is interference *management*, rather than interference *avoidance*. To develop such a MAC functionality, we identify link layer requirements for deploying MUD in a decentralized and self-organizing network in Sec. 4.2. Based on these requirements, a MAC protocol supporting MUD is proposed in Sec. 4.3. We analytically estimate possible throughput gains of this protocol in Sec. 4.4. Means to reduce the simulation complexity of multi-hop networks with MUD receivers without significantly degrading simulation accuracy are discussed in Sec. 4.5. Network simulations incorporating bit level signal processing are used to further investigate the achievable throughput improvements in Sec. 4.6. Then, Sec. 4.7 considers the case were the devices forming a multi-hop network are equipped with heterogeneous receivers with varying multi-user detection capabilities. Related work on multi-user detection in multi-hop networks is reviewed in Sec. 4.8. Finally, Sec. 4.9 concludes.

The contributions of this chapter have been partly published in [VKHB07, KVM+07, KVM+09].

## 4.1 Interference Cancellation in Multi-Hop Networks

This section reviews multi-user detection in order to provide some insights, and also to provide support for the arguments in the subsequent sections dealing with MAC protocol design.

It furthermore serves as a description of the physical layer module implemented in the simulation tool which was used for the bit level simulations presented later on.

### 4.1.1 Spread Spectrum and the Near-Far Problem

Spread spectrum techniques have long been used in wireless communications. By spreading a signal at the transmitter and despreading the signal at the receiver, the signal becomes more reliable against interference, in particular narrow-band interference. The transmission therefore becomes more robust under low SINR conditions, conversely allowing for higher transmission ranges.

In Code Division Multiple Access (CDMA) schemes [DGNS98], spreading is used to serve several users simultaneously with the same radio resources. Spread spectrum thus improves the system capacity by interference averaging. In cellular mobile networks, CDMA was introduced in 1993 in the USA as multiple access scheme of

the radio interface of the IS-95 standard [Rap02]. A closely related spread spectrum technique is Interleave Division Multiple Access (IDMA), where interleaver sequences are used to separate users [PLWL06].

A major caveat with spread spectrum techniques is the *near-far problem*. This problem occurs when a receiver attempts to despread a signal from a transmitter far away that is interfered by a signal with higher power level, typically from an interferer near by. In cellular CDMA systems, the general level of interference can be upper-bounded by admission control. On top, interference levels are equalized on the uplink (mobile phone to basestation) by fast power control such that the near-far problem is solved.

In contrast to cellular networks, it is difficult to apply power control for the purpose of facilitating spread spectrum in multi-hop networks. In the literature, both centralized and distributed proposals have been made (e.g. [AKKD01, MKR03, EE04, KK05]). Most of them require large signaling overhead, and it is questionable whether the fading characteristics of the channel allow for such schemes in the first place: Assuming that the channel coherence time is on the order of few packets at maximum, then this *a priori* prohibits a multi-hop distribution of control information for power control. Thus, even if neglecting the induced overhead for control information exchange, this exchange may not be able to follow the channel characteristics in time.

Furthermore, power control is fundamentally limited since multi-hop networks provoke near-far constellations that can generally not be resolved by power control. This can be the case even in the simple example of Fig. 2.3 on page 10. Assume that node $n_1$ is close to $n_4$, and $n_2$ is close to $n_3$. With equal transmission power, the power received by $n_2$ from $n_3$ is significantly larger than the power received from $n_1$, so $n_3$ will have to reduce its transmit power level. This, in turn, leads to a small SINR at $n_4$, advocating a decrease of the transmission power of $n_1$ as well. A distributed power control scheme would not converge in this example.

Power control may still be valuable for the purpose of energy conservation or topology control when based on long-term path loss conditions. It seems, however, questionable that power control can break ground for spread spectrum techniques in multi-hop networks with fast fading channels. Thus, many of the power control schemes in the literature should merely be considered as upper performance bounds for distributed approaches.

As a solution to the near-far problem problem, multi-user receivers capable of separating signals have been proposed [RM00, MT01, SG02, HYA03, TNV04, QZY05, ZZA$^+$06]. The existing work on using multi-user detection in multi-hop networks focuses on the improved detection capabilities as compared to conventional single-user detectors. The tendency is to consider simple MAC protocols such as slotted ALOHA [Gal85, RW99, SE02, QZC05]. In the following, it is argued that MUD benefits can only be leveraged by an appropriate MAC scheme.

To this end, it is necessary to provide a careful analysis of the requirements that must be met in order to benefit from multi-user detection capabilities. From this analysis it will become apparent that multi-user detection has to be used in connection with medium access schemes that are tailored to “interference by design”.

### 4.1.2 Multi-User Receivers

Multi-user receivers are designed so as to cope with interfering signals simultaneously transmitted by a multitude of devices [Ver03]. A variety of such receivers exists. They differ in their assumptions about the wireless communication system, their computational complexity, and their performance gap to the optimal Maximum Likelihood receiver. Fig. 4.1 provides an overview of main multi-user receiver categories, and also gives some examples [LR97].

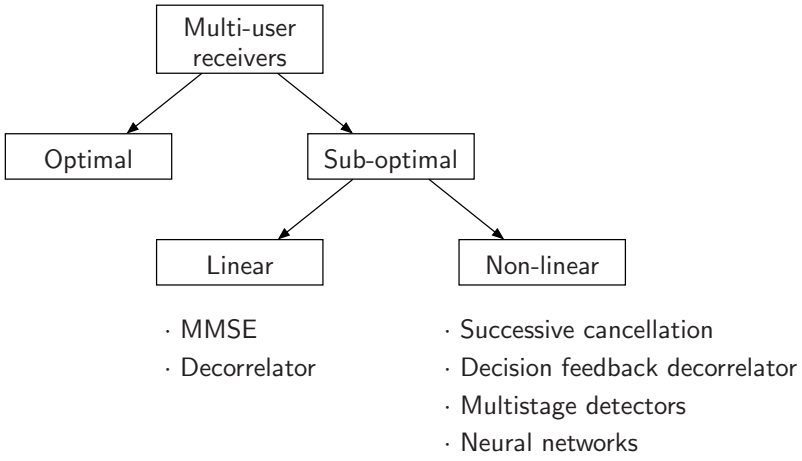



Figure 4.1: Categorization of multi-user receivers for interference cancellation.

In this work, successive interference cancellation is used because of its good tradeoff between complexity (it is significantly less complex than a Maximum Likelihood receiver) and performance.

### 4.1.3 Iterative Interference Cancellation

A multi-user detector can take on the following processing: First, detect the signal component with the highest power level (i.e. the signal coming from the closest transmitter). Then, subtract the estimated signal of the strongest transmitter and detect the second strongest component. Continue the process until all signals, or at least the signal(s) of interest, are detected.

A major drawback of such an approach is that the detection accuracy of the signal component with highest power may be erroneous depending on the noise level and the relative strength of the individual signal components. This error then propagates to the subsequently detected components as well.

Iterative interference cancellation attempts to avoid this drawback. It improves the signal component estimates by interchanging information between these estimates.

Let us consider the following example. Fig. 4.2 illustrates the small network that was already used in Fig. 2.3. Node $n_1$ transmits data using a signature #1 (CDMA code or

IDMA interleaver sequence), and node $n_3$ transmits data using a signature #2. The two transmissions can occur simultaneously without loss of data if the receiving nodes $n_2$ and $n_4$ apply MUD. Looking, by way of example, at the receiver $n_2$, it decodes the received signal $r(t)$ using two decoder branches in its MUD receiver. Only the signal component $\widetilde{s}^{(1)}(t)$, coded with signature #1, is a desired signal. It is the estimate of the signal $s^{(1)}(t)$ transmitted by $n_1$. The component $\widetilde{s}^{(2)}(t)$, which is coded with signature #2 and considered as interference by $n_2$, is also decoded, subtracted from the received signal, and discarded.

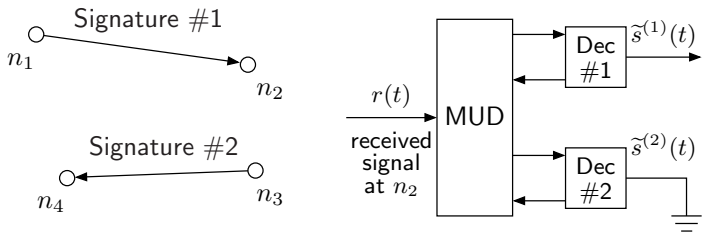



Figure 4.2: Interference cancellation by iterative multi-user detection in a multi-hop network. The interference from node $n_3$ is successively canceled at node $n_2$.

What is essential in iterative (successive) interference cancellation is that the decoders feed back information to the MUD circuitry so as to allow for an iterative processing. The estimates $\widetilde{s}^{(i)}(t)$ are thereby improved *iteratively*.

In general, spreading is necessary to make the process of multi-user detection robust. In principle, any CDMA or IDMA multiple access scheme can be used. This work focuses on IDMA, but the concept can be applied to CDMA systems as well. The transmitter and receiver processing is described in the following, based on the more detailed description provided in [KVM+07, KVM+09].

### Transmitter Structure for IDMA

The transmitter structure of a node $n_k$ is illustrated in Fig. 4.3.

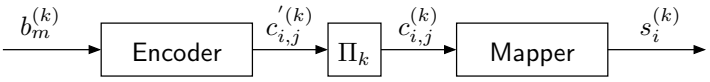



Figure 4.3: Transmitter structure.

The transmitter encodes information bits $b_m^{(k)}$ by any code, e.g. a convolutional code, linear block code, or repetition code. The obtained code word is denoted as $c_{i,j}^{\prime(k)}$.

We assume a transmitter-specific (user-specific) interleaver $\Pi_k$. The output of the interleaver is mapped onto a QAM/PSK signal constellation. This results in complex symbols $s_i^{(k)}$, where $c_{i,j}^{(k)}$ is the $j$-th bit of the symbol $s_i^{(k)}$.

### Channel Model

The channel is modeled as a finite length impulse response filter of order $L$. With channel taps $h_{\ell,i}^{(k)}$ between transmitter $n_k$ and the receiver, and a complex-valued zero-mean Gaussian noise $z_i$, the signal received from $K$ transmitters is

$$r_i = \sum_{k=1}^{K} \sum_{\ell=0}^{L} h_{\ell,i}^{(k)} \, s_{i-\ell}^{(k)} + z_i. \tag{4.1}$$

Note that subsequent sections assume a perfect synchronization between transmitters and receivers, and fast fading is not considered. The channel taps therefore only model the path loss. Nonetheless it has been shown in [KB05] that multi-user detection as applied here also works in realistic asynchronous environments. The corresponding transmission and propagation delays can be formally captured by the $h_{\ell,i}^{(k)}$. Furthermore, the channel taps can be defined to also reflect user-specific channelization codes.

### Receiver Structure and Decoding

The reception with successive interference cancellation can be carried out with a receiver as illustrated in Fig. 4.4.

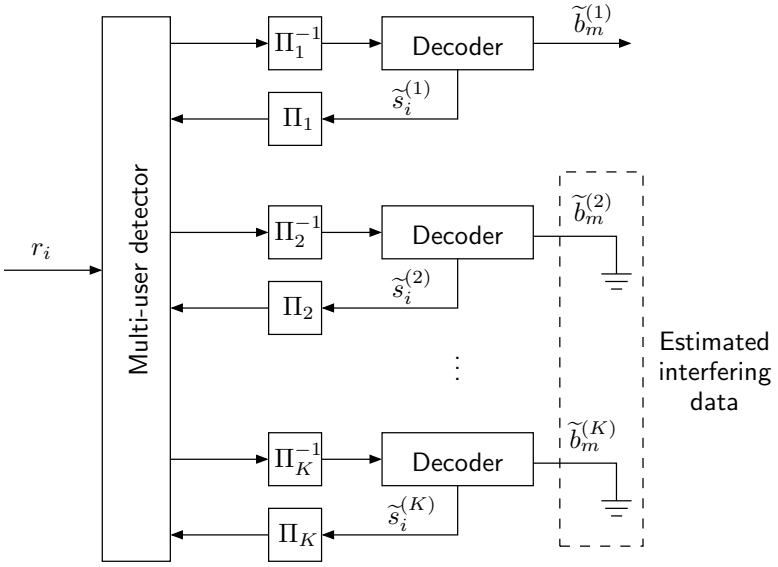



Figure 4.4: Receiver structure.

In order to decode the signal components of $K$ transmitters, the receiver must comprise at least $K$ decoder branches. If the number of decoder branches is larger than $K$, then the remaining branches are left unused. If it is smaller, then the receiver cannot cancel all interferers.

The receiver may decode and pass on more than one data stream $\widetilde{b}_m^{(k)}$, which means that it can intentionally receive data from more than one transmitter at the same time. Without loss of generality, the description of this section assumes that only one data stream is received at a time. Again without loss of generality, we define the signal from transmitter $n_1$ to be the signal of interest, and the signals from the remaining transmitters $n_k$, $k \in [2, \ldots K]$, as interference.

The intention of the receiver's processing is as follows:

- Perform soft interference cancellation $K$ times, i.e. for the desired signal and for $K-1$ interferers,
- repeat this process iteratively in order to improve the soft estimates,
- ideally, when deciding on hard estimates of the received signal, both MAI and Inter-Symbol Interference (ISI) is removed completely, the single-user bound thus reached.

Let the multi-user detector operate on a sliding window basis. Then, for estimating symbol $s_i^{(1)}$, the signals relevant for the window are as follows:

$$\begin{aligned} \boldsymbol{r}_i &= \sum_{k=1}^{K} \sum_{\ell=-L}^{L} \boldsymbol{h}_{\ell,i}^{(k)}\, s_{i-\ell}^{(k)} + \boldsymbol{z}_i \\ &= \boldsymbol{h}_{0,i}^{(1)} s_i^{(1)} + \underbrace{\sum_{\ell=-L,\ell\neq 0}^{L} \boldsymbol{h}_{\ell,i}^{(1)}\, s_{i-\ell}^{(1)}}_{\text{ISI}} + \underbrace{\sum_{k=2}^{K} \sum_{\ell=-L}^{L} \boldsymbol{h}_{\ell,i}^{(k)}\, s_{i-\ell}^{(k)}}_{\text{MAI}} + \boldsymbol{z}_i, \end{aligned} \tag{4.2}$$

with $\boldsymbol{r}_i = [r_i, \ldots, r_{i+L}]^{\mathrm{T}}$ and $\boldsymbol{z}_i = [z_i, \ldots, z_{i+L}]^{\mathrm{T}}$.

Assume that we have *estimates* for the signal components $s_i^{(k)}$, denoted as $\tilde{s}_i^{(k)}$. With these estimates – obtained from the previous iteration – and the received signal $\boldsymbol{r}_i$, we can compute

$$\tilde{\boldsymbol{r}}_i^{(1)} = \boldsymbol{r}_i - \sum_{k=1}^{K} \sum_{\ell=-L}^{L} \boldsymbol{h}_{\ell,i}^{(k)}\, \tilde{s}_{i-\ell}^{(k)} + \boldsymbol{h}_{0,i}^{(1)}\, \tilde{s}_i^{(1)}. \tag{4.3}$$

The estimates $\tilde{s}_i^{(k)}$ are soft bits, i.e. they are not finally decided to be $+1$ or $-1$, but rather have an accompanying certainty. The processing expressed in (4.3) is therefore called *soft interference cancellation*. For the first iteration, the $\tilde{s}_i^{(k)}$ are set to zero.

Assuming perfect estimates $\tilde{s}_i^{(k)} = s_i^{(k)}$, both MAI and ISI are completely removed, and the receive signal component

$$\tilde{\boldsymbol{r}}_i^{(1)} = \boldsymbol{h}_{0,i}^{(1)}\, s_i^{(1)} + \boldsymbol{z}_i \tag{4.4}$$

is isolated as if we were considering a single-user system. It is important to note that (4.3) requires the channel taps $h_{\ell,i}^{(k)}$ to be known. Proper channel estimation with respect to all $K$ transmitters is therefore critical.

Obtaining *estimates* on the transmitted information bits $b_m^{(1)}$ of the desired user $n_1$, denoted as $\tilde{b}_m^{(1)}$ in Fig. 4.4, requires the multi-user detector to compute log-likelihood ratios for the $c_{i,j}^{(k)}$. From these log-likelihood ratios, soft symbol estimates for the $\tilde{s}_i^{(k)}$ can be derived.

The decoders compute log-likelihood ratios for the deinterleaved $c_{i,j}^{(k)}$. This computation depends on the code in use. Finally, the decoder also computes log-likelihood ratios for the information bits $b_m^{(k)}$. Since we are only interested in the data symbols sent by transmitter $n_1$, it is sufficient to compute only those log-likelihood ratios resulting in the $\tilde{b}_m^{(1)}$.

A detailed description of the steps carried out by the multi-user detector and the decoders for IDMA and CDMA for different coding schemes can be found in [BCJR74, HOP96, BG96, DP97, RHV97, WP99, MP02, SH04b, KB06c, KB06b, KDUB07, KVM+07].

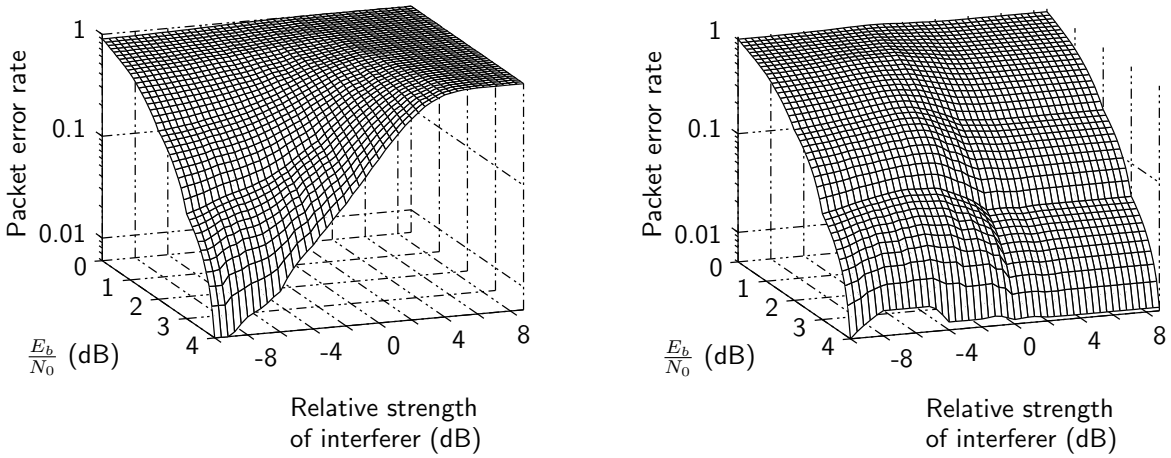
### Performance Aspects

It is up to the physical layer design to specify a number of iterations that results in a good tradeoff between computational efforts on the one hand, and the residual performance gap to the single-user bound on the other hand. To this end, bit level simulations for various channel conditions and user constellations (including different near-far factors) can be carried out, which then reveal the bit error rate progress over iterations.

Results of such simulations for a scenario with one interferer are shown in Fig. 4.5. In this diagram, the Packet Error Rate (PER) is drawn over the Signal to Noise Ratio (SNR) per bit $\frac{E_b}{N_0}$ of the desired signal, and the strength of the interferer relative to the desired signal. For instance, a relative strength of $0$ dB means that the desired signal and the interference signal are equally strong.

Fig. 4.5(a) shows the PER before the first MUD iteration. It is equivalent to a single-user detector, and shows poor performance unless the relative strength of the interferer is low. Fig. 4.5(b) shows the performance after four iterations. Interestingly, with a relative strength of the interferer above about $2$ dB, the PER is primarily limited by $\frac{E_b}{N_0}$, similar to the case of very weak interference (relative strength of $-10$ dB and below). With strong interference, the interference signal can be decoded reliably, and well separated from the desired signal in the MUD processing. We also note that there is a range roughly between $-10$ dB and $2$ dB of relative signal strength, where the PER deteriorates due to interference, and where four iterations are not sufficient to completely reach the single-user bound. The performance improvements in terms of the PER as compared to a single-user detector are drastic, still.

The system considered in Fig. 4.5 is an IDMA system with a rate $1/2$ memory $4$ standard non-recursive convolutional code followed by a rate $1/4$ repetition code, and QPSK modulation. The data block size is 128 bit. The results in [KVM+09] show that the interference cancellation of MUD can be less effective in systems where the spectral efficiency is higher.




(a) Before first iteration.

(b) After four iterations.

Figure 4.5: PER performance example of IDMA [KVM+09].

Such aspects can be considered in a cross-layer design of the physical layer and the MAC layer. In particular, the interference management may handle interference more strictly as the spectral efficiency increases. The cross-layer approach proposed in this chapter can be extended in this regard. However, the following focuses on an IDMA system with parameters as summarized above, and assumes that known interference can always be well canceled, as suggested by Fig. 4.5(b).

### Complexity Aspects

If the structure of the received signal $r_1(t)$ is unknown, the multi-user receiver may simply apply one decoder branch for each signature that may possibly be used in the network. However, it can be shown that in case $r_1(t)$ does not comprise a signal component with a signature that is in fact applied by the multi-user detector (e.g. a signature #3 in the example of Fig. 4.2), then this has a deteriorating effect on the estimation quality of the MUD output. Furthermore, the number of decoder branches that has to be operative at a time – along with the number of iterations performed before deciding on hard output symbols – determines the computational complexity of the MUD receiver. Finally, each of the decoder branches has to synchronize with the respective signal component. The state-of-the-art solution to this is to use a pilot sequence of synchronization symbols. Having no information about the structure of the received signal introduces great complexity in this respect. This is because the receiver then has to search for this sequence continuously, whereat a multitude of sequences may even occur simultaneously when several interferers are present.

In summary, the MUD receiver should be confronted with interference in a structured way, avoiding such uncertainty. This will be a general lead in the MAC protocol design of Sec. 4.3.

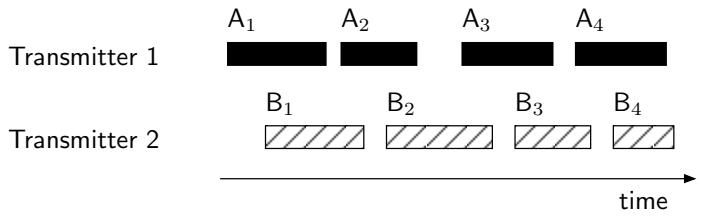
There are further interesting aspects with respect to complexity. As stated above, the complexity of the multi-user detector is determined by the number of parallel decoder branches and the number of performed decoding iterations. It is generally accepted that in multi-hop networks usually only few strong interferers have to be canceled [AWH07]. This is good on the one hand since it means that few decoder branches, i.e. MUD receivers with limited complexity, can already reduce the interference significantly and get close to single-user conditions. On the other hand, this means from a system perspective that the achievable throughput improvements over single-user detectors with strict medium access control are limited as well. At least it seems that tremendous performance improvements as reported in some literature focusing on the signal processing aspects of multi-user detection alone – neglecting network aspects – should be considered with care.

## 4.2 Requirements for Successive Interference Cancellation

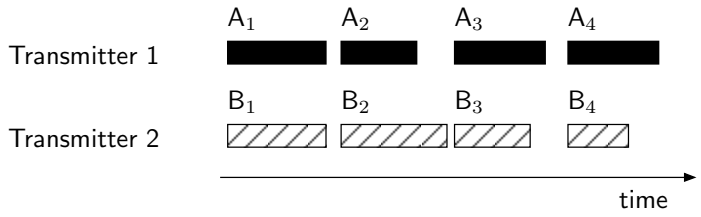
This section summarizes requirements that must be provided so that multi-user detection can be applied in multi-hop networks. The aspects of this section shall be considered in MAC layer design, and it therefore serves as connecting part between the review of multi-user detection in the previous section and the subsequent sections concerned with medium access.

The first major requirement for multi-user detection is time synchronization of the nodes. Since the multi-user detection requires block processing on a frame basis, without time synchronization, the decoding delay in packet reception may not be guaranteed to be limited. Fig. 4.6(a) illustrates an asynchronous case where two transmitters (e.g. nodes $n_1$ and $n_3$ in the example of Fig. 4.2) transmit data packets. For decoding packet $A_1$, an MUD capable receiver (node $n_2$ in the example) also decodes packet $B_1$. For decoding packet $B_1$, it also decodes packet $A_2$, *et cetera*. If the transmissions by the two transmitters never get decoupled in time while one wishes to fully exploit the MUD capability, the decoding delay and the required decoder memory size become infinite. Fig. 4.6(b) shows a synchronous case. For decoding packet $A_1$, only packet $B_1$ has to be decoded jointly. The decoding delay for packet $A_1$ is bounded to one data block length.

From a networking perspective, we have to ask how exact this synchronization must be. From the signal processing point of view, bit level synchronization – or even chip level synchronization – is not necessary for CDMA type transmission schemes, including IDMA. For instance, the Universal Mobile Telecommunications System (UMTS) uplink does also not require perfect synchronization. In [KB05] it has been shown that the detection of asynchronous transmissions with IDMA is possible. For our multi-hop scenario, we thus only require frame level synchronization. While methods for node synchronization are not discussed here, let us note that the accuracy of synchronization may be considered as a design parameter, trading off synchronization efforts with the




(a) Asynchronous.




(b) Synchronous.

Figure 4.6: Asynchronous and synchronous transmissions.

guard interval duration between frames. Further, we note that synchronization only has to be local: A synchronization drift across the network is harmless as long as the frames concurrently decoded by one MUD receiver are reasonably synchronized.

Another requirement for multi-user detection is signature knowledge at receiving nodes. For each signal that shall be considered by an MUD receiver, the receiver has to know the respective signature, i.e. the spreading code for CDMA or the interleaver sequence for IDMA. The receiver also has to know which interfering signals are actually contained in the received signal, i.e. it should not apply a signature which does not have a corresponding interference component. Consequently, a MAC protocol for multi-user detection should be able to exchange signature information for all signals that have to be considered by the MUD receiver.

Another aspect of the required knowledge about interfering sources is channel estimation. The estimated channels are used in the multi-user detector in order to estimate and subtract the contribution of each interfering signal to the overall received signal.

Yet another requirement is given by the limited number of MUD receiver branches. With $K$ decoder branches, a receiver can decode the desired signal along with $K-1$ interfering signals. This is a hardware/software limitation, irrespective of channel capacity limitations of the access scheme given by the spreading gain. A MAC protocol for multi-user detection should thus limit the MUD burden to $K-1$ strong interferers. For the special case of $K=1$, i.e. a receiver without MUD capability (a "weak" receiver), the MAC protocol should provide interference-free transmission opportunities, just as conventional MAC protocols. If weak receivers have to be protected, we further conclude that the MAC layer control information exchange must not demand multi-user detection. Instead, a common channel must be provided that is accessible without MUD capability.

In summary, the following requirements for MUD – denoted as **(R1)** to **(R6)** – influence the design of an MUD aware MAC protocol:

- Time synchronization on a frame level **(R1)**,
- knowledge about signatures in use **(R2)**,
- channel estimation with respect to all signal sources **(R3)**,
- knowledge about the start time and duration of packets **(R4)**,
- receiver-individual avoidance of interference beyond $K - 1$ sources **(R5)**, and
- protection of weak receivers having only single-user detectors **(R6)**.

## 4.3 MAC Protocol Design

### 4.3.1 Design Choices

Various mechanisms and protocols for medium access control have been developed for wireless communication systems. They can be roughly distinguished by the following characteristics [GL00].

**Centralized vs. distributed** With *centralized* MAC protocols, medium access is coordinated by a central control unit. In distributed multi-hop networks without central control units, such schemes are only possible if clustering and local master node selection mechanisms are in place. In general, this is a rarely followed design choice in the research community, mostly because of the inherent overhead and the limitations in mobile scenarios. Consequently, a distributed MAC solution shall be designed in this work.

**Scheduled vs. random access** The nature of multi-hop networks, and in particular ad hoc networks, suggests to share wireless resources in the time domain. That is to say that, since the nodes usually do not have data to transmit continuously and there is no clear distinction between "uplink" and "downlink" directions as in cellular systems, a multiplexing concept purely based on frequency division is inefficient. The division in time is based on either *scheduling* or *random access*. Distributed MAC protocols are generally based on random access (also referred to as *contention-based* access). Centralized approaches allow for both types of time division. In this work, a distributed random access scheme shall be designed.

**Transmission initiation** A data packet transmission may be initiated by a transmitter or by a receiver. In *receiver initiated* approaches, receivers should have information about the communication needs of other nodes, otherwise the protocols become inefficient. Most protocols for ad hoc and multi-hop networks assume *transmitter initiation* [JLB04], mainly because this is more flexible in terms of traffic relationships. In scenarios with highly saturated traffic, where each node has data to transmit at all times, this problem

does not exist. Since this assumption is not made in the following, a transmitter initiated design is preferred.

**User separation and duplexing** For user separation and duplexing, the channel can be shared in the time, frequency, space, and/or code domain. For the sake of simplicity, a single-channel system (one frequency band) is assumed throughput this work. Spreading with user-specific signatures along with multi-user detection is assumed in this chapter. Data transmissions that cannot be resolved by MUD are arbitrated in time. Duplexing between two communicating peers is performed in time-division as well.

**Signature assignment** For CDMA and IDMA, the spreading signatures must be assigned in a regular way. If each transmitter is assigned a fixed spreading signature which the transmitter uses for all transmissions, then the signatures are called *transmitter-based*. Alternatively, signatures can be *receiver-based*, *pair-wise*, or *per-transmission*. In this work, transmitter-based signatures are used. It has been shown in [KB06a] that IDMA interleaver sequences can be effectively derived from a node's unique ID.

Since this work adopts random access, means for avoiding collisions and – because data collisions can never be avoided completely – collision resolution must be found. This is a peculiarity of wireless communications, and primarily due to the half-duplexing nature of wireless transceivers. Mechanisms in this regard are

- *carrier sensing* for passive collision avoidance,
- active *collision avoidance* (either by out-of-band signaling such as busy tones used e.g. in [HD02], or by control handshakes such as the Request to Send (RTS)-Clear to Send (CTS) mechanism of IEEE 802.11 DCF, sometimes referred to as "virtual" carrier sensing), and
- *backoff mechanisms* for regulating access attempts.

### 4.3.2 Receiver-Based MAC Paradigm

MAC protocols typically implement a blocking scheme in order to limit channel access and avoid data packet collisions. The most important protocol in this regard used today, the IEEE 802.11 Distributed Coordination Function (DCF) protocol, implements both carrier sensing and collision avoidance: A transmitter first must sense the medium interference free before making an access attempt. The access attempt is then initiated by transmitting an RTS control message. If no error occurs, the receiver responds with a CTS message. This control handshake may be omitted for small packets for efficiency reasons.

The "blind" blocking scheme using RTS and CTS messages is not suitable for multi-user detection. It turns any node receiving either one of these messages into a so-called

*exposed terminal*. This means that such a node gets blocked and may not send or receive data, irrespective of whether or not it would actually disturb other transmissions.

If physical layer capabilities are to be exploited, the decision on whether or not channel access can be granted should therefore not be based on physical carrier sensing or on collision avoidance blocking an entire area. Instead, this decision should depend on the signal processing capabilities of concurrent receivers, multi-user detection in our case.

To this end, the proposed MAC protocol for multi-user detection, called *MUD-MAC*, adopts the following receiver-based paradigm:

> Each transmission must be *announced* by a transmitter, and any co-located concurrent receiver may *object* to the transmission depending on its MUD capabilities. If no objection occurs, the transmitter may proceed with data transmission.

With this objection mechanism, carrier sensing as primary means for collision avoidance is not required in MUD-MAC. By the announcement-objection paradigm, a high level of transmission parallelism can be achieved since only concurrently *receiving* nodes can block other nodes. This effectively eliminates the exposed terminal problem.

Another known issue in MAC protocol design are *hidden terminals*. Any node that may potentially cause disruptive interference at a receiver and that is not detectable at the transmitter (by carrier sensing) is called *hidden*. Eventually, hidden terminals are the motivation to use MAC protocols that go beyond carrier sensing. With MUD-MAC, hidden terminals are effectively avoided by the possibility for receivers to object to other transmissions.

### 4.3.3 MUD-MAC Protocol Description

#### Signaling messages and frame structure

For implementing the announcement-objection principle, the following messages are defined: An announcement message `ANN`, an objection message `OBJ`, payload data blocks, and an acknowledgment message `ACK`.

For each of these messages there is a corresponding time slot when the message is to be transmitted, separated by an Interframe Space (IFS). These time slots preserve a frame structure (requirement **(R1)** of Sec. 4.2) as shown in Fig. 4.7. It is assumed here that neighboring nodes are synchronized accordingly. The frame structure is continuously repeated over time. Multi-user detection can be applied during the data slot. During the slots for `ANN`, `OBJ` and `ACK` messages no MUD is applied. This is in order to satisfy the requirement **(R6)** that all signaling messages must be accessible without MUD.

Collisions can occur during the `ANN`, `OBJ` and `ACK` slots. When several `ANN`s collide at the intended receiver, the subsequently following vain data transmissions have a direct impact on the overall throughput. In order to reduce the probability of `ANN` collisions, nodes start transmitting the `ANN` with random delay. For this purpose the `ANN` slot is preceded by a number of $n_M$ "minislots". In this work, this parameter

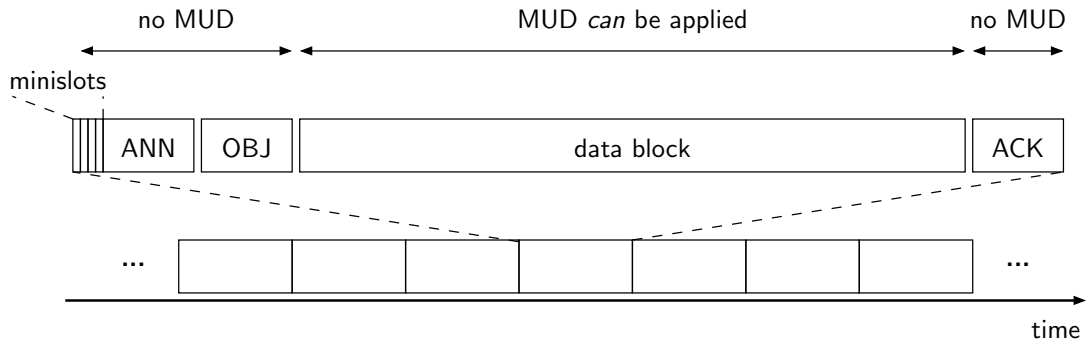



Figure 4.7: MUD-MAC frame structure, not to scale.

is chosen as $n_M = 4$. Each node transmitting an ANN randomly selects one of these minislots, and starts transmission in that minislot. During the minislots, up to their own chosen minislot, nodes intending to transmit an ANN listen for ANNs from other nodes. If a node senses an ANN from a different node, it loses the ANN contention resolution and backs off for a random duration. This backoff is a random exponential backoff without carrier sensing. The purpose of the exponential backoff, where backoff times are increased in an exponential manner as a node repeatedly loses the contention resolution, is the stabilization of medium access at high access rates.

The probability of ACK collisions is considerably smaller than for ANNs. In this regard, MUD-MAC does not aim for perfect control, and means for collision avoidance similar to the ANN minislots are not defined. Colliding OBJs are by far more frequent, but are not severe, as long as a node can in this case infer that an OBJ occurred at all. The latter is always assumed in the following.

### Medium access and cross-layer interaction

A node that wants to access the channel announces this intention using an ANN message in the next ANN slot. The ANN contains the ID of the transmitter, the ID of the intended receiver, and the data packet size (requirement **(R4)**). The ID is preferably the hardware MAC address, as it is for instance included in the current IEEE 802.11 DCF messages. Signatures are transmitter-based, and the transmitter ID readily specifies the used MUD signature (requirement **(R2)**). If MUD-MAC is applied on top of CDMA, transmitter-specific spreading codes defined for the UMTS uplink in [3G 02] may be used. For the case of IDMA, which is actually chosen for our simulations, there has been a suitable proposal for generating user-specific interleavers from the received ID [KB06a]. Then, by including the transmitter ID in the ANN, the receiver can determine the used signature.

Nodes receiving the ANN use it for estimating channel gains relative to the transmitter (requirement **(R3)**).

Neighboring concurrent receivers, i.e. nodes that will receive a data packet in the following data slot, can object to the announced transmission using an OBJ message. The

OBJ contains the ID of the announcing node. Whether or not a node objects depends on its MUD capabilities. An OBJ is issued when the announced data transmission would corrupt the parallel communication. If this is not the case, the estimated channel gain is used for an additional decoder branch of the multi-user detector, and no OBJ is issued. If no more decoder branches are available, an OBJ will again be sent (requirement **(R5)**). The announcement-objection mechanism is the key MAC functionality for providing cross-layer interaction between medium access and MUD. Further details of when such an OBJ is issued are described below.

After transmitting an ANN, a transmitter listens for OBJs. The transmitter abstains from transmitting data if it decodes an OBJ containing its ID, or senses (but cannot decode) an OBJ in the corresponding slot. It then backs off and contends for ANN transmission in a later MAC frame. In case no OBJ occurred and the data transmission was successful, the data packet is acknowledged by the receiver using an ACK message.

Even when split into several data blocks, only one ANN and one ACK have to be exchanged for one data packet. Determining an appropriate data block size is subject-matter of Sec. 4.4.

### Objection criterion

An important aspect of the proposed protocol is the assessment of received announcements. A node receiving a data packet listens, in between the data blocks of the packet, for ANNs in the corresponding ANN slot. Three situations may occur:

- The receiver does not sense any ANN. In this case, it does not transmit an OBJ.
- The receiver decodes an ANN. In this case, it checks whether its MUD capabilities are able to deal with the expected additional interference. If this is not the case, the receiver will subsequently transmit an OBJ, indicating the transmitter ID which was included in the ANN.
- The receiver senses, but cannot decode, a signal during the ANN slot. In this case, it pessimistically assumes that a subsequent data transmission could constitute corruptive interference. It thus issues an OBJ not indicating a transmitter ID.

The most complex case is the second: The receiver will issue an OBJ if it has no more unused MUD branches available. This is always the case for a "weak" receiver. If MUD signatures were not unique for a transmitter – which we however assume throughout this work – then code collision can be a further reason for an OBJ. Finally, the receiver will issue an OBJ in case it has been addressed by the ANN, since it is already busy with receiving a data packet. As a variant in future work, extensions of the protocol may be able to allow for simultaneous receptions by one and the same receiver.

### Backoff mechanisms

Similar to other MAC protocols, MUD-MAC provides backoff mechanisms for contention resolution and failure recovery. Backoffs are entered by transmitters after

receiving an ACK, in the case of losing the ANN contention resolution, in case of receiving an OBJ, and in the case of a missing ACK. In the latter three cases, the backoff is exponentially increased if the backoff reason occurs repeatedly for one packet. All backoffs are simple wait backoffs, i.e. they do not involve carrier sensing.

## 4.4 Protocol Parameterization and Analysis

For a well-performing design of the MUD-MAC protocol, the data block size (cf. Fig. 4.7) has to be chosen carefully.

In order to fit into the data block frame, data packets are split into a number of $b$ data blocks. A large data packet may thus require several frames. The last data block of a data packet is completed with dummy bits, if necessary.

This section determines an appropriate block size by looking at analytical throughput bounds for dense topologies, i.e. a contention area. The throughput bounds do not only serve the important aspect of block size determination, but are also part of the performance evaluation of MUD-MAC with respect to throughput. A further analysis of the throughput performance is provided in Sec. 4.6 based on simulations. The overall network throughput is chosen as performance figure since it is one of the most essential network measures related to medium access. Other measures, such as transmission delay or energy consumption, are important as well, but not considered in this work.

In this section, we evaluate the MUD-MAC protocol and, for comparison, the IEEE 802.11 Distributed Coordination Function (DCF) protocol, 1999 Edition [IEE97], with the RTS/CTS reservation mechanism. We will be looking at the transmission of data packets with a fixed size of $N_b = 8192$ bits.

The throughput bound of a MAC protocol for dense topologies is governed by the signaling overhead per data packet, and the degree of parallelism in transmitting information. MAC protocols with carrier sensing and blocking do not allow for parallel transmission in a contention area. For MUD-MAC, the number of parallel transmissions is upper-bounded by the number of data blocks, $b$, into which packets of size $N_b$ are split. This is because only $n_A = 1$ ANN can successfully initiate a data transmission per frame, as illustrated in Fig. 4.8.

In the following, we combine the above statements on transmission parallelism with per-packet overhead of control messages. Based on this, we derive throughput bounds for both 802.11 and MUD-MAC. The actual throughput can only be smaller than these bounds, since backoff mechanisms, failure recovery and retransmissions add overhead.

Considering a packet size of $N_b$ bits and the different message and time durations that are necessary to transmit a packet, the throughput bound $\widehat{R}_{802.11}$ for 802.11 can be expressed as

$$\widehat{R}_{802.11} = \frac{N_b}{3 \cdot t_{\mathrm{SIFS}} + t_{\mathrm{DIFS}} + t_{\mathrm{RTS}} + t_{\mathrm{CTS}} + t_{\mathrm{DATA}} + t_{\mathrm{ACK}}}, \tag{4.5}$$

where $t_{\mathrm{SIFS}}$ and $t_{\mathrm{DIFS}}$ are the Short Interframe Space (SIFS) and Distributed Coordina-

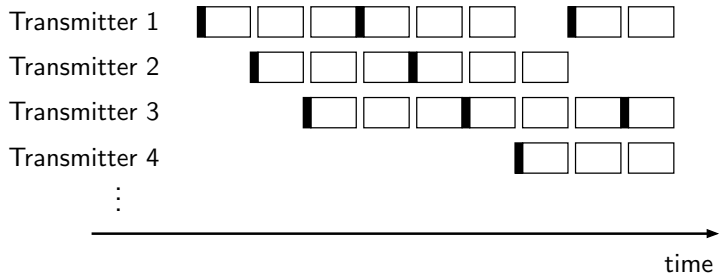



Figure 4.8: With only $n_A = 1$ ANN slot (depicted solid) per frame as assumed throughout the analysis, the beginning of only one packet can be announced per frame. When packets are split into $b = 3$ data blocks as illustrated here, at most $b = 3$ simultaneous transmissions are possible in a contention area.

tion Function Interframe Space (DIFS) of the IEEE 802.11 standard, and $t_{\mathrm{RTS}}$, $t_{\mathrm{CTS}}$ and $t_{\mathrm{ACK}}$ are the durations of the RTS, CTS and ACK packets of the standard, respectively. All time durations are fixed except the data packet duration $t_{\mathrm{DATA}}$, which depends on $N_b$ and the data transmission rate.

For MUD-MAC, a throughput bound can be given in a similar fashion. The number of decoder branches, $K$, limits the number of parallel transmissions from the physical layer perspective (cf. Eq. (4.3)). From the MUD-MAC perspective, as just shown, it is additionally limited to $b$. For our MUD-MAC design, in order to fully exploit the potential interference cancellation assuming that all receivers have $K$ decoder branches, we thus have to ensure $b \geq K$. The throughput bound $\widehat{R}_{\mathrm{MUD-MAC}}$ for MUD-MAC is given by

$$\widehat{R}_{\mathrm{MUD-MAC}} = K \cdot \frac{\frac{N_b}{b}}{(n_M + 4) \cdot t_{\mathrm{IFS}} + t_{\mathrm{ANN}} + t_{\mathrm{OBJ}} + \frac{N_b}{b \cdot R_D} + t_{\mathrm{ACK}}}, \tag{4.6}$$

where $R_D$ is the data transmission bit rate, $t_{\mathrm{IFS}}$ is the inter-frame spacing between the slots of the frame structure, and $n_M$ is the number of minislots used for the collision resolution of ANNs, each with duration $t_{\mathrm{IFS}}$.

A possible extension to the above described MUD-MAC frame structure are *multiple* ANN slots (back-to-back) per frame, each with preceding contention minislots. A number of $n_A$ ANN slots can allow for further transmission parallelism and thus increase network throughput:

$$\widehat{R}'_{\mathrm{MUD-MAC}} = n_A \cdot K \cdot \frac{\frac{N_b}{b}}{n_A \cdot (n_M \cdot t_{\mathrm{IFS}} + t_{\mathrm{ANN}}) + 4 \cdot t_{\mathrm{IFS}} + t_{\mathrm{OBJ}} + \frac{N_b}{b \cdot R_D} + t_{\mathrm{ACK}}}. \tag{4.7}$$

A similar modification could be made with respect to multiple ACK slots in order to reduce the probability for ACK collisions. This is thus coupled with choosing $n_A > 1$, i.e. in the case of a possibly very high number of parallel transmissions. In this work,

Table 4.1: Slot durations.

| 802.11 | MUD-MAC |
|---|---|
| $t_{\mathrm{RTS}} = 192\,\mu\mathrm{s} + \frac{160\,\mathrm{bit}}{R_C}$ | $t_{\mathrm{ANN}} = 192\,\mu\mathrm{s} + \frac{160\,\mathrm{bit}}{R_C} + n_M \cdot 20\,\mu\mathrm{s}$ |
| $t_{\mathrm{CTS}} = 192\,\mu\mathrm{s} + \frac{112\,\mathrm{bit}}{R_C}$ | $t_{\mathrm{OBJ}} = 192\,\mu\mathrm{s} + \frac{112\,\mathrm{bit}}{R_C}$ |
| $t_{\mathrm{DATA}} = \frac{N_b}{R_D}$ | $t_{\mathrm{DATA}} = \frac{N_b}{b \cdot R_D}$ |
| $t_{\mathrm{ACK}} = 192\,\mu\mathrm{s} + \frac{112\,\mathrm{bit}}{R_C}$ | $t_{\mathrm{ACK}} = 192\,\mu\mathrm{s} + \frac{112\,\mathrm{bit}}{R_C}$ |
| $t_{\mathrm{SIFS}} = 10\,\mu\mathrm{s}$ | $t_{\mathrm{IFS}} = 20\,\mu\mathrm{s}$ |
| $t_{\mathrm{DIFS}} = 50\,\mu\mathrm{s}$ | |

however, we restrict MUD-MAC to the case of $n_A = 1$, leading to rather pessimistic performance results.

For numerical analysis, we consider the slot durations of Tab. 4.1. The transmission bit rate of control messages is denoted as $R_C$. Since we are interested in an upper bound on the throughput, we assume $b = K$ in the following.

As an illustration of the above formulas, Fig. 4.9 shows the throughput limit of the MUD-MAC approach, $\widehat{R}_{\mathrm{MUD-MAC}}$, in comparison to the throughput limit $\widehat{R}_{802.11}$ of the single-user detection system of 802.11. The control signaling bit rate is $R_C = 1\,\mathrm{Mbit/s}$.

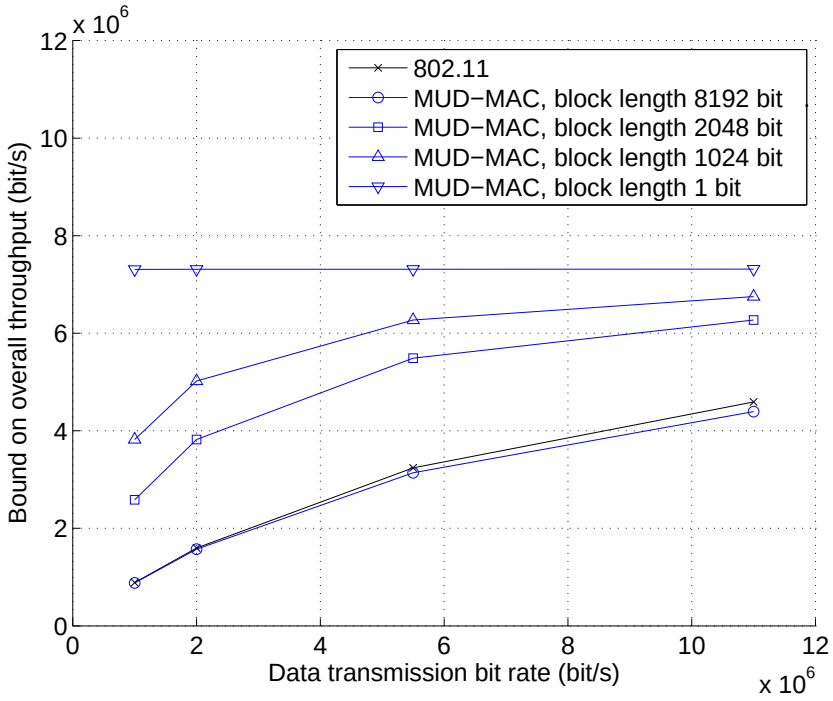



Figure 4.9: Bounds on overall throughput in a contention area for 802.11 and MUD-MAC as a function of the data transmission bit rate $R_D$.

The diagram shows that a high number of data blocks $b$ per packet, corresponding to a small data block size $\frac{N_b}{b}$, is preferable because of higher throughput. For MUD-MAC with $b = 1$ block per packet, the throughput bound is almost identical to 802.11.

This is because similar amounts of control signaling are necessary per payload packet. As $b$ approaches 8192, the throughput limit for MUD-MAC is shifted toward higher throughput.

In Fig. 4.9, we disregarded for the moment that the MUD performance depends on the block length on which MUD is applied. Taking this into account, a small block length (one bit in the extreme case of $b = 8192$) is of course impractical. Furthermore, we would require $K = 8192$ for the upper-most throughput bound in Fig. 4.9.

Based on Fig. 4.9, we can trade off data rate with signal processing aspects. As a first outcome of this section, we choose $b = 4$. The receivers should then have at least four decoder branches, which is reasonable for receiver design. The resulting data block length is 2048 bit, which is enough for a CDMA or IDMA scheme to perform well. It is in fact a conservative choice, keeping in mind that the results of Fig. 4.5 have been obtained for a block length of 128 bit.

We will use $R_D = 2$ Mbit/s and a transmission bit rate for control signaling of $R_C = 1$ Mbit/s for all simulations in this chapter. As can be seen in Fig. 4.9, the corresponding throughput bounds for a contention area are $1.6$ Mbit/s for 802.11 and $3.8$ Mbit/s for MUD-MAC. Hence, as a second outcome of this section, we can state that looking at throughput bounds the throughput gain of MUD-MAC over 802.11 is a factor of 2.4. This will be a lead for the network simulations in Sec. 4.6.

The performance bounds given here could actually be exceeded by using one announcement frame for announcing several, e.g. periodic, data packets. Such an extension could be very helpful for applications with small packet sizes, and in particular for real-time and multimedia services with video and voice data. The ultimate overall throughput for such an extension would then not be limited by $b$ any more, but only be governed by the number of available MUD branches of the receivers, access scheme parameters such as the processing gain, and the system bandwidth. This extension slightly complicates the protocol, and requires the channel coherence time to be sufficiently long. In this work, we assume the channel to be constant for the duration of one data packet only, and do not consider this extension.

## 4.5 Joint Simulation of Networking and Signal Processing

For simulating MUD-MAC, the simulation tool described in appendix C was used. It is an event-based simulator written in C++, calling a Matlab module for physical layer simulations. In the physical layer module, random bits are generated, encoded, transmitted and decoded. The simulator thus has a high level of detail across layers. Unfortunately, running the full Matlab implementation for evaluating multi-hop networks results in extremely long simulation run times. As a solution to this problem, this section discusses means to reduce the computational complexity, while still ensuring realistic results.

A first simplification of numerical efforts is the assumption of error-free transmission

in case of high SINR conditions. For both 802.11 and MUD-MAC, we assume that packets are transmitted uncorrupted when the SINR exceeds 10 dB.

A second simplification is motivated by the performance of iterative multi-user detection. Pertinent research shows that the single-user bound is usually approached in case all interference sources are known [WP99, MP02, SH04b, PLWL06]. Here, interference knowledge comprises the signatures in use, and the respective channels. According to the findings in this research, we assume that the single-user bound is reached when this information is available.

In many cases, not all interference sources are known. We can capture this more general situation by the notion of an *effective* SINR. It is defined in the following.

Assume that the signal to be received comprises both known and unknown interference. The SINR can then be calculated by

$$\text{SINR} = \frac{p}{p_{\text{known}} + p_{\text{unknown}} + p_z}, \tag{4.8}$$

where $p$ is the power of the desired signal, $p_{\text{known}}$ is the overall power of all known interference sources, $p_{\text{unknown}}$ is the overall power of all unknown interference sources, and $p_z$ is the noise power. The second simplification discussed above corresponds to setting $p_{\text{known}}$ to zero. In doing so, we assume the MAI term of (4.2) to disappear. The resulting SINR, called *effective* SINR here, is

$$\text{SINR}_{\text{eff}} = \frac{p}{p_{\text{unknown}} + p_z}. \tag{4.9}$$

If the number of interference sources exceeds the number of decoder branches of the receiver, then those interference components that cannot be assigned to a decoder branch also contribute to $p_{\text{unknown}}$.

The first simplification regarding high SINR can be applied on top of the second simplification. Then, if $\text{SINR}_{\text{eff}}$ exceeds 10 dB, transmissions are assumed to be error-free. If this is not the case, Matlab is called for bit level simulations of multi-user detection, and based on these simulations packet error probabilities are fed back to the network level simulations.

In order to test the simplifications, some of the data points in subsequent diagrams were simulated without any of these simplifications. The results are labeled “full PHY”. They closely match the results obtained with the simplifying assumptions.

## 4.6 Throughput Performance

In this section, the MUD-MAC protocol is analyzed in terms of throughput. By means of simulations, we compare MUD-MAC for an IDMA system with the standard IEEE 802.11 DCF protocol. The considered scenarios comprise single-hop and multi-hop networks.

Table 4.2: Simulation parameters.

| | |
|---|---|
| Control signaling bit rate $R_C$ | 1 Mbit/s |
| Data bit rate $R_D$ | 2 Mbit/s |
| Number of MUD branches $K$ | 5 |
| Number of ANN minislots $n_A$ | 4 |
| Packet size $N_b$ | 8192 bit |
| Data block size | 2048 bit |
| Transmission power | 100 mW |
| Decoding sensitivity | -81 dBm |
| Carrier sensing sensitivity | -91 dBm |
| Carrier frequency | 2 GHz |
| System bandwidth | 22 MHz |
| Noise temperature | 295 K |
| Path loss exponent | 3.0 |
| Simulated time | 60 s |
| Statistics reset time | 10 s |

### 4.6.1 Modeling and Parameters

Important modeling assumptions and parameters (cf. Tab. 4.2) used for simulations are summarized in the following. They apply to all simulations discussed in subsequent sections, except for the number of MUD branches. The latter is varied in Sec. 4.7.

For MUD-MAC, we assume that the network is synchronized to slots, and disregard possible overhead for achieving synchronization. The channel is modeled as described in Sec. 2.3.5.

The system parameters have to be chosen carefully to provide a fair comparison between 802.11 and MUD-MAC. For 802.11, we adopt the Direct Sequence Spread Spectrum (DSSS) scheme with the 11-chip Barker code for spreading, with Differential Quadrature Phase Shift Keying (DQPSK) resulting in a data rate of $R_D = 2$ Mbit/s [IEE97]. For IDMA in connection with MUD-MAC, a rate 1/11 repetition code and Quadrature Phase Shift Keying (QPSK) is applied. Forward Error Correction (FEC) is applied neither to 802.11 nor to MUD-MAC.

### 4.6.2 Single-Hop Scenarios

#### Simulation of a Particular Scenario

The analysis of the upper bound on the effective data rate in Sec. 4.4 did not comprise any backoff or failure recovery mechanisms. We would still, however, expect a similar performance gain of MUD-MAC over 802.11, since both use similar backoff mechanisms.

In the first set of simulations, we use an artificial setup where nodes are located in

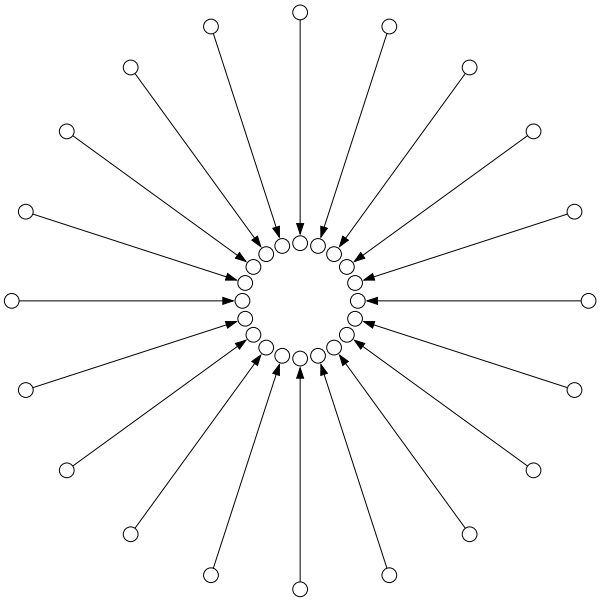


Figure 4.10: The particular topology with $N = 40$ nodes. The 20 transmitters are located on a circle with radius 50 m, the 20 receivers are located on a circle with radius 10 m. All nodes are within communication range. The arrows indicate traffic relationships.

circles, as shown in Fig. 4.10. In order to avoid effects of traffic relationship interdependencies, we divide nodes into an equal number of transmitters and receivers. Each transmitter (outer circle) has associated its own receiver (inner circle).

The traffic model assumes Poisson distributed arrivals of equally sized packets ($N_b = 8192$ bit). The offered traffic is varied by changing the mean packet inter arrival time. Simulations have been repeated and averaged, all shown results exceed a confidence level of 95% for a confidence interval of $\pm 5\%$ around the estimated average, assuming Gaussian distributed results. The transient phase at the beginning of the simulations is excluded from the statistics by resetting statistics after $10$ s of simulated time.

The number of decoder branches, $K = 5$ for each node, is larger than the number of data blocks, $b = 4$, which each data packet is split into. The number of parallel transmissions in a contention area is thus not limited by MUD capabilities, but it is limited to $4$. Consequently, the analysis of Sec. 4.4 with $b = 4$ can be applied to this scenario.

The simulation results are shown in Fig. 4.11. As the offered traffic increases, the overall throughput of 802.11 saturates at about 1.5 Mbit/s. For MUD-MAC, this is the case at about 3.8 Mbit/s. Beyond the point where the protocols saturate, the transmission queues of the devices are typically filled up. In summary, MUD-MAC is able to serve a 2.5-times higher offered traffic than 802.11 in this scenario.

The circles indicate the results obtained with the full physical layer implementation. The results closely match the simplifying simulations.

The throughput results for this particular scenario are quite encouraging: MUD-MAC is intended for multi-hop networks with high medium access contention, which is the

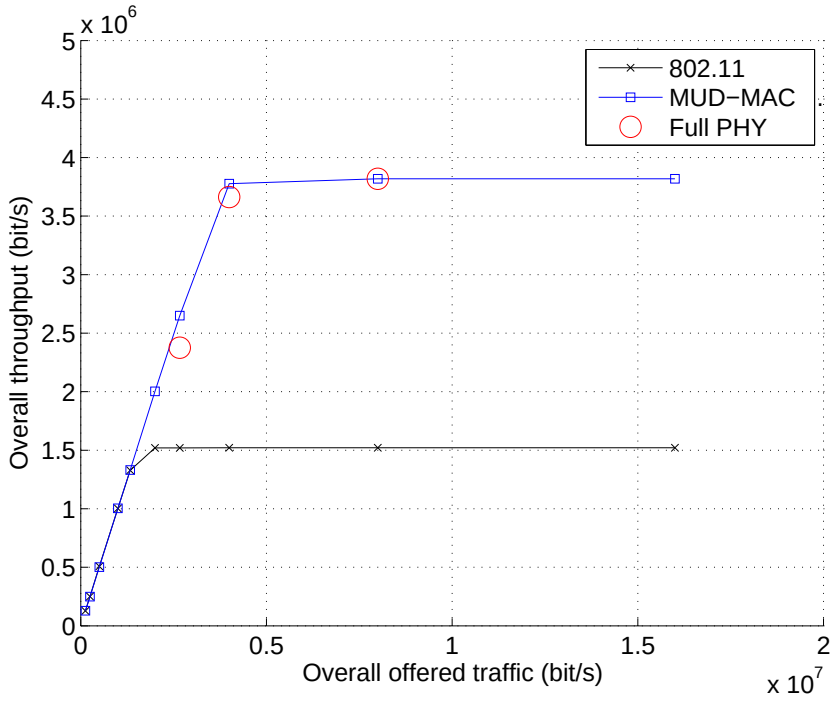



Figure 4.11: Overall throughput in the network of Fig. 4.10.

case in this scenario. The results depicted in Fig. 4.11 show that MUD capabilities can reduce the need for MAC layer contention resolution. In this simulation setup, the resulting performance gain in terms of the overall throughput is almost equal to the performance gain estimated in Sec. 4.4.

The following will assess the single-hop performance of MUD-MAC in a more complex setting.

### Simulation of Random Networks

Random networks are simulated in order to confront the MUD-MAC protocol with a large variety of traffic relationships and transmission-reception constellations, and to confront the multi-user detector with a large variety of near-far factor combinations. As any computer-based network simulation, the results of course do not exhaustively reveal the potential and limitations of the proposed scheme.

The networks are random in the sense that nodes are randomly and homogeneously placed on a network area of $50\,\mathrm{m} \times 50\,\mathrm{m}$.

In the particular scenario discussed above, the traffic offered to the network was varied by changing the average packet inter arrival time. In contrast to this, here the average inter arrival time is kept constant. Instead, the offered traffic is varied by placing different numbers of nodes in the system area.

In order to avoid isolated nodes when the number of nodes is low, the system area of $50\,\mathrm{m} \times 50\,\mathrm{m}$ ensures that all nodes are within transmission range. Consequently, only

single-hop traffic occurs.

Each node generates traffic, randomly choosing one of the nodes as sink. The same traffic model is used as with the particular scenario. After a transient phase which is excluded from the statistics, the traffic model parameterization lets all nodes have backlog traffic (filled queues) at all times.

Concerning the behavior of 802.11, it is known that the overall throughput degrades after a peak as the number of contending nodes is increased. It was verified that this is the case with the simulation tool developed in the course of this work when a packet size distribution other than a fixed packet size is used. However, for the traffic model assumed here, the overall throughput does not degrade significantly with an increasing number of nodes. This accords with the literature [Bia00], and holds for the case of rather large packet sizes when the RTS/CTS mechanism is used.

We expect from MUD-MAC to provide stable overall throughput as the number of contending nodes gets large. Moreover, we hope for a similar performance gain over 802.11 as in the particular scenario.

Fig. 4.12 shows the obtained results. MUD-MAC achieves a stable throughput even for high node densities, and outperforms 802.11 by a factor of about 2.3 for 100 nodes. The performance gain is thus very similar to the particular scenario, and again agrees with the throughput improvements estimated in Sec. 4.4.

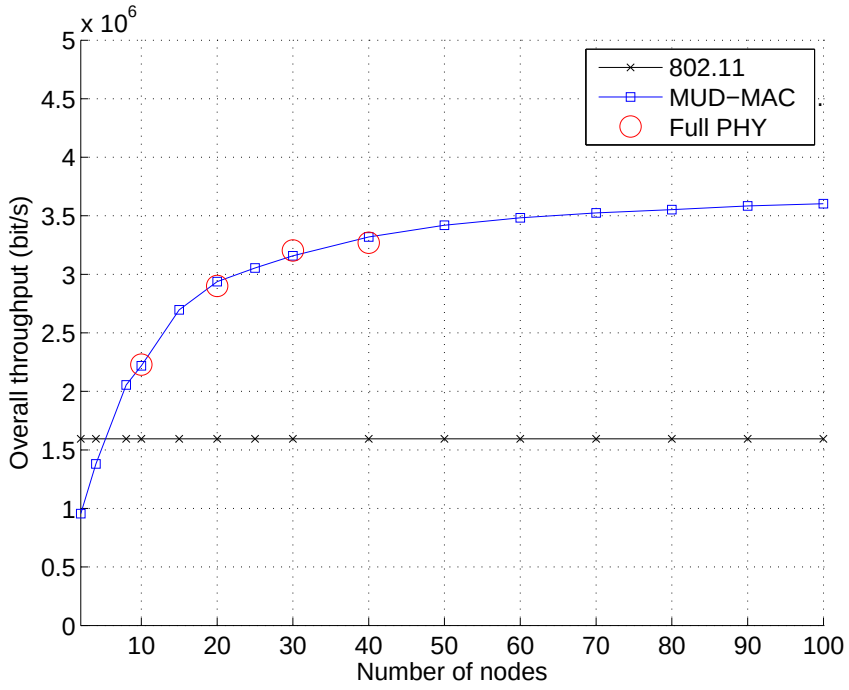



Figure 4.12: Overall throughput in the single-hop scenario.

### 4.6.3 Multi-Hop Scenarios

The single-hop scenarios of the previous section were characterized by a high node density, where the available decoder branches were well used. This section analyzes the throughput performance of MUD-MAC in a scenario with moderate node density. Besides the exact throughput performance, we are also interested in a proof-of-concept, demonstrating that MUD-MAC can be applied in multi-hop networks.

In our multi-hop network considered here, $N = 200$ nodes are homogeneously placed in a square system area of $1000\,\mathrm{m} \times 1000\,\mathrm{m}$. A sample realization of such a random network is illustrated in Fig. 4.13.

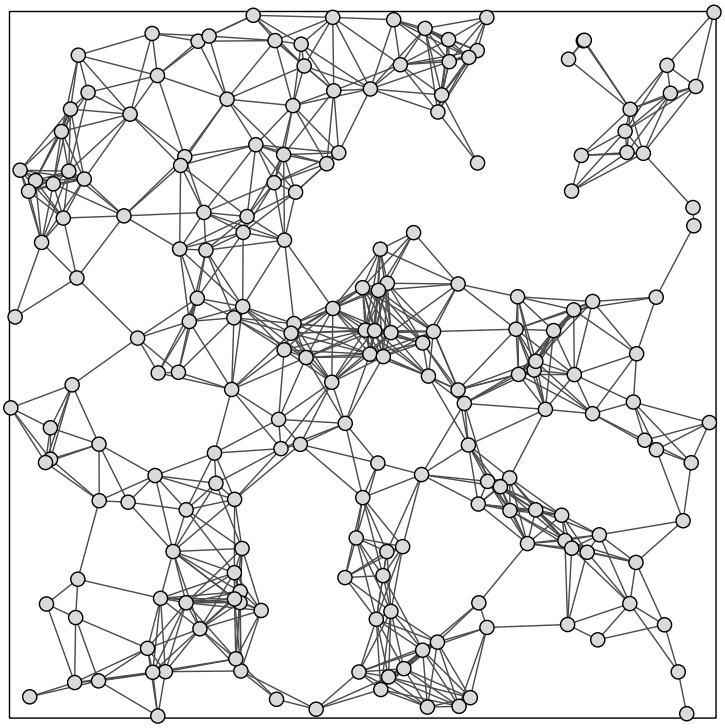


Figure 4.13: Sample random network with $N = 200$ nodes on $1000\,\mathrm{m} \times 1000\,\mathrm{m}$.

While the number of nodes is fixed, we vary the number of data flows in the network. In the simulations, the number of flows is varied between 20 and 200. A corresponding number of nodes is randomly selected as transmitters, and each transmitter chooses one of the remaining 199 nodes as sink. Packets are forwarded on the shortest path. The traffic model is the same as in the previous section, with a mean packet inter arrival time of $0.81920\,\mathrm{s}$. This results in a nominal bit rate of 10 kbit/s per flow.

The end-to-end per-flow throughput for both MUD-MAC and 802.11 is shown in Fig. 4.14. On average, about 120 flows are necessary for MUD-MAC to pile up backlog traffic in the packet queues, as compared to about 40 flows with 802.11. This means that MUD-MAC can sustain a significantly higher number of flows in the network than 802.11. The absolute throughput numbers suggest a similar benefit of MUD-MAC in multi-hop networks as in the single-hop case.

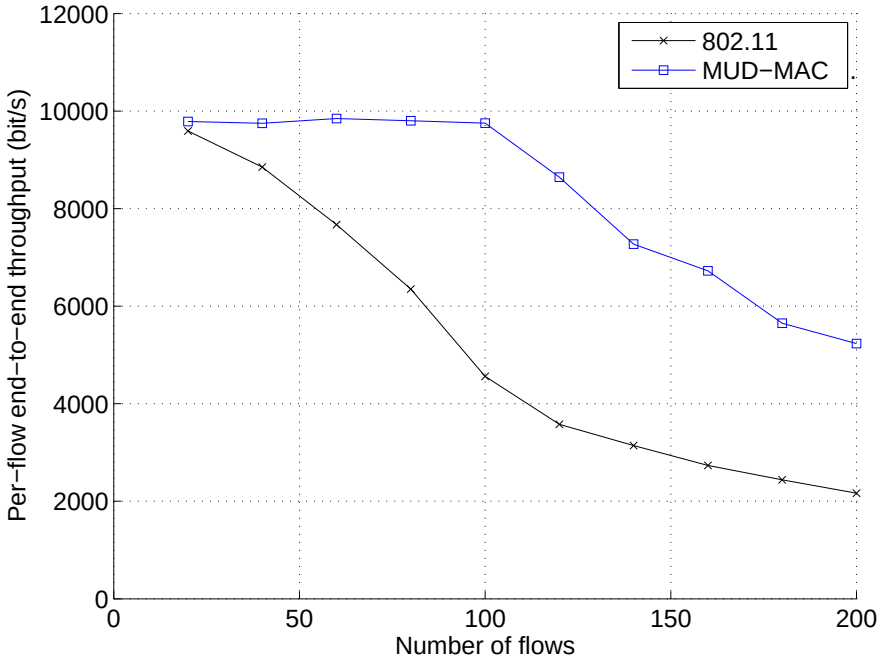



Figure 4.14: Throughput in the multi-hop scenario, varying number of data flows.

## 4.7 Heterogeneous Receiver Complexity

In previous sections we have analyzed the throughput of MUD-MAC when all receivers in the network have the same number of MUD branches. In this section we emphasize a specific MUD-MAC property: its ability to operate in scenarios with nodes having different MUD capabilities, and in particular its ability to protect weak receivers. This implies that nodes with MUD capability should not increase their own throughput to the disadvantage of nodes without MUD capability. We ask the following questions:

> Do weak receivers obtain a fair share of the system capacity?
>
> Does the per-node throughput depend on the MUD capabilities?
>
> How does the overall network throughput behave as the number of weak receivers increases?

In order to analyze this, further simulations in the scenario shown in Fig. 4.10 were performed. The overall offered traffic is unchanged in this section and amounts to $0.4 \cdot 10^7$ bit/s. According to Fig. 4.11, this is a traffic load slightly beyond what is sustainable by MUD-MAC in this scenario.

In contrast to the simulations presented in Sec. 4.6.2, four receivers out of the 20 receivers here have one, two, three, four, and five decoder branches, respectively. The simulation results for this setup are depicted in Fig. 4.15(a), showing the throughput of each individual receiver. The receivers are indicated in the order of their clock-wise location in the circular setup.

It can be observed that the throughput of the weak receivers, roughly $0.8{\cdot}10^5$ bit/s, is higher than what we would expect in a scenario with weak receivers only ($1{\cdot}10^6/20 \approx 0.5{\cdot}10^5$ bit/s, to be shown in Fig. 4.16). Furthermore, this throughput is slightly higher than the per-node throughput of $1.6{\cdot}10^6/20 \approx 0.8{\cdot}10^5$ bit/s for 802.11 (cf. Fig. 4.11).

From these observations we can conclude that, with MUD-MAC, the MUD capable nodes are not diminishing the throughput of weak nodes, but are even helping the weak receivers. By comparing the throughput of all nodes, we further realize that MUD-MAC provides for fairness among nodes with different MUD capabilities. A very similar throughput is obtained for all nodes, irrespective of whether a node has one, two, three, four or five MUD branches. As expected, there is no difference between four and five branches, since the medium access is limited to $b = 4$ parallel transmissions.

Fig. 4.15(b) shows results obtained with a slightly different setup. Here, five receivers have two decoder branches, another five receivers have three decoder branches, and the remaining 10 receivers have four decoder branches available. Again, the throughput achieved by the different nodes is similar, and only slightly higher for nodes with a higher number of decoder branches. The main difference is that the improved MUD capabilities as compared to the previous setup (there are no weak receivers with only single-user detection capability) lead to a general increase in throughput for all receivers.

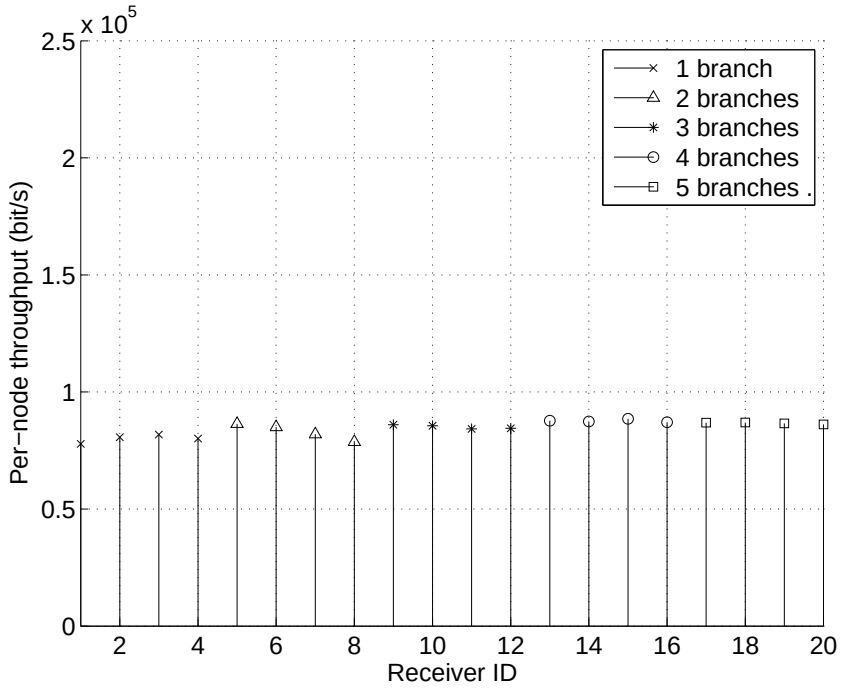
In order to analyze the impact of weak users numerically, we look at another simulation experiment. Here, we consider two types of receivers only: weak receivers and MUD receivers with five decoder branches. The results in terms of overall throughput are depicted in Fig. 4.16, where the number of weak receivers is varied from 0 to 20. From the graph we can see how the overall throughput in our network is reduced when MUD receivers are one-by-one replaced with weak receivers. The more weak receivers the network contains, the lower the number of parallel transmissions, and consequently the lower the overall throughput.

The throughput for 20 weak receivers corresponds to the throughput in Fig. 4.12 for the case of two nodes. In both cases, no gains by spatial reuse of radio resources can be achieved by MUD. On the other hand, MUD-MAC can provide a considerable increase in spatial reuse in heterogeneous networks where some nodes have simple single-user receivers.

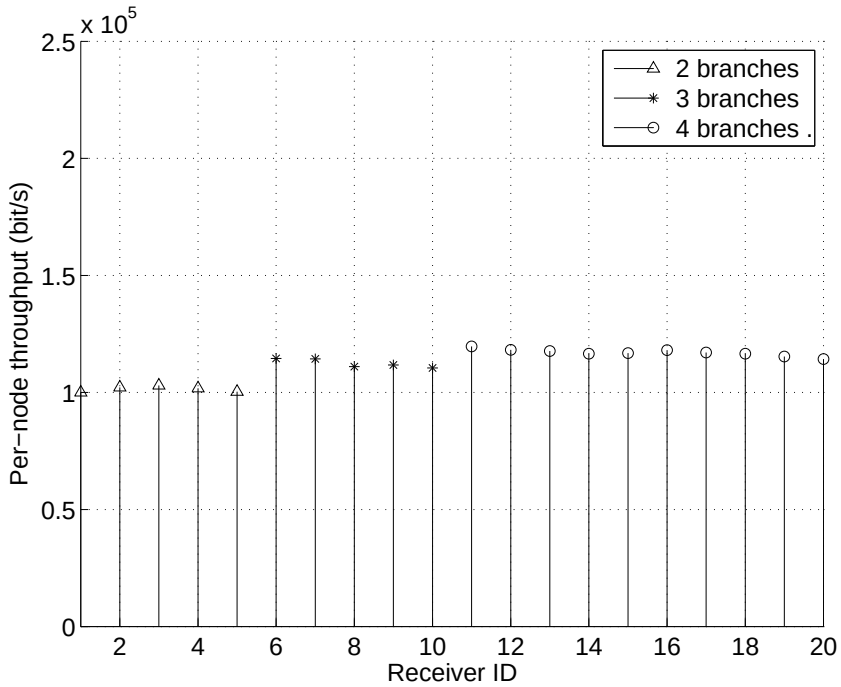
## 4.8 Related Work

A good overview paper discussing various aspects of spread spectrum in multi-hop (and, in particular, ad hoc) networks is provided in [AWH07].

While the work on spread spectrum communications in multi-hop networks is manifold, there is a general lack of work on multi-user detection in such networks. In [SE02], multi-user detection with CDMA is discussed in connection with routing, but mainly concerned with power control issues. Some further work considers MUD with slotted ALOHA [RW99, QZY05, QZC05], scheduled approaches [SG02], or requires exact




(a) Scenario including single-user detectors.




(b) Scenario with two to four decoder branches.

Figure 4.15: Per-node throughput with MUD-MAC for the scenario of Fig. 4.10 with heterogeneous MUD capabilities.

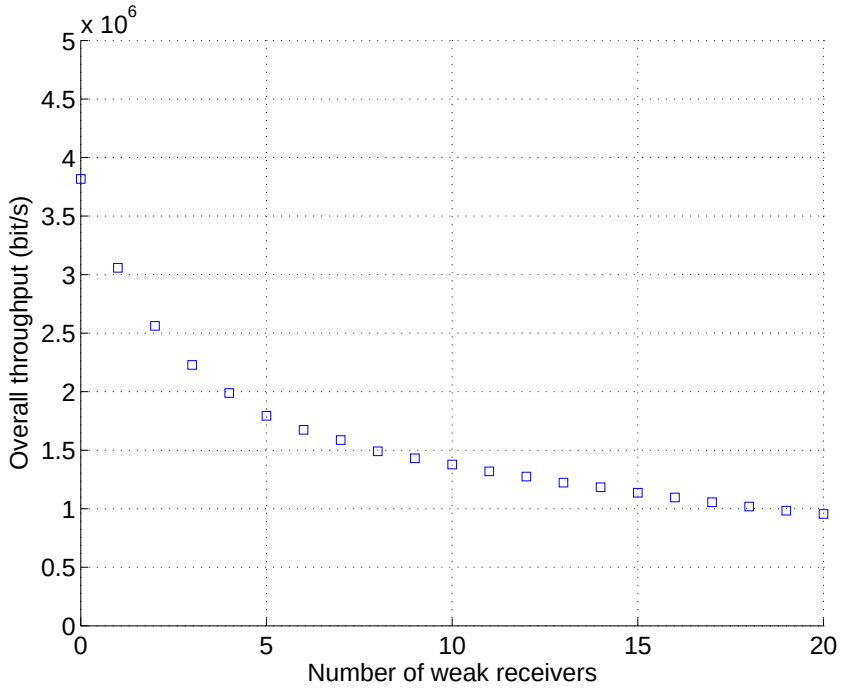



Figure 4.16: Impact of the number of weak receivers on the overall throughput.

location information of nodes [RM00]. Somewhat related concepts are with respect to using spatial multiplexing in multi-hop networks. Examples for research in this direction are [LCZ06, CLZ06].

A concept of announcing prospective interference in spread spectrum communications was used in [Yeh04], but not applied to multi-user detection. The need to make the channel access decision dependent on receivers instead of transmitters is also highlighted in [HA07]. There, the notion of a guard zone around receivers is used to maximize the transmission capacity of ad hoc networks with spread spectrum communications.

In [OWJR05], a novel MAC protocol for ad hoc networks with CDMA is proposed. It focuses on subdividing the available spectrum into channels, and allocates mobile devices to these channels in an interference-aware manner.

Medium access control in the context of multi-user detection in ad hoc networks is discussed in [ECS+07] and [LHL+08]. The authors of [ECS+07] describe a MAC protocol using RTS/CTS messages and distributed scheduling, and report significant throughput improvements and latency reductions as compared to existing systems.

## 4.9 Summary

This chapter discussed the use of multi-user detection in wireless multi-hop networks, and proposed a complete MAC layer framework for MUD.

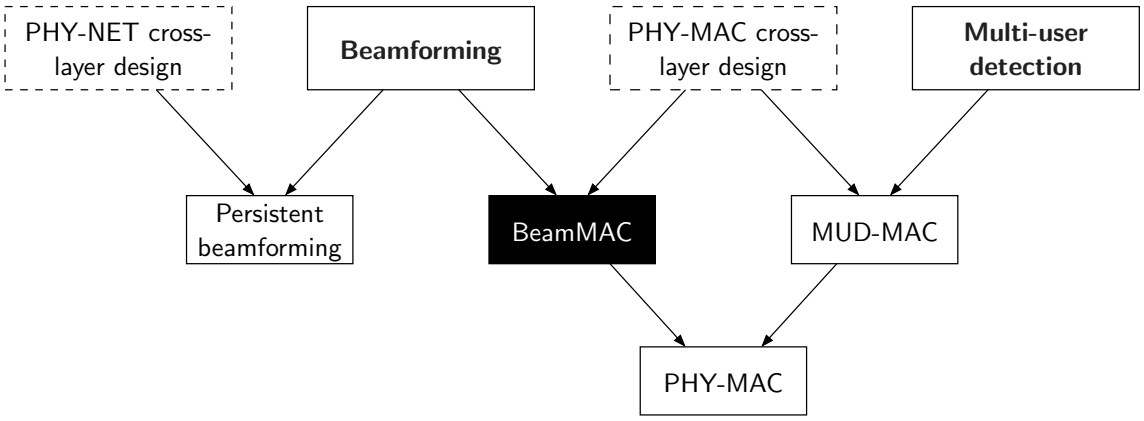
As first contribution, we started with the identification of six requirements for multi-user detection that influence MAC layer design. We found that these requirements are with respect to time synchronization, signature knowledge, proper channel estimation, knowledge about start time and duration of data packets, interference limitation according to the number of decoder branches, and the particular protection of single-user detectors.

The second contribution was the design of the MUD-MAC protocol. As opposed to previous work in the field, it adopts an announcement-objection paradigm to effectively implement multi-user detection in multi-hop networks. Its design is derived from the MAC layer requirements identified before. MUD-MAC can be applied in fully decentralized multi-hop networks since it does not assume any centralized interference management or power control. It also does not require the clustering of nodes into quasi-cellular sub-networks.

As a third contribution, the performance of MUD-MAC was estimated by throughput bounds, and thoroughly analyzed by extensive simulations. To this end, both a comprehensive simulation tool (cf. appendix C) and means to reduce the computational complexity of link level simulations were necessary. The computer-based analysis was with respect to the throughput performance, and considered different single-hop and multi-hop scenarios. The simulation results confirmed a throughput improvement over a single-user detection system of a factor of about 2.4, which was estimated from the protocol structure. Particular attention was paid to heterogeneous networks, where not all nodes have (equal) multi-user detection capabilities. It was found that MUD-MAC works well in such scenarios, and provides fairness between nodes with different MUD capabilities. As the number of nodes with simple single-user detectors increases, the throughput does not drop, but rather decreases gradually. For reasons of complexity, battery life time and monetary costs, such heterogeneous networks are likely to become reality if multi-user detection gets adopted for multi-hop networks. To the best of our knowledge, this is the first attempt to analyze the co-existence of single-user and multi-user detectors in a multi-hop environment.

Further work is necessary to develop advanced objection criteria. As indicated in [KVM+09], different transmission schemes lead to different multi-user detection performance. While a certain objection criterion may work well with one particular transmission scheme, it may be too optimistic – or too restrictive – with other transmission schemes. The MUD-MAC protocol could be extended to support multiple simultaneous packet receptions. This could further improve the performance of the overall system, in particular in networks with meshed traffic relationships. Another possible extension is with respect to multiple ANN slots. Their effect is captured by equation (4.7), but not yet analyzed in simulations. MUD-MAC also deserves further analysis with respect to its impact on transmission delays, and its performance under QoS constraints. Finally, it would be very interesting to implement the proposed scheme in hardware. Continuous advances in signal processing technology and the trend toward software-defined radio may soon ease rapid prototyping of cross-layer approaches such as the one proposed in this chapter.

# 5 Medium Access and Beamforming Antennas




Beamforming antennas can improve the performance of cellular wireless networks in terms of data rates and system capacity. In this chapter, we are interested in using beamforming antennas in multi-hop networks. The question is whether beamforming can provide similar gains in multi-hop networks as in cellular networks.

This chapter is concerned with the design of efficient MAC protocols for multi-hop networks with beamforming antennas. The intention is to adapt the beamforming pattern of mobile devices per-packet, thereby improving the throughput and reliability of wireless transmissions. This is in contrast to chapter 3, where beamforming antennas were used in a persistent manner.

Several aspects of beamforming antennas can contribute to throughput improvements in multi-hop networks. One aspect is the spatial decoupling of co-located transmitter-receiver pairs. By dividing signals in space, radio resources can be reused more aggressively than with omnidirectional antennas. A second aspect is link adaptation. High antenna gains at receivers and/or transmitters improve the link budget, and thereby allow for higher-order modulation and coding schemes providing higher data rates. Finally, high antenna gains may be translated into larger transmission ranges, leading to reduced hop distances. This, in turn, requires a smaller number of retransmissions for an end-to-end packet delivery, potentially improving throughput and reducing delays.

These three aspects cannot be exploited independently. For instance, allowing for larger transmission ranges limits the modulation scheme in use. In this chapter, we focus on the aspect of spatial reuse only. The topic of link adaptation is generally disregarded in this work in order to not complicate things. Effects of beamforming on hop distances have already been discussed in chapter 3.

Having described benefits of beamforming above, there is also a note of caution to be sound. Per-packet beamforming should not be implemented as mere physical layer adaptation. This is because MAC layer functionalities, such as carrier sensing and collision avoidance signaling, are conceptionally contradictory to beamforming antennas. These and further issues are summarized in this chapter. From the conclusions drawn, a novel MAC layer approach for beamforming antennas, called *BeamMAC*, is described and analyzed.

At first, typical issues with per-packet beamforming are demonstrated in Sec. 5.1. A thorough survey of previously proposed MAC protocols is provided in Sec. 5.2. Motivated by this previous work, requirements for an improved MAC protocol for realistic beamforming antennas are identified in Sec. 5.3. Based on these requirements, the BeamMAC protocol is described in Sec. 5.4. Its throughput performance is analyzed in Sec. 5.5. Sec. 5.6 is devoted to how beamforming and multi-user detection could be facilitated by an integrated MAC approach. Related work on per-packet beamforming in multi-hop networks is reviewed in Sec. 5.7, with a focus on capacity issues. Sec. 5.8 provides a summary and concludes this chapter.

Major concepts presented in this chapter have been previously published in [VBH06]. The survey part of this chapter is based on [VB05].

## 5.1 Conceptional Analysis of Medium Access with Beamforming

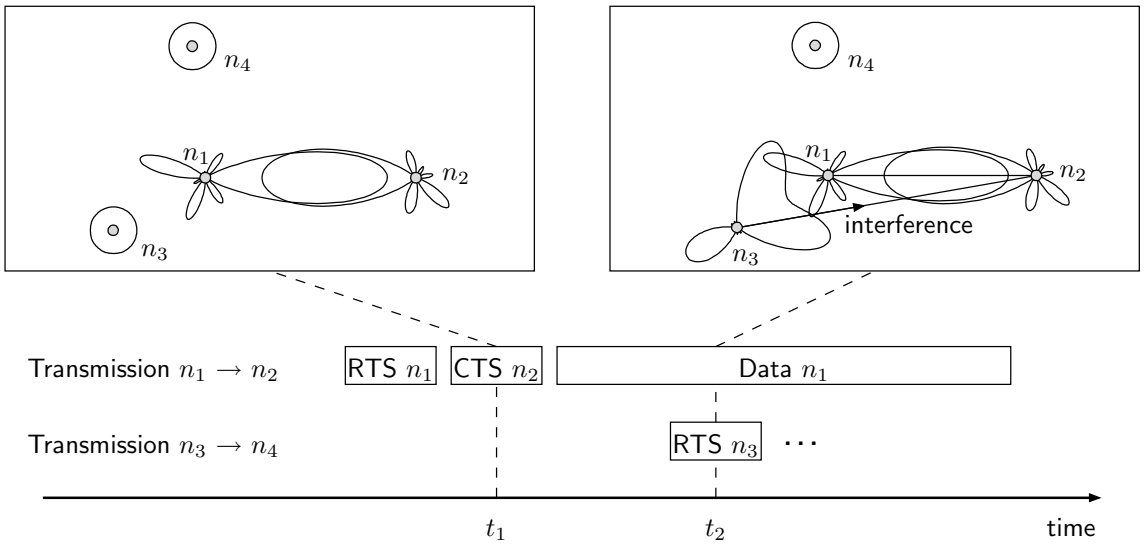
Chapter 4 showed that using multi-user detection requires a specific MAC layer design. The present section motivates MAC protocols that are specifically designed for beamforming antennas.

Some of the problems that occur with beamforming antennas can be explained best by showing sample scenarios. This is done in the following, where the focus is on using beamforming antennas in connection with the IEEE 802.11 DCF protocol.

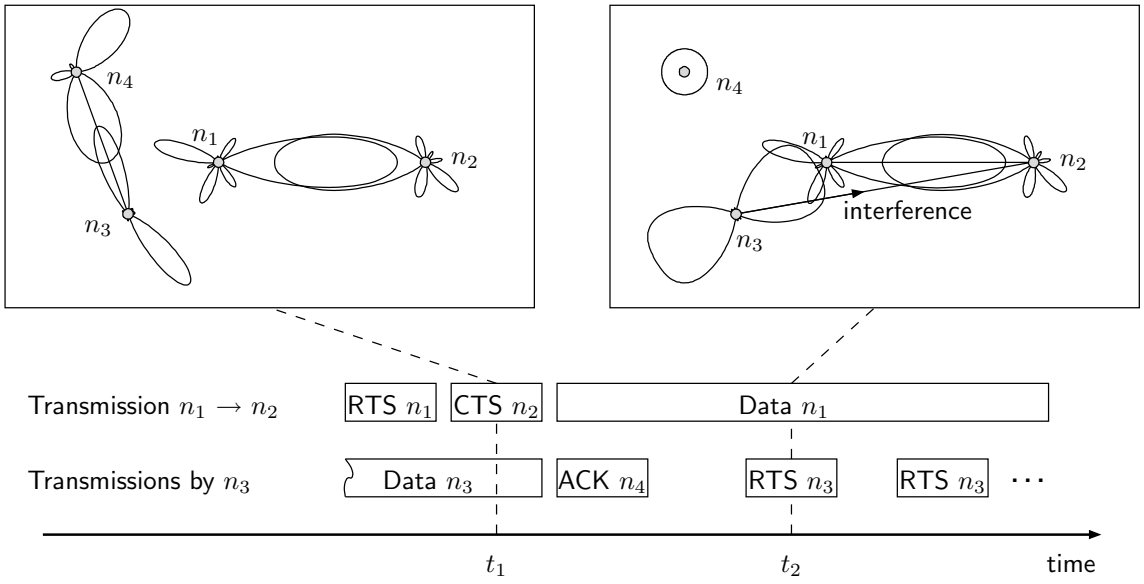
Fig. 5.1 illustrates mixed scenarios with UCA and ULA antennas with different numbers of antenna elements. As explained in the introduction, the drawn patterns are linear antenna gains. They do not illustrate transmission ranges. When all nodes are in omnidirectional mode in Fig. 5.1(a), all nodes are within communication range except $n_2$, which then can only communicate with $n_1$. The goal is to use the beamforming antennas to spatially resolve the transmissions from node $n_1$ to $n_2$ and from node $n_3$ to $n_4$.

When using the ALOHA protocol, i.e. carrier sensing only, node $n_3$ is not able to sense a data transmission of $n_1$ due to the small antenna gain of $n_1$ in the direction of $n_3$. A data transmission of $n_3$ would thus cause strong interference at node $n_2$ after beamforming toward $n_4$.

Even when using 802.11 for collision avoidance, the RTS and CTS messages exchanged by $n_1$ and $n_2$ would not block node $n_3$. This is illustrated in Fig. 5.1(a). When $n_2$ sends the CTS to $n_1$ at time $t_1$, $n_3$ is idle and receives in omnidirectional mode. It therefore




(a) Hidden terminal problem due to switching between omnidirectional and directional mode.




(b) Hidden terminal problem due to re-adapting the beamforming pattern.

Figure 5.1: Examples for how beamforming impairs the 802.11 collision avoidance.

is not able to decode the CTS. Then, as it beamforms toward $n_4$ to transmit an RTS at time $t_2$, its antenna array (ULA with $m = 6$ antenna elements in this case) exhibits an unintentionally strong antenna gain toward $n_2$.

This interference alone may be too weak to let the decoder of $n_2$ be unable to correctly receive the data transmission from $n_1$. In a multi-hop context, though, a plurality of nodes in the vicinity of $n_2$ may contribute to the interference level due to such MAC layer deficiencies, not only at $n_2$, but also mutually. The MAC protocol then cannot provide for transmission protection and stable throughput in the network.

With 802.11 and omnidirectional antennas, this situation would not occur since $n_3$ would be blocked by the RTS of $n_1$.

Another example is illustrated in Fig. 5.1(b). The figure shows how the collision avoidance of 802.11 can be affected by asynchronous switching of beamforming patterns, leading to unheard control messages.

In this example, $n_1$ and $n_2$ initiate a transmission using an RTS/CTS handshake. This handshake is not detected by $n_3$, as it is beamforming to $n_4$ (time $t_1$). After receiving an acknowledgment from $n_4$, $n_3$ wishes to transmit data to $n_1$. For that purpose, it beamforms toward $n_1$ and transmits an RTS (time $t_2$). At this time, it is neither aware of the fact that it causes significant interference at $n_2$, nor that $n_1$ is not able to respond to RTS messages. At worst, $n_3$ retransmits RTS messages until a count-out occurs.

With 802.11 and omnidirectional antennas, this situation would again not occur since $n_1$ would be blocked by $n_3$, and the two simultaneous communications at $t_1$ would be prohibited.

The example in Fig. 5.1(b) illustrates a further MAC layer problem with beamforming antennas which is referred to as *deafness* in the literature. Assume that $n_1$ and $n_2$ have quite an amount of data to exchange. If $n_1$ never switches back to omnidirectional mode in-between packets, but rather keeps the beamforming direction toward $n_2$, then $n_3$ will not be detected by $n_1$ for a long time period. The MAC layer access arbitration and fairness mechanisms then do not work any more, and no transmissions from $n_3$ to $n_1$ can take place until $n_1$ has finished its communication with $n_2$. On a networking level, $n_3$ may even decide to remove $n_1$ from its neighbor list. Beamforming can therefore even affect the consistency of neighbor discovery in a multi-hop network.

In summary, the problems described above are mostly due to the fact that adaptive beamforming can impair the exchange of MAC layer control messages. This is one of the reasons that motivated persistent beamforming as discussed in chapter 3.

The intention of using beamforming antennas is to allow for a higher density of simultaneous transmissions, and a reduced occurrence of exposed nodes. As shown by the above examples, this can easily capsize and then lead to excessive interference and loss of data, by far outweighing the gains achieved. In the course of this work, using beamforming antennas with 802.11 was investigated by way of simulations. Without showing numerical results here, let us note that throughput and delay significantly deteriorated when applying beamforming to 802.11. Changing certain technicalities in the PHY-MAC interaction – such as various kinds of mixing omnidirectional control packet exchange with directional data transmission – did not improve the situation.

Similar observations have been made by other researchers in this field. Since around the year 2000, this motivated quite a number of MAC protocols for beamforming antennas, which are surveyed in the following section.

## 5.2 Survey and Classification

This section provides an overview and survey of recently proposed MAC protocols for beamforming antennas in multi-hop networks.

Many of these protocols are extensions to the IEEE 802.11 DCF protocol (Sec. 5.2.2). For multi-hop networks it may be necessary to entirely re-conceive the MAC problem and develop novel, "non-802.11" MAC protocols (Sec. 5.2.3). Characteristics that are considered to be different from 802.11 comprise in particular scheduled approaches, protocols which do not rely on virtual carrier sensing, and protocols which are not based on the notion of blocking an area by RTS/CTS messages. The literature on such protocols is rather limited.

The authors of the surveyed protocols reported performance improvements over protocols for omnidirectional antennas in terms of network capacity and per-node data rates. These improvements typically amount to a factor of roughly $1.5$ to $5$. This section does not intend to compare protocols with respect to numerical results, since the modeling assumptions made by the authors are very diverse.

### 5.2.1 MAC Layer Issues with Beamforming Antennas

**Direction information** With omnidirectional antennas, the angular direction of a neighbor is not relevant. This changes in the case of beamforming antennas. For a transmitter, it is important to know where the receiver is located, so as to point the main lobe of the antenna pattern in this direction. For higher spatial reuse of radio resources, it is important for transmitters to know where not to point to, and for receivers to know in which directions interference sources are located. Some of the MAC protocols for directional antennas proposed in the literature assume that GPS information is available. This is a questionable assumption. Furthermore, it is generally accepted that direction information drawn from location information is inappropriate in the presence of shadowing and reflection in the signal propagation path. Using antenna arrays for DOA estimation is therefore more practical. With DOA estimation, the complexity of the algorithm, the accuracy of the results, and the time to estimate the direction of incoming signals are important aspects. A comprehensive review of DOA estimation methods can be found in [God97].

**Side lobe pattern** It is important to account for the side lobe pattern of beamforming antennas, in particular when it cannot be controlled arbitrarily (e.g. because of limited DOFs). With some of the previously proposed MAC protocols, this is done in a very explicit way. With such protocols, the medium access decision of a node is made based

on information about the entire antenna pattern of neighboring nodes. This is only practical if the antenna pattern can be described with few parameters, as it is the case when assuming very simple patterns (e.g. keyhole patterns). With realistic antenna patterns (as exemplified in appendix A), exchanging such large amounts of information between nodes is prohibitive. Thus, MAC design should not be required to be aware of explicit beamforming patterns on a protocol level. Information obtained from signal strength estimation should be sufficient.

**Carrier sensing** In many MAC protocols, carrier sensing is used to avoid interference. The problem with carrier sensing is that it is location-dependent, i.e. a prospective transmitter cannot accurately assess the signal power level at a receiver. With beamforming, carrier sensing is even less meaningful.

**Array rotation and changing beamforming pattern** The performance of MAC protocols is generally related to the system dynamics. If all traffic relationships and wireless channels were static, efficient scheduled approaches – implemented in a distributed manner – could be used. In dynamic multi-hop networks, contention-based approaches are preferred. Array rotations and frequently changing beamforming patterns add to the system dynamics, and may lead to protocol malfunctions. This has already been illustrated in connection with Fig. 5.1.

### 5.2.2 Extensions to 802.11 DCF for Beamforming

The protocols summarized in the following are based on the IEEE 802.11 DCF protocol. They typically comprise the well-known procedure of RTS and CTS messages, also known as virtual carrier sensing. Another concept known from 802.11 is using a so-called Network Allocation Vector (NAV). It keeps track of the time a node has to remain in a blocked state upon overhearing an RTS or CTS message. With a number of directional protocols, the NAV vector is extended to maintaining direction specific blocking information. It is then commonly referred to as Directional Network Allocation Vector (DNAV).

One of the first modifications of the DCF for directional antennas was proposed with the *Directional MAC* (D-MAC) protocol [KSV00]. It is a rather straightforward extension of the 802.11 protocol. With D-MAC, RTS, data, and Acknowledgment (ACK) packets are sent directionally. Alternatively, RTS packets are sent omnidirectionally if a transmitter is not blocked in any direction by the DNAV. This reduces the probability of control packets collisions.

Another early protocol was the *Multi-Hop RTS MAC* (MMAC) protocol [CYRV02]. It facilitates fully directional links – i.e. links that require both the transmitter and the receiver to beamform toward each other – by omnidirectional signaling over multiple hops. To this end, a special type of RTS packet is used, whose reception does not affect DNAV state information.

The *Tone-based directional MAC* (ToneDMAC) protocol [CV03a] seeks to indicate deafness to blocked transmitters and thus increase fairness. Using omnidirectional tones after the data/ACK exchange, both the data transmitter and the receiver indicate that they were recently engaged in a communication. Thus, a neighboring node can realize deafness if it overhears a tone from its intended receiver. In this case, the node reduces its contention window to the minimum value and has thus a fair chance to win the next channel contention. Multiple tones are used, and a node has to be able to identify the transmitter of a tone. It is assumed that a set of tone frequencies and durations is available. Further, the transmitter of a tone can be determined with a certain probability by means of its unique identifier and a hash function. The authors propose to exploit location information about neighboring nodes in order to reduce the probability of tone mismatching.

In [NYYH00], Nasipuri et al. focus on interference reduction by directional transmission of data packets. Their protocol seems to require all nodes to maintain the antenna array orientation at all times. Both RTS and CTS messages are sent omnidirectionally. The control packet exchange is augmented by a mechanism for determining angular directions. Throughput improvements by a factor of 2 to 3 are reported. The authors observe that the mobility of nodes does not affect throughput. However, they note that their mobility model only allows for position changes at discrete intervals of time.

In the protocol of Sanchez et al. [SGZ01], CTS and data packets are transmitted using directional antennas. The authors argue that an omnidirectional transmission of RTS packets reduces the hidden terminal problem, whereas a directional transmission reduces the exposed terminal problem. The authors propose two corresponding schemes of their protocol. By means of simulation, they conclude that a directional transmission of RTS packets generally outperforms the omnidirectional scheme. The authors further elaborate on the performance impact of the carrier sensing threshold. They conclude that, without carrier sensing, the hidden terminal problem dominates the exposed terminal problem.

The *Directional Virtual Carrier Sensing* (DVCS) approach [TMRB02] deploys a directional RTS/CTS exchange and a timer controlled DNAV table. The location information is cached whenever a node overhears signals. This information is used when setting up a communication. If a node does not have up-to-date location information about its intended receiver in its cache, or if no CTS packet was received upon directional transmission of an RTS packet, the RTS packet is sent omnidirectionally. The authors refer to this mechanism as DOA *caching*. Power control is assumed in order to not extend the transmission range beyond the omnidirectional range. This work is one of the few studies where mobility is considered in the simulations. The authors report that, in the case of mobility, physical carrier sensing brings about a dramatic performance improvement, in particular if nodes experience an accumulated interference due to numerous concurrent transmissions. In such cases, where nodes may fail to receive control packets (RTS or CTS) successfully, physical carrier sensing can effectively help to avoid collisions in the authors' scheme.

The MAC protocol of Lal et al. [LTR+02] is receiver-based and allows for simultaneous receptions from different transmitters (spatially resolved at the receiver). A particular

aim is to increase the throughput at nodes that lie on many active paths, so-called "bottleneck" nodes. A receiver synchronizes a number of packet receptions from other nodes. This is done by polling neighboring nodes by means of periodically transmitted Ready to Receive (RTR) messages. The subsequent RTS/CTS handshake (RTS packets contain training sequences for the purpose of beamforming) and data transmission is fully directional.

The *Dual Busy Tone Multiple Access with Directional Antennas* (DBTMA/DA) protocol [HSSJ02] adapts the Dual Busy Tone Multiple Access (DBTMA) protocol [HD02] to directional antennas. The idea here is to transmit tones directionally in addition to the directional transmission of control packets. This is motivated by the authors' observation that the performance of MAC protocols relying on the RTS/CTS scheme deteriorates in cases of control packet collisions. With DBTMA/DA, a "transmit busy tone" and "receive busy tone" is transmitted along with data transmission and reception, respectively. This provides a means for alleviating the hidden and exposed terminal problem. Numerical work showed remarkable effects on throughput and end-to-end delay. The authors also investigate the omnidirectional use of busy tones as well as hybrid schemes. The directional transmission of transmit busy tones turns out to be superior over the omnidirectional transmission, since it helps to avoid the exposed terminal problem. The performance comparison of directional and omnidirectional receive busy tones is less straightforward. In fact, a tradeoff exists. Omnidirectional receive busy tones do not result in new hidden terminal problems, but reduce spatial reuse. In contrast, directional receive busy tones provide better spatial reuse, but suffer from new hidden terminal problems. It should be noted that, by using busy tones, the maintenance of DNAV tables can be avoided. On the other hand, dual transceivers are necessary to implement busy tones.

A MAC protocol for "full exploitation of directional antennas" is proposed by Korakis et al. in [KJT03]. The most distinct protocol aspect is the rotational directional transmission of RTS packets (sweeping). This ensures a directional communication setup without requiring knowledge about the receiver location. However, multiple RTS packets transmitted for a single data packet degrade the MAC performance. Another major aspect is the use of a location table, maintaining the identity of each detected neighbor, the beam index on which it can be reached, and the corresponding beam index used by the neighbor. The location table is updated upon each packet reception and keeps track of the angular direction of neighboring nodes. The location table is further used for keeping track of beam directions that are skipped in the circular RTS transmission. The method seems to be particularly vulnerable to antenna rotation. The maintenance of location tables induces overhead and may not be suitable for dynamic scenarios. Another MAC mechanism using rotational beamforming was proposed by Roy et al. [RSB+03].

The *Smart-802.11b protocol* [SS04a] is based on 802.11 and deploys beamforming, a DOA estimation algorithm, and nulling. A transmitter must transmit a sender tone and must not transmit the data packet before receiving a receiver tone. The sender and receiver tones serve as a substitute for the conventional RTS/CTS exchange. Both tones are transmitted directionally. The sender tone is used for beamforming at the

receiver in a way that allows both for beamforming toward the transmitter and nulling toward interferers. A receiving node estimates the direction of the transmitter as the direction with the maximum received signal strength. As compared to RTS packets, no destination node can be indicated by the sender tone. Thus, a receiver may receive a data packet that was not intended for it. Tones preceding control packet transmissions are also used in [RSB+03].

### 5.2.3 Alternative MAC Concepts for Beamforming

There is limited work on complete MAC layer re-design for beamforming antennas in multi-hop networks. Research in this direction is summarized in the following.

The *Receiver-Oriented Multiple Access* (ROMA) protocol [BGLA02] is the only scheduled MAC protocol discussed here. While relying on local two-hop topology information, ROMA splits nodes into transmitters and receivers, which are paired together for maximum throughput. This allows for transmission SDMA and reception SDMA. The separation into transmitters and receivers is carried out in a random fashion. The ROMA protocol comprises the steps of priority assignment, transmission/reception mode assignment, hidden terminal avoidance, and selection of simultaneous receivers (for transmitting nodes) and simultaneous transmitters (for receiving nodes). In this survey, ROMA is the only protocol having explicit means for supporting QoS. A weight associated with each link reflects the data flow demand governed by upper layers of the transmitter and is used for contention resolution.

The *Direction-of-arrival MAC* (DOA-MAC) protocol [SS03] is a time-slotted approach based on the slotted ALOHA protocol. With DOA-MAC, time slots are broken into three minislots. During the first minislot, each transmitter transmits a tone toward the intended receiver. The receiving nodes run a DOA estimation algorithm and lock their beam toward the strongest signal. Furthermore, the receivers form nulls toward all the other identified directions. Data packets are transmitted during the second minislot. A packet is not acknowledged in the third minislot if the receiving node was not the intended receiver. False packet receptions may occur since no RTS/CTS handshake is performed before data transmission. Besides throughput, additional performance measures would be interesting for the evaluation of DOA-MAC, in particular the probability of unintended packet receptions and the probability of deadlocks. The authors proposed a similar approach in [SS04a], called *Smart-Aloha*. As an enhancement of DOA-MAC the authors implemented a single-entry cache scheme. It allows a receiver to beamform toward the second strongest signal if the receiver was not the intended receiver of a packet transmitted by a node providing the strongest signal.

In [SVGD02, SdV04], Stine et al. provide a framework for collision resolution in multi-hop networks, called *Synchronous Collision Resolution* (SCR). In [Sti06], it is applied to smart antennas and beamforming. SCR assumes that the wireless channel is slotted. At the beginning of each slot, transmitters contend for medium access using signals, and the surviving contenders go on with an RTS/CTS exchange. Different kinds of smart antenna techniques are applied, and it is highlighted that SCR can be combined

with spread spectrum techniques as well. While [Sti06] is one of the most thorough and illuminating articles on medium access and smart antennas in multi-hop networks, two statements are arguable. The first one is that directional antennas were not able to increase the signal transmission range over omnidirectional antennas. The second is that smart antennas were not able to both point a main beam toward a source of interest while simultaneously suppress interference from undesired directions. The paper does, however, assume advanced antenna techniques where the antenna weights are adapted to both the intended signal and the sources of interference. This "optimized reception" is not explained in detail.

In [GSD03], Grace et al. propose *Directional Synchronized Unscheduled Medium Access* (DSUMA), which uses the signaling approach of [SVGD02]. With DSUMA, the signals are relayed in the two-hop neighborhood of a node. No RTS or CTS messages are used since it is assumed that the collision resolution using signals is sufficient to avoid interference. It even results in node blocking such that, after collision resolution signaling, "promotion" signals are necessary to increase spatial reuse again.

### 5.2.4 Categorization

Main properties of the discussed MAC protocols are summarized in Tab. 5.1, 5.2 and 5.3, to allow for a better comparison of existing work in this field. Entries are left blank in case of absent information in the publications.

Tab. 5.1 indicates which message type is sent using omnidirectional (o) or directional (d) antenna mode. The usage of tones and "signals" (short transmissions to be sensed by other nodes, not containing payload data), which may serve different purposes, is indicated in the last column.

Tab. 5.2 lists the used antenna type, along with the modeling assumptions of the respective papers. It shows how the authors model side lobes, whether they exploit adaptive nulling, and whether they assume a static antenna orientation.

Looking back to chapter 3, we can see that persistent beamforming leads to long, fully directional links that happen to occur even without node coordination. "Fully directional" means that such links only occur if two nodes point their main lobes toward each other. For MAC protocols, where nodes are usually in omnidirectional mode when idle, such fully directional links are hard to establish. The nodes could either arbitrarily beamform in some direction when idle (similar to persistent beamforming), or coordinate fully directional links by signaling over multiple hops. The first solution is detrimental to MAC protocols for beamforming, the second causes significant signaling overhead. Therefore, only one of the surveyed MAC protocols seeks to establish fully directional links. It may be possible to establish fully directional links by implementing modulation and coding schemes with bit rates that are not sufficient for communication, but which allow for enough signaling to set up the receiver for subsequent directional reception. Also, it may be possible to carry out DOA estimation at the receiver on the basis of a signal level well below the decoding threshold. Such involved aspects are generally not considered in existing work on networking and medium access in

■ Table 5.1: Use of omnidirectional (o), directional (d) and rotational directional (rot) antenna patterns on transmission (TX) and reception (RX) for physical carrier sensing (PCS), messages, and tones by the transmitter (t) and receiver (r).

| | PCS | RTR | | RTS | | CTS | | Data | | ACK | | Tones | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | TX | RX | TX | RX | TX | RX | TX | RX | TX | RX | t | r |
| D-MAC | d | – | – | o/d | o | o | o | d | o | d | o | – | – |
| MMAC | d | – | – | d | o | d | d | d | d | d | d | – | – |
| ToneDMAC | d/o | – | – | d | o | d | o | d | o | d | o | o | o |
| Nasipuri et al. | o | – | – | o | o | o | o | d | d | – | – | – | – |
| Sánchez et al. | o | – | – | d(o) | o | d | d | d | d | – | – | – | – |
| DVCS | o/d | – | – | d/o | d | d | d | d | d | d | d | – | – |
| Lal et al. | | o/d | o | d | d | d | d | d | d | d | d | – | – |
| DBTMA/DA | o | – | – | o | | d | | d | | – | – | d/o | d/o |
| Korakis et al. | o | – | – | rot | o | d | o | d | d | d | | – | – |
| Smart-802.11 b | – | – | – | – | – | – | – | d | d | d | d | d | d |
| ROMA | – | – | – | – | – | – | – | d | d | – | – | – | – |
| DOA-MAC/ Smart-Aloha | – | – | – | – | – | – | – | d | d | d | d | d | – |
| SCR | – | – | – | o/d | o/d | d | d | d | d | d | d | o | – |
| DSUMA | – | – | – | – | – | – | – | d | o | | o | o | – |

multi-hop networks.

The discussed papers show that quite a variety of antenna models is used in the networking community:

- Switched beam antennas, where a node may select one beamforming pattern from a set of predefined patterns. With this model, different assumptions about side lobes and the susceptibility to interference are made:
  - Some researchers assume ideally sectorized beams which do not overlap, and where the antenna gain in angular directions outside the selected beam is zero. This is also referred to as *flat-top* model.
  - Others assume a similar model where beams are sectors of increased antenna gain, with reduced (but not zero) antenna gain in other directions. This is then a switched beam version of the *keyhole* model (cf. Sec. 3.3.1).
  - In other work, the patterns are assumed to be quite realistic phased-array patterns with a main lobe and several side lobes, e.g. linear arrays in broadside or end-fire operation. It is, however, usually neglected that the antenna pattern changes as the beamforming direction is switched. Instead, it is assumed that the pattern is simply rotated as-is.
- Steered beam antennas, where a node can point a main lobe to an arbitrary direction. The antenna pattern is then usually modeled as keyhole.

■ Table 5.2: Antenna types and modeling assumptions.

| | Antenna type | Side lobe model | Adaptive nulling | Direct. range extension | Fully direct. links | Static antenna orient. |
|---|---|---|---|---|---|---|
| D-MAC | switched | ideal sect. | n/a | – | – | – |
| MMAC | steered | keyhole | n/a | yes | yes | – |
| ToneDMAC | switched | sidelobes | n/a | yes | – | – |
| Nasipuri et al. | switched | ideal sect. | n/a | – | – | yes |
| Sánchez et al. | switched | keyhole | n/a | – | – | – |
| DVCS | adaptive | realistic | – | – | – | – |
| Lal et al. | adaptive | realistic | yes | – | – | yes |
| DBTMA/DA | switched | ideal sect. | n/a | – | – | yes |
| Korakis et al. | switched | ideal sect. | n/a | yes | – | yes |
| Smart-802.11 b | adaptive | realistic | yes | yes | – | – |
| ROMA | adaptive | none | ideal | – | – | – |
| DOA-MAC/ Smart-Aloha | adaptive | realistic | yes | yes | – | – |
| SCR | adaptive | realistic | yes | – | – | – |
| DSUMA | switched | ideal sect. | n/a | – | – | – |

– Realistic antenna patterns considering the antenna configuration (usually ULA or UCA) and signal processing methods are, in the context of multi-hop networks, used by few researchers only. Some apply gain maximization, others adaptive nulling techniques. Sometimes, however, the capabilities of adaptive antennas are overestimated and it is assumed that any interference can be suppressed by adaptive antennas.

Changes in the orientation of an antenna array occur as a node changes its moving direction or when the mobile device is rotated. Several papers postulate that antenna orientations remain unchanged, even in the presence of mobility. This may be uncritical as long as DOA estimation is performed for each individual packet, and the packet duration is small compared to the angular speed of the antenna. Assuming static antenna orientations becomes critical as soon as long-term directional MAC state information is maintained.

Tab. 5.3 categorizes the protocols by distinguishing scheduled protocols from random access schemes. The table also shows whether node synchronization is necessary for the protocol. Further, it indicates protocols with receiver initiated data transmissions, as opposed to transmitter initiation.

As shown in Tab. 5.3, many protocols assume known receiver locations (or, at least, angular directions) at the time of transmission initiation. Most protocols consider the task of DOA estimation to be part of medium access control, at least for beamforming at the receiver. With switched beam antennas, beam selection at the receiver is mentioned

■ Table 5.3: Protocol classification, direction information and mobility.

| | Scheduled | Synchronous | Receiver initiated | Known receiver location | DOA estimation | SDMA TX | SDMA RX | Mobility in simulations |
|---|---|---|---|---|---|---|---|---|
| D-MAC | – | – | – | yes | – | – | – | – |
| MMAC | – | – | – | yes | yes | – | – | – |
| ToneDMAC | – | – | – | yes | select | – | – | – |
| Nasipuri et al. | – | – | – | – | select | – | – | yes |
| Sánchez et al. | – | – | – | yes | select | – | – | – |
| DVCS | – | – | – | yes | yes | – | – | yes |
| Lal et al. | – | RX | yes | – | yes | – | yes | – |
| DBTMA/DA | – | – | – | – | select | – | – | – |
| Korakis et al. | – | – | – | – | select | – | – | – |
| Smart-802.11 b | – | – | – | yes | yes | – | – | – |
| ROMA | yes | yes | yes | | yes | yes | yes | – |
| DOA-MAC/ Smart-Aloha | – | yes | – | yes | yes | – | – | – |
| SCR | – | yes | – | yes/– | yes | – | – | – |
| DSUMA | – | yes | – | yes | | – | – | – |

in several papers, and always considered to be an accurate and faultless procedure.

Two of the papers assume that a transmitter and/or a receiver can handle multiple simultaneous transmissions using array processing, i.e. some sort of SDMA as used in cellular networks.

Tab. 5.3 further indicates whether mobility is considered in computer simulations, the primary means for protocol analysis in the literature.

## 5.3 Requirements for Per-Packet Beamforming

In Sec. 4.2, requirements for a MAC protocol design for multi-user detection have been summarized. For beamforming, the requirements are less definite. As shown in the survey of Sec. 5.2, even protocols very similar to the well-established 802.11 DCF protocol are conceivable for beamforming antennas. Yet, the design of such protocols is often based on assumptions that may be critical in a practical realization: availability of GPS location information, static antenna orientations, or idealistic antenna patterns and signal propagation conditions, just to mention a few. This section postulates preferable conditions of a realistic cross-layer design for beamforming antennas, focusing on MAC layer design.

As exemplified in Sec. 5.1, the additional dynamics introduced by per-packet beamforming can lead to MAC protocol malfunction. Maintaining beamforming patterns for

predefined durations and re-adapting antenna weights synchronously can significantly improve timing subtleties and avoid malfunctions. As a first condition, we therefore call for time synchronization on a packet frame level, with beamforming patterns only being changed in between frames. The synchronization accuracy does not have to be on a bit level, but the better the synchronization, the smaller the guard intervals in between frames can be designed.

As a further requirement, antenna arrays should be free to move and rotate. Long-term MAC layer state information about the directions of communication peers and interferers should be avoided. On the other hand, we consider the cost of frequent DOA estimation operations bearable, and assume that mobility and channel conditions allow us to rely on DOA information for the duration of at least one data packet.

A node cannot determine the beamforming pattern of other nodes unless it is explicitly signaled. Some MAC protocols for beamforming antennas assume this by exchanging main lobe direction information. Such information exchange causes signaling overhead and, to be useful, may require known array orientations obtained from compasses. It seems generally more practical to avoid such explicit pattern exchange, and instead only use implicit information that can be obtained from signal strength estimation.

MAC protocols for beamforming must be functioning under realistic channel conditions. In the *analysis* of MAC protocols for beamforming antennas it may be acceptable to consider only simplified wireless channels. Simplifications can be with respect to LOS conditions without multi-path propagation. What is important, though, is that such assumptions should never be preconditions by concept. Examples for prohibitive design are protocols relying on Global Positioning System (GPS) location information, because the direction of arriving signals in Non-Line of Sight (NLOS) channels is not the one calculated from explicitly exchanged GPS location information.

In summary, the following requirements for MAC protocol design for beamforming antennas are postulated:

- Time synchronization on a frame level **(R1)**,
- incorporation of DOA estimation and adaptive beamforming into the medium access mechanism, no usage of location information **(R2)**,
- exploitation of DOA information if available, without requiring it **(R3)**,
- implicit estimation of antenna gains using signal strength estimation, no explicit exchange of beamforming patterns **(R4)**,
- applicability to NLOS and multi-path channels **(R5)**,
- possibility to use both main lobe pointing and adaptive nulling **(R6)**,
- interoperability of nodes with different antenna array setups and different beamforming implementations **(R7)**, and
- protection of nodes without beamforming capabilities **(R8)**.

## 5.4 Protocol Design for Beamforming Antennas

This section proposes a MAC protocol for beamforming antennas, called *BeamMAC*. It is motivated by the fact that previous MAC protocols for beamforming antennas do not comply with the requirements as summarized in Sec. 5.3, and that new solutions are therefore necessary.

Similar design choices as described in Sec. 4.3.1 exist. Since the intended use case of BeamMAC in terms of the network scenario is similar to MUD-MAC, we again aim for a distributed, random access, transmitter initiated, time duplex MAC solution.

### 5.4.1 Receiver-Based MAC Paradigm

As discussed in Sec. 4.3.2, a node cannot determine the multi-user detection capabilities of neighbors, unless they are explicitly signaled. Similarly, it is impossible to implicitly determine the beamforming capabilities of neighboring nodes. Therefore, the effectiveness of the announcement-objection paradigm used in MUD-MAC is exploited in BeamMAC as well. The idea is again to make the medium access decision, which has to be made when arbitrating channel use, be incumbent on concurrent receivers.

Let us re-state the receiver-based announcement-objection paradigm for beamforming:

> Each transmission must be *announced* by a transmitter, and any co-located concurrent receiver may *object* to the transmission depending on its beamforming capabilities. If no objection occurs, the transmitter may proceed with data transmission.

The idea behind this approach is that a transmitter has to use the same antenna weights for transmitting an announcement as it is going to use for data transmission. Neighboring nodes can therefore use the announcement message both for beamforming adaptation and for deciding on whether or not to object to the transmission. The following details the BeamMAC implementation of this paradigm.

### 5.4.2 BeamMAC Protocol Description

#### Signaling messages and frame structure

With BeamMAC, medium access is organized using announcement messages (`ANN`) and objection messages (`OBJ`). Successful data transmissions are acknowledged by receiving nodes using acknowledgment messages (`ACK`).

The channel is time-slotted, and a dedicated time slot exists for each of the message types. The frame structure (requirement **(R1)** of Sec. 5.3) is depicted in Fig. 5.2.

A number $n_M$ of "minislots" is used to reduce the probability of `ANN` collisions. Each transmitter intending to access the channel picks one of the minislots. If no neighboring node started transmission before this slot, the transmitter starts transmitting an `ANN` in this slot. This functionality is the same as with MUD-MAC (chapter 4).

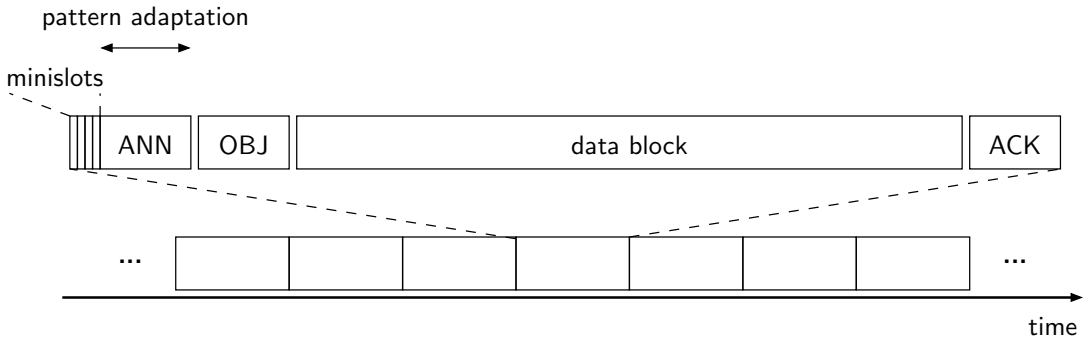



Figure 5.2: BeamMAC frame structure, not to scale.

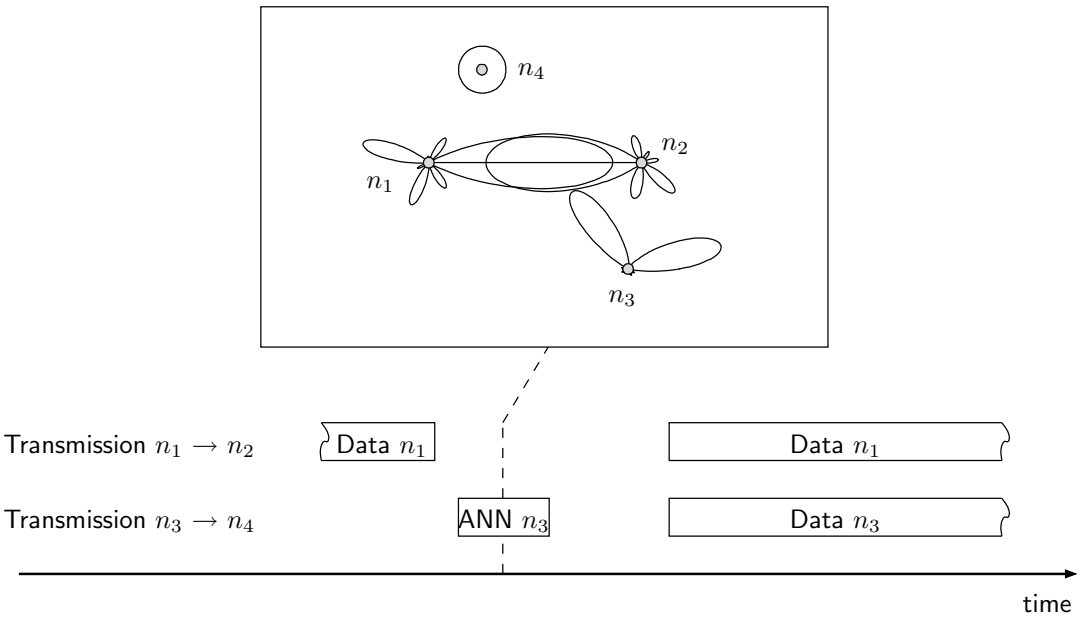
### Medium access and cross-layer interaction

The BeamMAC functionality is explained using two sample scenarios (Fig. 5.3). The scenarios are heterogeneous in the sense that the nodes are equipped with different antenna configurations (UCA and ULA with different numbers of antenna elements), according to what is demanded by requirement **(R7)** of Sec. 5.3.

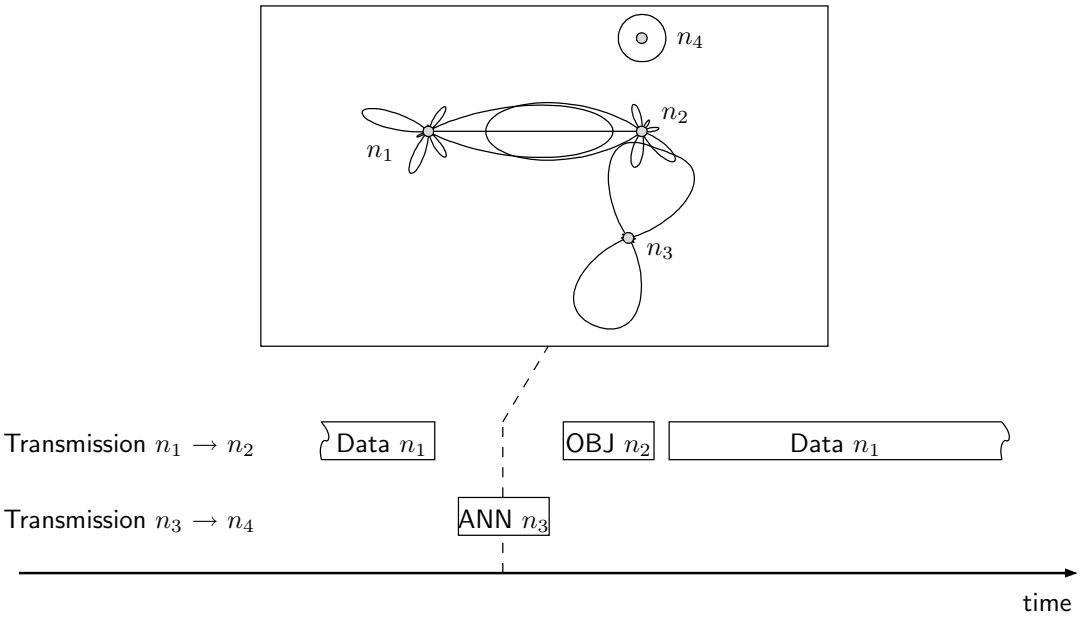
A successful channel access is exemplified in Fig. 5.3(a). In this example, a data transmission from node $n_1$ to $n_2$ is taking place. Since the data packet transmitted by $n_1$ is larger than one data block of a frame (Fig. 5.2), the packet is split into several data blocks, and the data packet transmission thus extends over several frames. Two of the blocks are illustrated in Fig. 5.3(a). In between these two blocks, in the `ANN` slot of the frame structure, node $n_3$ attempts to access the channel by transmitting an `ANN` message containing the ID of $n_4$. To this end, it beamforms toward $n_4$, and it is here assumed that the angular direction of $n_4$ is known by $n_3$. If it was not known, or if $n_3$ was not equipped with a beamforming antenna, $n_3$ could transmit the `ANN` in omnidirectional mode (requirements **(R3)** and **(R8)**, respectively).

During the `ANN` slot, all nodes receiving data listen for `ANN` messages. So does node $n_2$ in the present example. The low antenna gain of $n_3$ toward $n_2$ lets the `ANN` message not be detectable at $n_2$. If $n_2$ could detect it, it would estimate the signal strength of the `ANN` message and compare to the signal strength of the data transmission from $n_1$. In the first case, the `ANN` would simply remain undiscovered. In the second case, the signal strength estimation by $n_2$ would suggest that $n_3$ is either far away or has a low antenna gain toward $n_2$ (requirement **(R4)**). Either way, $n_2$ would not expect corruptive interference and would therefore not object, i.e. not send an `OBJ` in the `OBJ` time slot following the `ANN` time slot. A more rigorous discussion of the objection criterion is provided below.

For $n_4$, the `ANN` message indicates a subsequent data transmission directed to it, and it estimates the DOA of the `ANN`, adjusts its antenna weights, and beamforms toward $n_3$ (requirement **(R2)**). If $n_4$ is not able to estimate the angular direction or is not equipped with a beamforming antenna, it remains in omnidirectional mode for data reception (requirements **(R3)** and **(R8)**, respectively).




(a) Channel access mechanism in a high spatial reuse case.




(b) Announcement-objection mechanism in an interference avoidance case.

Figure 5.3: Illustration of the BeamMAC message exchange.

Node $n_3$ proceeds with data transmission in the subsequent data slot since it did not receive an OBJ upon the ANN. The two data transmissions $n_1 \rightarrow n_2$ and $n_3 \rightarrow n_4$ can occur simultaneously.

The nodes $n_3$ and $n_4$ keep their antenna weights until all data blocks constituting the data packet – and the ACK message by $n_4$ in case of successful decoding of all data blocks – have been transmitted. Then, the nodes return to omnidirectional mode when entering the idle state.

The example of Fig. 5.3(b) shows a case where BeamMAC avoids interference. In this example, the antenna gain of $n_3$ toward $n_2$ is not negligible. By estimating the signal strength of the ANN from $n_3$ (requirement **(R4)**), $n_2$ expects corruptive interference if $n_3$ would proceed with data transmission. Node $n_2$ therefore transmits on OBJ containing the ID of $n_3$, and $n_3$ abstains from proceeding with data transmission. Node $n_4$ uses the ANN to adjust its beamforming, but switches back to omnidirectional mode after not receiving data in the first data block following the ANN. In case it can decode the OBJ sent by $n_2$, it switches back immediately after. Upon receiving the OBJ, $n_3$ enters a backoff state and retries channel access at a later time.

### Objection criterion

Whenever a node is in the role of a receiver (i.e. it will subsequently receive a data block) and detects an ANN message, it checks whether it is indicated as receiver in this ANN. If so, it transmits an OBJ in the subsequent OBJ time slot, since we do not assume multi-user detection capabilities here. If it is not indicated as receiver, it computes the expected SINR

$$\mathrm{SINR}_{\mathrm{exp}} = \frac{p}{p_I + p_{\mathrm{ANN}} + p_z}, \tag{5.1}$$

where $p$ is the power of the desired signal determined from the previous data block, $p_I$ is the prevailing interference power at the receiver, $p_{\mathrm{ANN}}$ is the expected additional interference power estimated from the reception of the ANN, and $p_z$ is the ambient noise power. Since ANNs are transmitted using the same antenna weights and the same frequency band as data packets, the channel characteristics for ANNs are substantially the same as for the intended data transmission. Hence, given a predetermined transmit power, $p_{\mathrm{ANN}}$ can be used to estimated the additional interference that will occur if the receiver does not object to the ANN.

From $\mathrm{SINR}_{\mathrm{exp}}$, the data block size and the modulation scheme in use, the receiver can compute a block error probability for the subsequent data block. The receiver issues an OBJ with this very probability. This means that, if the expected block error probability is high, the receiver is likely to object.

A simplified objection criterion could be to specify a required minimum SINR, and to object to the announced transmission if $\mathrm{SINR}_{\mathrm{exp}}$ falls short of this required SINR.

#### Backoff mechanisms

The same backoff mechanisms as described in connection with MUD-MAC (Sec. 4.3.3) are used for BeamMAC.

## 5.5 Throughput Performance

In Sec. 4.4, a formal analysis of throughput bounds was given for MUD-MAC. For BeamMAC, such an analysis is difficult to derive since throughput improvements depend on antenna configurations, signal propagation environments, the spatial constellation of the nodes, and traffic relationships. In an ideal case, beamforming antennas can completely decouple all transmissions, and the throughput in a contention area with $N$ nodes is $N/2$ times the throughput of a system with omnidirectional antennas. In the worst case, the spatial constellation of the nodes does not allow for concurrent transmissions at all.

In order to obtain insights on the throughput performance of BeamMAC in random networks, computer-based simulations were carried out. The results are summarized in this section.

The same modeling assumptions hold as in Sec. 4.6, except that multi-user detection is not applied. The same parameters as summarized in Tab. 4.2 on page 94 were used.

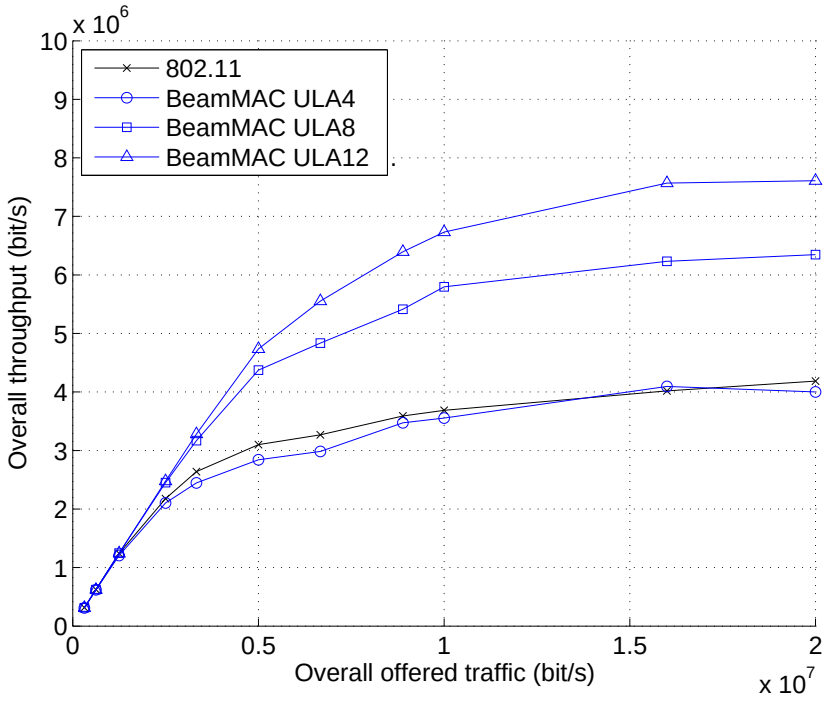
### 5.5.1 Single-Hop Scenarios

Let us begin with single-hop scenarios with $N = 50$ nodes being randomly and homogeneously placed on a network area of $500\,\mathrm{m} \times 500\,\mathrm{m}$. It is assumed that, in a given scenario, all nodes have the same antenna configuration, either UCA or ULA antennas with $m = 4, 8$ or 12 antenna elements.

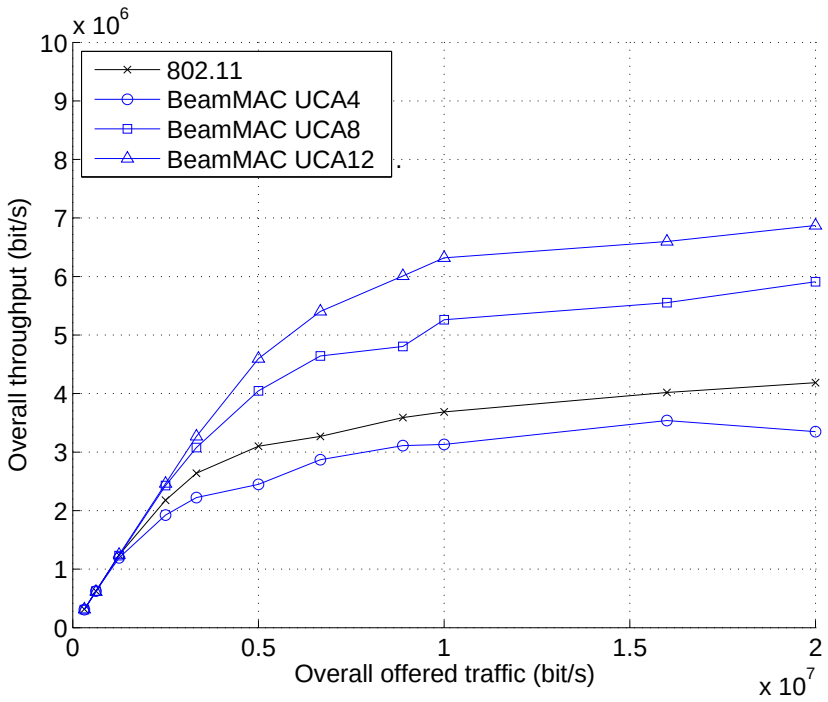
All nodes generate traffic with a fixed packet size of $N_b = 8192$ bit. The average packet inter arrival rate is varied in order to generate different traffic loads in the network. All traffic is single-hop, destined to a randomly chosen neighbor.

The results obtained for ULA and UCA antennas are shown in Fig. 5.4. The improvements over the non-beamforming 802.11 protocol are, with saturated traffic, roughly a factor of 1.5 for ULA8, 1.8 for ULA12, 1.4 for UCA8, and 1.6 for UCA12 antennas.

For antennas with $m = 4$ antenna elements, no improvements or even worse performance is observed. In this case, the beamforming does not allow for a significantly better spatial reuse of radio resources. Another reason is the fact that `ANN` messages are not acknowledged by some sort of CTS message as with 802.11. This is to say that, while BeamMAC effectively protects *ongoing* transmissions, there are cases in which a receiver – after successfully receiving an `ANN` message – is affected by interference from the first data block on. This can lead to vain packet transmissions which have to be repeated by the sender. With $m = 8$ and 12 antenna elements, the link budget is




(a) ULA antennas.




(b) UCA antennas.

Figure 5.4: Throughput in the single-hop scenario.

generally high, and it therefore rarely happens that a receiver cannot capture a signal after beamforming to the sender.

There are different possibilities to circumvent weak performance when $m$ is small. A drastic change of BeamMAC would be to include CTS messages. These could be transmitted in a separate slot before or after the OBJ slot, or simply within the OBJ slot to avoid a degradation of the throughput bound. A moderate amendment of BeamMAC would be to include fairness mechanisms into the objection criterion. For instance, a node suffering from existing interference at the start of the first data block could react more aggressively (by sending OBJs) to future ANNs causing interference. This could improve the fairness of bandwidth sharing in a contention area. Another possibility for improvements is to use adaptive nulling at receiving nodes in a precautionary manner. This means that, even when in idle state, nodes nullify the antenna gain toward currently active transmitters, so as to be ready for reliable signal reception. A minor change improving the throughput would be to notify the transmitter about vain packet transmissions. This could be implemented by an OBJ transmitted by the receiver after the first corrupted data block, thereby acting as negative acknowledgment. Such amendments are, however, not analyzed in this work, and may be subject-matter of further research.

### 5.5.2 Multi-Hop Scenarios

For analyzing multi-hop networks, $N = 200$ nodes are placed homogeneously at random on a square system area of $1000\,\mathrm{m} \times 1000\,\mathrm{m}$. It is the same network scenario that was already exemplified in Fig. 4.13 on page 98.

Each node is equipped with a ULA antenna. Different numbers of antenna elements were used in the simulations. For varying the traffic load of the network, only a subset of the nodes generates traffic. These nodes are selected at random, along with a randomly selected sink. Then, shortest paths are computed for each source-sink pair, and the traffic generation is started. The traffic model assumes a continuous flow of packets with a fixed packet size of $N_b = 8192$ bit and negative exponentially distributed packet inter arrival times with an average value of $0.8192\,\mathrm{s}$. The average per-flow source rate is thus 10 kbit/s.

Fig. 5.5 shows the average end-to-end throughput per flow obtained for ULA antennas with $m = 4$, 8, and 12 antenna elements, and compares to 802.11.

With $m = 4$, the same effects as described in the single-hop case occur. If nodes in the center of the network particularly suffer from too aggressive spatial reuse, this can affect the end-to-end data transmission of several flows. This explains the drastic performance degradation in this case.

The results for $m = 8$ and $m = 12$ antenna elements show a significant performance improvement over 802.11 with omnidirectional antennas in terms of the number of sustainable flows. While 802.11 can only handle 20 flows, BeamMAC can virtually serve the data created by 40 flows with ULA8 antennas, and 80 flows with ULA12 antennas.

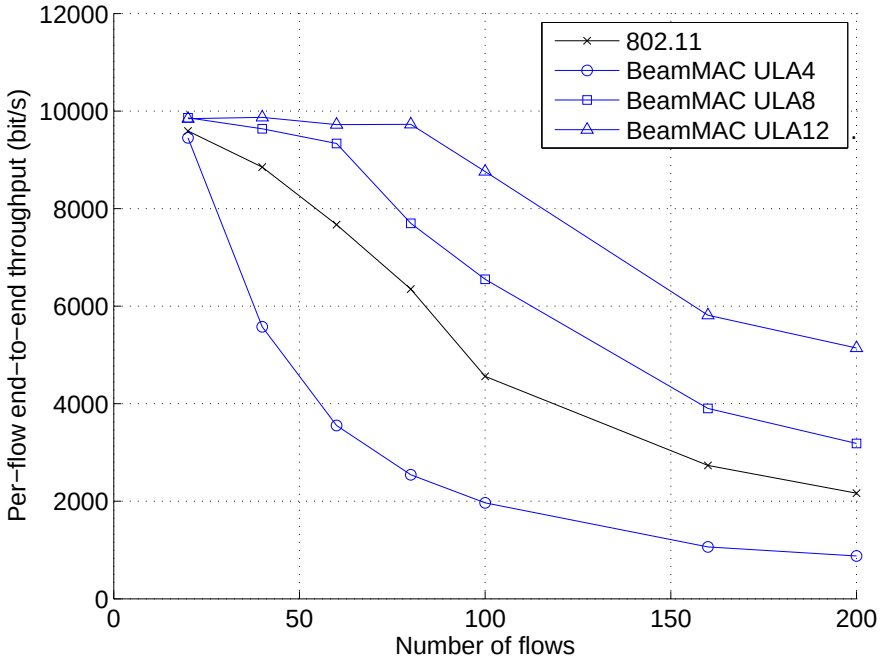
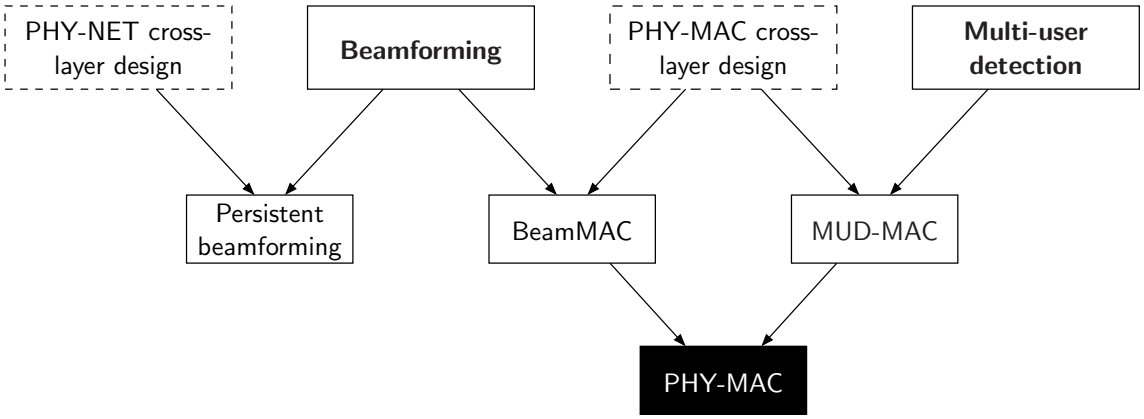



Figure 5.5: Throughput in the multi-hop scenario, varying number of data flows.

## 5.6 Integration of Beamforming Antennas and Multi-User Detection




In Sec. 5.5, we have seen that BeamMAC can be used with various antenna array configurations. In section Sec. 4.7, MUD-MAC was successfully applied in scenarios with heterogeneous MUD capabilities. A consequent question is to ask how both physical techniques could be incorporated in a cross-layer approach.

With the announcement-objection paradigm of this work, the integration of beamforming and multi-user detection is possible. The same access paradigm, signaling

messages, frame structure and backoff mechanisms are used with both BeamMAC and MUD-MAC. The two concepts could therefore be integrated into a protocol supporting different kinds of physical layer signal processing. Let us call such a protocol *PHY-MAC* here.

The integration of beamforming and multi-user detection has two major facets. The first facet is how networks with beamforming capable nodes on the one hand and MUD capable nodes on the other hand can benefit from both techniques, and how the different nodes interact. The second facet concerns nodes having *both* beamforming and MUD capabilities. In this case, the question is how both capabilities can be carried into effect at the same time, or how they are traded off depending on the environment.

The following briefly discusses MAC layer issues of the two facets. A physical layer approach can be found in [MC05], discussing antenna combining joint with multi-user detection.

### Interaction with Beamforming Capable Nodes

When a node without beamforming capabilities wishes to communicate with a node capable of beamforming, this is straight forward with PHY-MAC. Signal capture and decoding is possible in the usual way since beamforming only affects the link budget of a transmission. Furthermore, omnidirectional antennas just resemble yet another (omnidirectional) gain pattern, which is incorporated in the announcement-objection mechanism and the SINR-based objection criterion (5.1) in the same manner as directional patterns. Consequently, no requirements have to be imposed on nodes without antenna arrays in order to communicate with nodes having beamforming capabilities.

### Interaction with MUD Capable Nodes

For MUD-MAC, we assumed transmitter-based spreading codes or interleaver sequences. The same would be the case in a PHY-MAC design. We therefore require each receiver – even nodes with only single-user detectors – to be able to apply any sequence that may be used in the network. Not all simultaneously of course, in the case of limited MUD capabilities. The latter can again be assured by the announcement-objection paradigm. This has been successfully shown in the heterogeneous case studies of Sec. 4.7, verifying requirement **(R6)** of Sec. 4.2.

### Nodes with Multiple Capabilities

Interesting research questions occur when a node has different physical layer capabilities to pick from. For instance, with PHY-MAC, a node may handle interference announced by an `ANN` message either by multi-user detection, or by placing an antenna null in the corresponding direction. The decision on which option to choose may depend on the number of decoding branches already in use, the number of antenna

elements available, the angular direction of the additional interference and the angular direction of the desired signal, and the power levels of the different signal components. Complexity and energy efficiency aspects may also be relevant.

The interplay between beamforming and MUD can be based on the expected performance. For instance, MUD may be preferred if the relative strength of interferers is high, and MUD is likely to perform well. On the other hand, interference suppression by adaptive beamforming upon receiving an `ANN` (or, by a transmitter, upon receiving an `OBJ`) may be preferred over MUD when the angular spread of the incident signal is expected to be small, and when the angular direction can be determined reliably. It may also be preferred when the received signal is equally strong as interfering signals, which is when MUD has difficulties canceling interference. The DOA estimation for adaptive nulling may then still be reliable, since it is performed upon receiving an `ANN`, not during the data slot. Using `ANN` minislots for PHY-MAC (similar to MUD-MAC and BeamMAC), the collision probability for `ANN` messages is low, leading to favorable SINR conditions for DOA estimation.

Further work is necessary to formulate the tradeoff between beamforming and multi-user detection more thoroughly. Results of such work could lead to respective objection criteria, which advance over those proposed in this work.

Finally, let us note that there is also a connection between the persistent beamforming schemes of chapter 3 and PHY-MAC. In networks with low node density, it may be preferable to use persistent beamforming for connectivity improvements, and only use multi-user detection for interference cancellation on a packet basis. On the other hand, if connectivity is not an issue, both MUD and beamforming capabilities may be used for dealing with interference.

## 5.7 Related Work

The MAC protocols proposed for beamforming antennas, including BeamMAC, aim at increasing the capacity of multi-hop networks. Existing work on the capacity of multi-hop networks, such as [GK00, TG03], has to make quite simplifying assumptions (e.g. in terms of the signal collision model), or is limited to particular network topologies (e.g. linear or circular node setups). This is due to the wide range of aspects having an effect on throughput and capacity. There is not even a definite notion of "capacity". It may be defined as single-hop or multi-hop measure, consider or not consider node distances, could possibly be based on outage probabilities, etc. Also, there are no results on absolute capacity limits as available for individual links since Shannon's theory of communication [Sha48]. Since there is no commonly accepted and comprehensive mathematical foundation for the capacity of multi-hop networks, it is difficult to estimate the effects of beamforming.

Cautious results in [PS03] estimate a maximum increase of the stable throughput in the order of $\log^2(n)$, with $n$ being the number of transmitter-receiver pairs. This demands complex signal processing with multiple beams both at transmitters and receivers. The authors achieved this result by formulating a multi-commodity flow problem.

The authors of [YPK03] and [YPKAS07] extend the work in [GK00] to beamforming antennas, modeled as ideal sector antennas with a beam width of $\alpha_t$ and $\alpha_r$ at the transmitter and the receiver, respectively. Their analysis suggests a capacity gain of $\sqrt{\frac{2\pi}{\alpha_t \alpha_r}}$ for arbitrary networks, and $\frac{4\pi^2}{\alpha_t \alpha_r}$ for random networks.

In [SR03b], the authors report capacity gains of multi-hop networks using beamforming by a factor of up to about $20$. These gains depend on the beam width and suppression ratio for the keyhole model, and the number of antenna elements for ULA antennas. In [SR03a], the same authors study the asymptotic capacity behavior for the keyhole antenna model, phased array antennas and fully adaptive antennas. These antenna models are combined with the methodology of [GK00]. The authors give upper bounds for sustainable transmission rates for different collision models. They conclude that the increase in the number of antenna elements necessary to improve the scaling laws may not always be feasible.

To sum up, beamforming antennas are considered to improve the capacity of multi-hop networks. They are, however, not expected to circumvent their general scaling problem.

## 5.8 Summary

Beamforming antennas can improve the efficiency of multi-hop networks. This chapter provided a detailed survey of MAC protocols for beamforming antennas, which is a research topic that received considerable attention since around the year 2000. In general, a fundamental tradeoff between spatial reuse and packet collisions resides in the surveyed protocols. From this research, several open question remained. This is in particular due to critical assumptions often made, such as static antenna orientations or idealistic antenna patterns.

As a second contribution, we identified eight requirements for beamforming that influence MAC layer design. We found that time synchronization, DOA estimation, exploitation of DOA information, avoiding the exchange of explicit beamforming patterns, applicability to NLOS and multi-path channels, the possibility to use main lobe pointing and adaptive nulling, the interoperability of heterogeneous antenna configurations, and the protection of nodes without beamforming capabilities are important design criteria.

The third contribution was the application of the announcement-objection paradigm to beamforming antennas, resulting in the BeamMAC protocol. It is a MAC protocol for self-organizing multi-hop networks, designed to meet the requirements defined in this chapter.

The throughput performance of BeamMAC was analyzed as a fourth contribution. In both single-hop and multi-hop networks, throughput improvements over 802.11 with omnidirectional antennas have been observed. For beamforming with $m = 8$ antenna elements, these improvements amount to a factor of about 1.5, and about 1.7 for $m = 12$. With only $m = 4$ elements, vain packet transmissions due to high spatial

reuse have a deteriorating effect on the throughput, in particular in the multi-hop case. Further work is necessary in this connection. Several possibilities to circumvent this problem can be incorporated into BeamMAC, as discussed in Sec. 5.5.1.

The fifth and final contribution of this chapter was a discussion of how beamforming and multi-user detection can be integrated in a cross-layer framework. This is left to future research.

Further work should also comprise the exploitation of adaptive nulling, by nullifying the antenna gains in undesired directions instead of transmitting OBJs messages. For instance, in Fig. 5.3(b), alternatively to issuing an OBJ $n_2$, may re-adapt its antenna weights upon receiving the ANN and nullify the antenna gain toward $n_3$, the potential source of inference (requirement **(R6)**). By this sort of cooperative behavior, the two transmissions $n_1 \rightarrow n_2$ and $n_3 \rightarrow n_4$ would then be possible simultaneously.

In the analysis of BeamMAC, we did not consider extended transmission ranges in order to analyze throughput effects isolated from topological effects of beamforming. This should be explored further. Throughout this chapter, known angular directions of the receiver were assumed. This is a valid assumption for continuous data transmission, but this information may not be available at the beginning of a communication. As discussed earlier, ANN messages are simply transmitted omnidirectionally in this case with BeamMAC. The effect on the throughput could be analyzed in future work, in particular for mobile scenarios. A necessity of further research is also with respect to a performance evaluation in NLOS environments (requirement **(R5)**), and a delay analysis in the context of QoS requirements.

# 6 Conclusions

This work contributed to the research on multi-hop networks by considering two key signal processing techniques from a networking perspective: adaptive beamforming and multi-user detection.

As a first main contribution, we showed that network layer information can effectively be used to adapt the beamforming characteristic of antenna arrays so as to improve connectivity and throughput. The resulting approach to persistent (i.e. long-term) beamforming is significantly easier to implement than per-packet beamforming, since it does not affect MAC functionality, and does not require (frequent) DOA estimation. The proposed distributed schemes lead to a path probability close to $1$ in different homogeneous and clustered network scenarios, and also significantly increase the number of node disjoint paths. We also proposed a persistent approach to adaptive interference suppression. Its beneficial impact on interference directly translates into promising throughput improvements on a network scale. Results on optimal beamforming configurations described in this work can be used as performance bounds for distributed approaches. These optimizations could be extended in further work to not only optimize node degrees, but also the path probability.

As another contribution, we broke new ground in applying emergent behavior to the partitioning of multi-hop networks. We showed that simple local interactions can effectively identify network partitions. The application of emergent behavior to beamforming and for improving the connectivity of multi-hop networks is left to future work.

The second main topic of this thesis was the application of multi-user detection to multi-hop networks. We answered questions with respect to basic requirements for MUD in this type of network, contributed by MUD-MAC as novel MAC layer framework for MUD, and showed that throughput improvements by a factor of about $2.4$ are possible with MUD-MAC. We demonstrated the applicability of MUD-MAC to networks with heterogeneous receiver complexity, and highlighted aspects of future work.

The third contribution was the application of the MUD-MAC paradigm to beamforming antennas, and – eventually – to networked nodes with both beamforming and multi-user detection. The resulting BeamMAC protocol was shown to effectively improve the spatial reuse of radio resources when realistic antenna models with high directivity are used.

A major task of further research is the integration of beamforming and multi-user detection. Since different approaches are conceivable as shown in this work, the question is how these techniques could be used in an optimal manner.

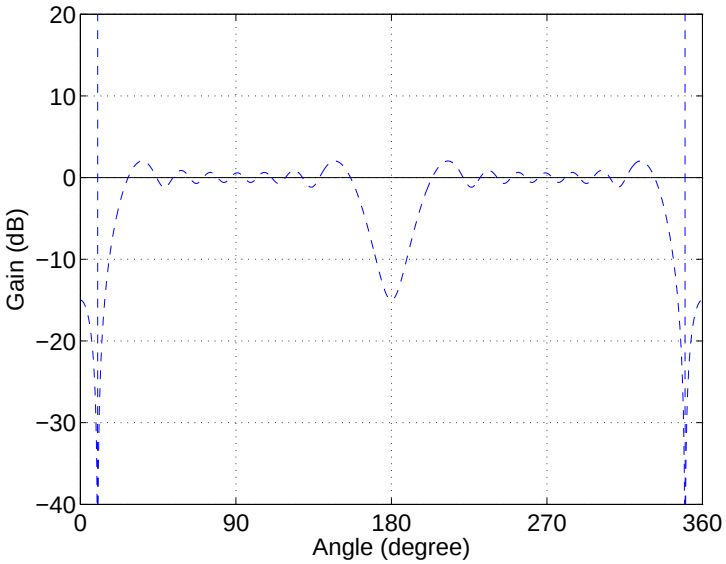
The work on beamforming has assumed that all devices are equipped with antenna arrays. This is not possible for all wireless products, in particular small hand-held

hardware. However, prototype work has already shown that ULA antennas for beamforming can be incorporated in laptop computers [LFE07].

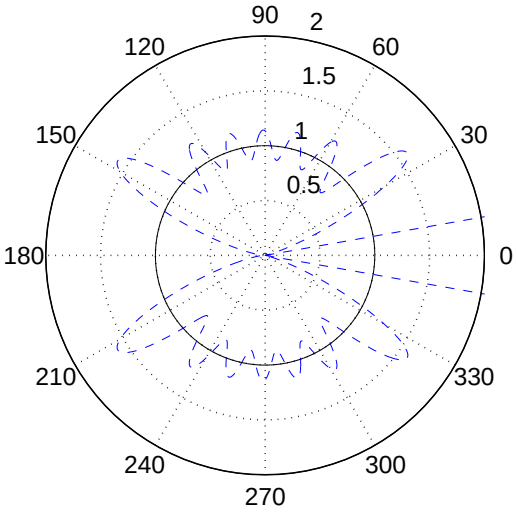
Recent work on advanced signal processing in multi-hop networks shows a strong tendency toward solutions requiring network synchronization (e.g. [SdV04, TG06, CLZ06]). It would be interesting to conduct further research analyzing whether this is a generally beneficial or unavoidable paradigm.

# A Sample Antenna Patterns

### Linear Array Antennas

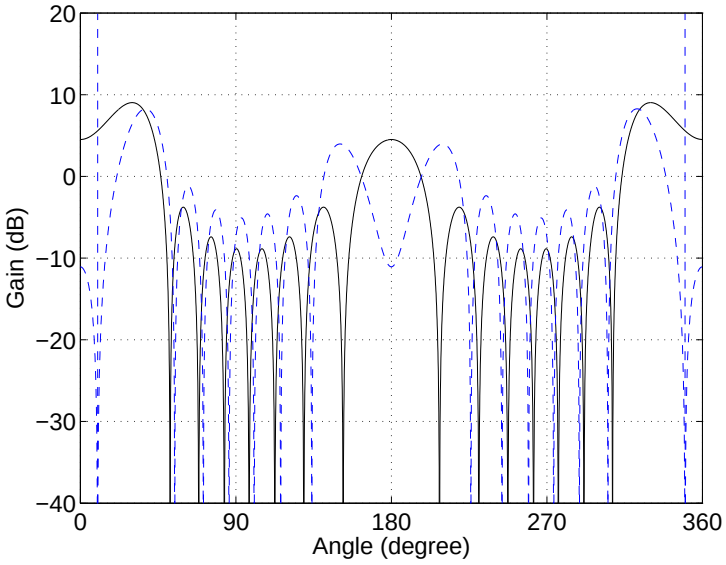



(a) Gain in dB.

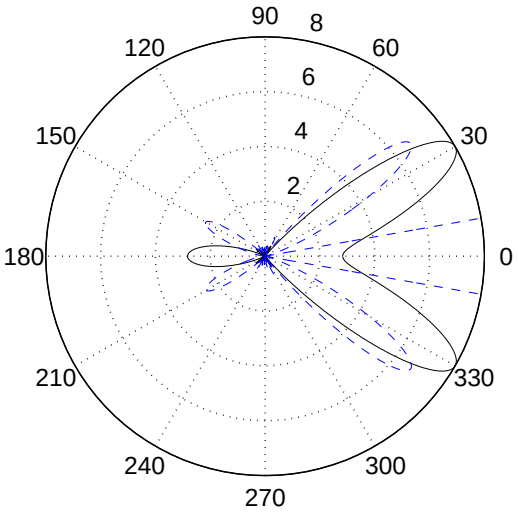



(b) Linear Gain.

Figure A.1: Quiescent (solid) and approximate (dashed) pattern of ULA, $m = 8$ elements, omnidirectional quiescent pattern, null direction $10^\circ$ (dashed line).



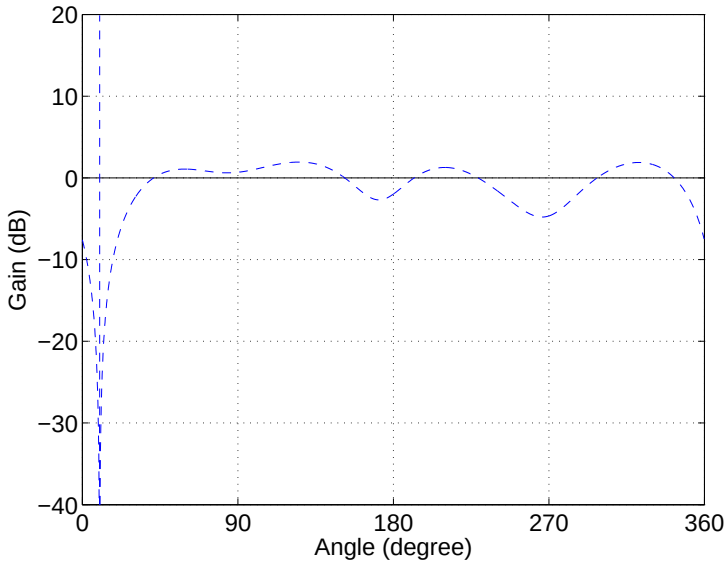

(a) Gain in dB.



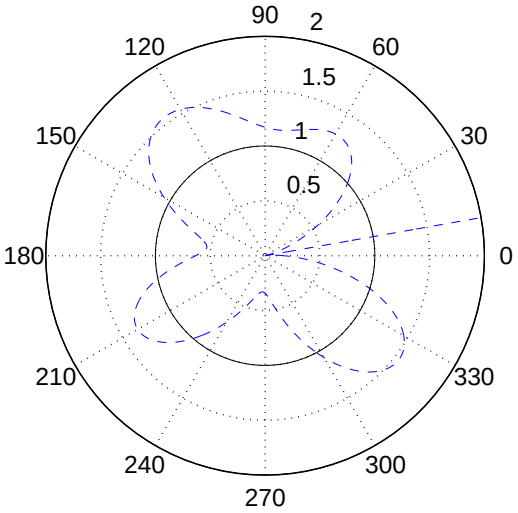

(b) Linear Gain.

Figure A.2: Quiescent (solid) and approximate (dashed) pattern of ULA, $m = 8$ elements, beamforming direction $\theta_0 = 30°$, null direction $10°$ (dashed line).

### Circular Array Antennas

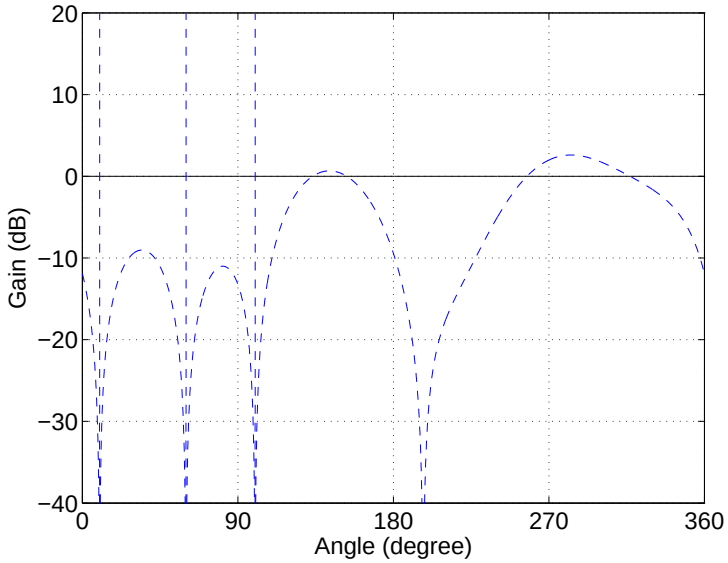



(a) Gain in dB.

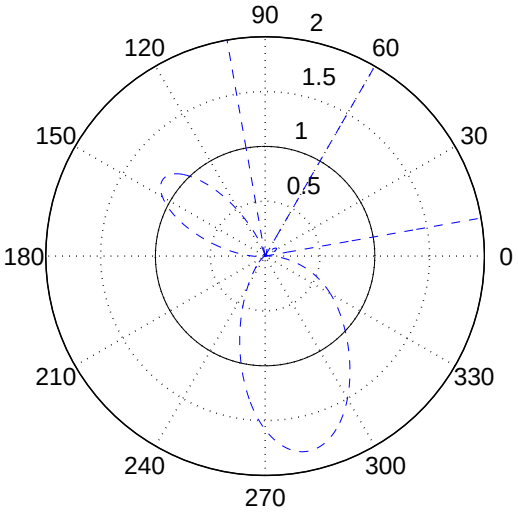



(b) Linear Gain.

Figure A.3: Quiescent (solid) and approximate (dashed) pattern of UCA, $m = 4$ elements, omnidirectional quiescent pattern, null direction $10^\circ$ (dashed line).

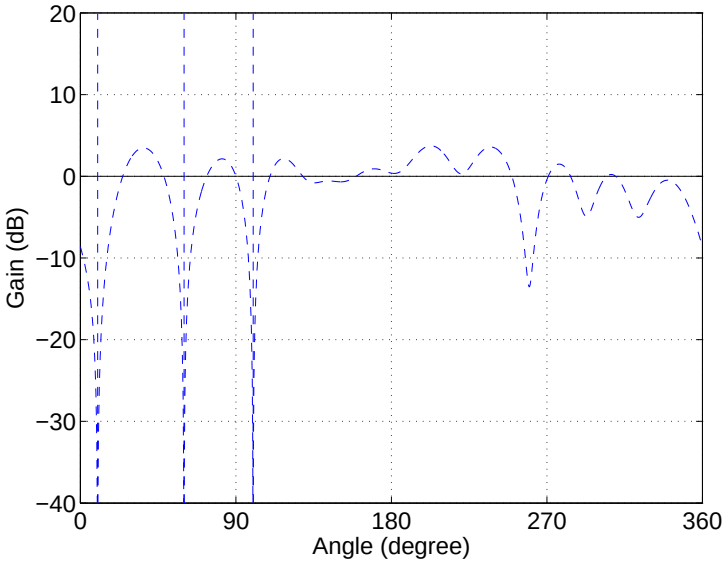



(a) Gain in dB.

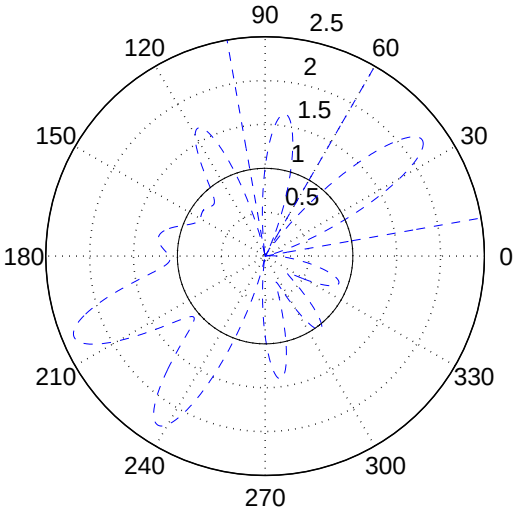



(b) Linear Gain.

Figure A.4: Quiescent (solid) and approximate (dashed) pattern of UCA, $m = 4$ elements, omnidirectional quiescent pattern, null directions $10°$, $60°$, $100°$ (dashed lines).

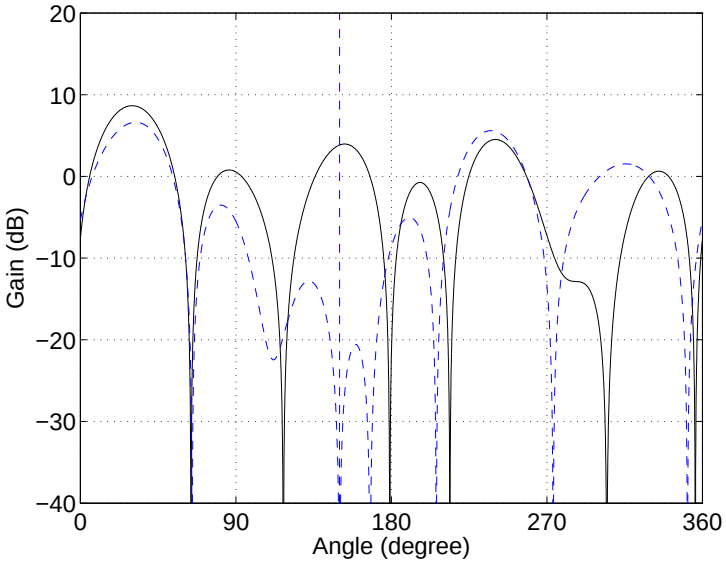



(a) Gain in dB.

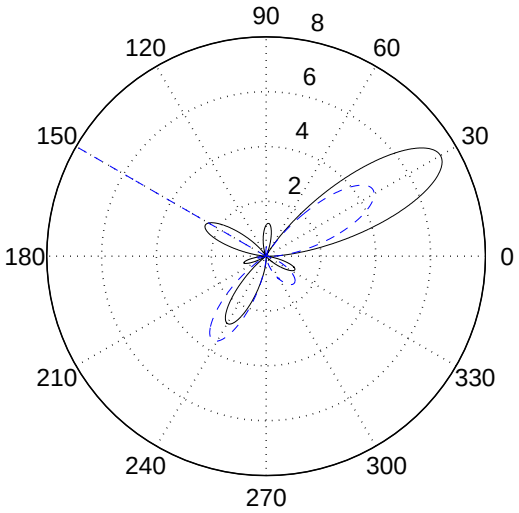



(b) Linear Gain.

Figure A.5: Quiescent (solid) and approximate (dashed) pattern of UCA, $m = 12$ elements, omnidirectional quiescent pattern, null directions $10°$, $60°$, $100°$ (dashed lines).




(a) Gain in dB.




(b) Linear Gain.

Figure A.6: Quiescent (solid) and approximate (dashed) pattern of UCA, $m = 8$ elements, beamforming direction $\theta_0 = 30°$, null direction $150°$ (dashed line).

# B Linear Formulation of the Partitioning Problem

The partitioning constraint matrix $\mathbf{A}$ of dimension $N^2 \times (p \cdot N^2)$ is, with unity matrices $\underset{\sim}{\mathbf{1}}$ of dimension $N^2 \times N^2$:

$$\mathbf{A} = \begin{pmatrix} \underset{\sim}{\mathbf{1}} & \underset{\sim}{\mathbf{1}} & \dots & \underset{\sim}{\mathbf{1}} \end{pmatrix}. \tag{B.1}$$

Additional optimization constraint matrices exemplified for a network with $N = 4$ nodes:

$$\mathbf{A}' = \left(\begin{array}{cccc|cccc|cccc|cccc} 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \end{array}\right), \tag{B.2}$$

$$\mathbf{A}'' = \left(\begin{array}{cccc|cccc|cccc|cccc} 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right). \tag{B.3}$$

The solution variables $\overline{\mathbf{x}}_{(i)}$ are constructed from $\mathbf{x}_{(i)}$ by concatenating $N$ duplicates of $\mathbf{x}_{(i)}$ and setting, for each zero entry of $\mathbf{x}_{(i)}$, the corresponding duplicate to zeros. For instance, when nodes 2 and 4 are assigned to partition $i$ in a four node network, then

$$\mathbf{x}_{(i)} = \begin{pmatrix} 0 & 1 & 0 & 1 \end{pmatrix}^{\mathrm{T}}, \text{ and} \tag{B.4}$$

$$\overline{\mathbf{x}}_{(i)} = \left(\begin{array}{cccc|cccc|cccc|cccc} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{array}\right)^{\mathrm{T}}. \tag{B.5}$$

# C Simulation Software

In the course of this thesis, a comprehensive software tool for the simulation of multi-hop networks was developed. It is an integrated tool that was used for all simulation-based analysis of topology, path life time, and protocol-related issues of this work. Most of the illustrations of this work, in particular those containing beamforming patterns, have been directly exported from the simulator. Only the numerical optimizations were computed with separate tools.

The major goal was to implement realistic interference, signal capture, beamforming, and multi-user detection models. Existing open simulators, such as the Network Simulator ns-2, have deficiencies or missing functionality in this respect, and were therefore not used. The developed simulator does, however, not implement multi-path channels.

The simulator is written in C++, an object-oriented high-level programming language. Overall, 41 classes and 33,500 lines of C++ code have been developed for the simulator.

The simulator is designed such that it can be compiled on Linux, Windows and Mac OS platforms. The following libraries are used:

- Qt (Trolltech), used for the graphical user interface, network visualization, and XML file input/output,
- GSL, used for collecting simulation data, and
- Matlab, used for bit level simulations.

The Matlab code for bit level simulations was developed by a research partner. It is integrated into network level simulations by calling the Matlab engine during simulation. Using this call, near-far factors, the data block length, the modulation scheme, spreading factors, and multi-user detection parameters are supplied to Matlab, which reports packet error information back to C++.

Matlab was further used for post-processing of simulation data, such as the computation of averages and confidence intervals. For the implementation of random processes, the Mersenne Twister random number generator [MN98] was used.

### Software Structure

The basic structure of the simulator is illustrated in Fig. C.1, showing the relationship between its main C++ objects.

The event-based simulation is handled by one instance of `CSimulator`, which schedules a queue of `CEvent` objects. `CSimulator` manages a set of `CNode` objects, which implement one wireless device each. For graph-related algorithms, such as shortest path tree computations, a `CGraph` object is constructed from this set of nodes. The

class `CSimDialog` is used to create a user interface, constructing a graph visualization using `CVisualGraph`.

All the modeling of wireless communication is implemented in `CNode` and its member objects. The main OSI layers are implemented by a respective class. Various options exist for each layer, as illustrated in Fig. C.2. Since MAC protocols typically use backoff mechanisms, a separate class `CBackoff` exists. The beamforming algorithms are implemented in `CPattern`, and beamforming state information is maintained in `CBfLayer`.

Teletraffic is generated by `CNode` objects as data flows. When a node not only creates own traffic but also forwards other traffic (multi-hop communication), then it maintains a plurality of `CFlow` objects, one for each end-to-end communication. Per-flow statistics, such as packet drop rates or packet delays, are stored in `CFlowStats`.

Translatory and rotary motion is implemented in inheriting functions of `CMobility` and `CRotation`, respectively. Since each node instantiates one of these classes, scenarios with heterogeneous mobility patterns, such as basestation and fixed relay scenarios with static nodes, can be simulated.

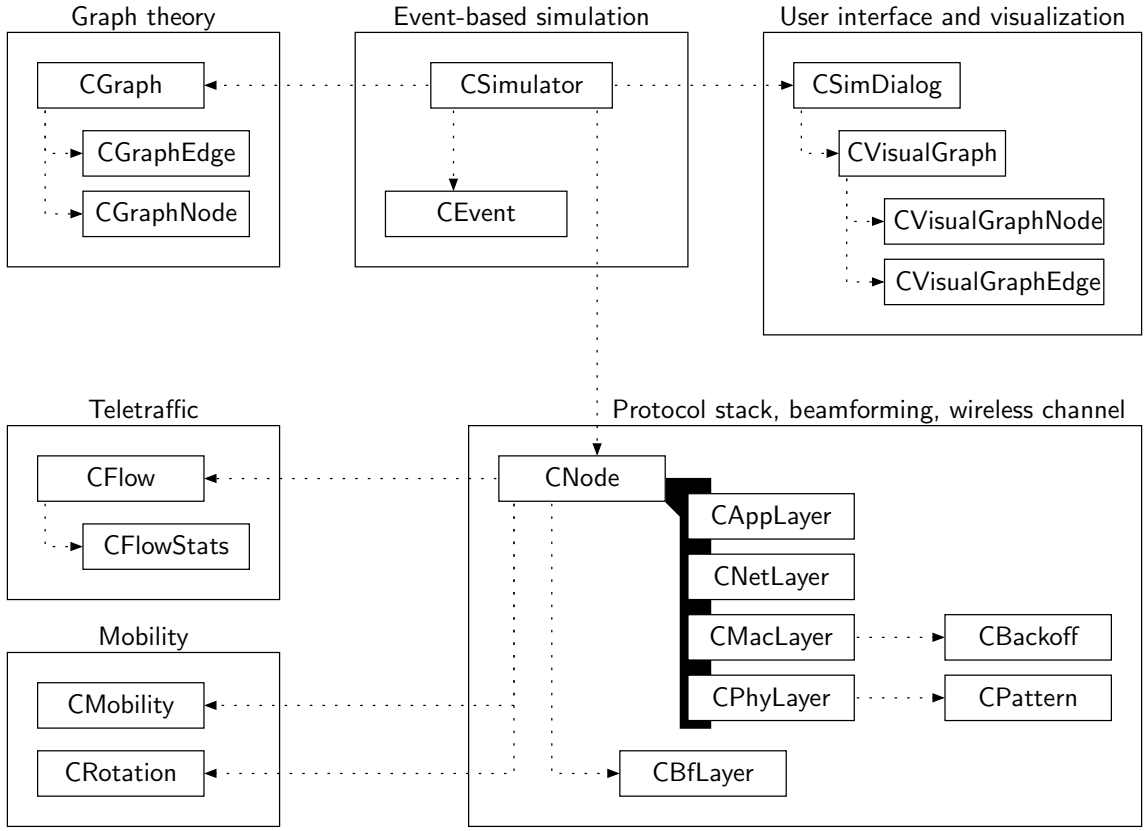



Figure C.1: Main C++ classes of the simulator. Arrows indicate instantiation of objects.

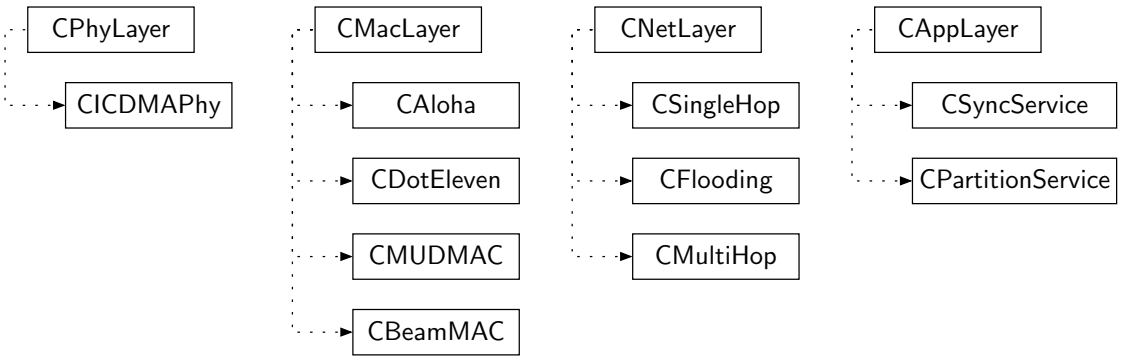



Figure C.2: Base classes and inheriting classes implementing OSI layers. Arrows point to inheriting classes.

### Parameterization

The simulation can be parameterized by 155 different command line arguments. They specify the simulation scenario, beamforming parameters, channel and transceiver parameters, power levels, the selection of OSI layer implementations, traffic parameters, MAC protocol parameters, mobility model parameters, general simulation settings, and options for visualization, debugging, and data output.

### Sample Topology File

The following is an example for a topology file that can be supplied to the simulator. It is the very topology file that was used for the first scenario of Fig. 5.3 on page 120.

It specifies node IDs (`id`), node coordinates (`x`, `y`), numbers of MUD decoder branches (`mud`), antenna types (`a`), numbers of antenna elements (`m`), orientations of antenna arrays relative to the x-axis (`arraydir`), and data flows (`flow`).

Each data flow is characterized by the node ID of the sink (`sink`) and the traffic type (`type`). By indicating a traffic type other than `BEST_EFFORT` (`BROWSING`, `DOWNLOAD`, `VIDEO_STREAM`, `VOIP_LQ`, `VOIP_HQ`, or `VIDEO_TEL`), hard-coded values for the traffic model (packet size and packet inter arrival time statistics) and QoS constraints (in terms of bit rates and latencies) can be specified.

```
<!DOCTYPE scenarioML>
<scenarioML>
    <node>
        <id>1</id>
        <x>255</x>
        <y>220</y>
        <mud>1</mud>
        <a>1</a>
        <m>8</m>
        <arraydir>0.0</arraydir>
        <flow>
            <sink>2</sink>
```

```
            <type>BEST_EFFORT</type>
        </flow>
    </node>
    <node>
        <id>2</id>
        <x>335</x>
        <y>220</y>
        <mud>1</mud>
        <a>1</a>
        <m>6</m>
        <arraydir>5.0</arraydir>
    </node>
    <node>
        <id>3</id>
        <x>330</x>
        <y>180</y>
        <mud>1</mud>
        <a>2</a>
        <m>4</m>
        <arraydir>70.0</arraydir>
        <flow>
            <sink>4</sink>
            <type>BEST_EFFORT</type>
        </flow>
    </node>
    <node>
        <id>4</id>
        <x>275</x>
        <y>255</y>
        <mud>1</mud>
        <a>1</a>
        <m>4</m>
        <arraydir>0.0</arraydir>
    </node>
</scenarioML>
```

# D Symbols and Abbreviations

## Symbols

| | |
|---|---|
| $a'_{n_1,n_2},\ a''_{n_1,n_2}$ | indicator variables |
| $A$ | magnitude of antenna weight |
| $\mathbf{A},\ \mathbf{A}',\ \mathbf{A}''$ | constraint matrices |
| $b$ | number of data blocks per data packet |
| $b_m^{(k)}$ | information bits transmitted by node $n_k$ |
| $\widetilde{b}_m^{(k)}$ | estimated information bits received from node $n_k$ |
| $b_{n_1,n_2}$ | indicator variables |
| $b'_{n_1,n_2},\ b''_{n_1,n_2}$ | indicator variables |
| $\mathbf{c}$ | mesh distance vector |
| $c_{i,j}^{(k)}$ | code word after interleaving |
| $c_{i,j}^{\prime(k)}$ | code word |
| $c'_{n_1,n_2},\ c''_{n_1,n_2}$ | indicator variables |
| $c_1,\ c_2,\ c_3$ | algorithm parameters |
| $d$ | network diameter |
| $d_{h,n_i,n_j}$ | hop distance between node $n_i$ and node $n_j$ |
| $d_{\max}$ | critical distance |
| $d_{\min}$ | critical distance |
| $d_{n_i,n_j}$ | physical distance between node $n_i$ and node $n_j$ |
| $d_r$ | maximum transmission range |
| $d_0$ | reference distance |
| $D$ | random variable for the network diameter |
| $E_b$ | energy per bit |
| $f$ | number of hops needed to flood a network |
| $f(\theta,\phi)$ | radiation pattern of an antenna element |
| $F$ | random variable for number of hops needed to flood a network |
| $F(\theta,\phi)$ | array factor |
| $g$ | linear antenna gain |
| $g_0$ | quiescent pattern |
| $g_a$ | approximate pattern |

| | |
|---|---|
| $G$ | antenna gain in dBi |
| $G_M$ | antenna gain of main lobe in dBi |
| $G_S$ | antenna gain of side lobe in dBi |
| $h$ | hop distance |
| $h_{\mathrm{TTL}}$ | time-to-live value for mobility detection |
| $h_{l,i}^{(k)}$, $\boldsymbol{h}_{l,i}^{(k)}$ | channel taps |
| $h_{n_i,n_j}$ | channel coefficient between node $n_i$ and node $n_j$ |
| $H$ | random variable for the hop distance |
| $J_e$ | criterion function |
| $k_{n_i,n_j}$ | number of node disjoint paths between node $n_i$ and node $n_j$ |
| $K$ | number of decoder branches |
| $l_{n_i,n_j}$ | binary indicator variable for link between node $n_i$ and node $n_j$ |
| $L$ | order of finite length channel |
| $m$ | number of antenna elements |
| $M_1$ | number of antenna gain nulls |
| $M_2$ | order of antenna gain nulls |
| $M$ | variable for mobility detection |
| $n_A$ | number of announcement slots per frame |
| $n_i$ | node with ID $i$ |
| $n_K$ | number of main lobe directions in main lobe sweeping |
| $n_M$ | number of announcement minislots |
| $N$ | number of nodes |
| $\mathcal{N}$ | set of nodes |
| $N_b$ | packet size in bit |
| $N_0$ | noise power spectral density |
| $p$ | receive power of desired signal / number of partitions |
| $p_{\mathrm{ANN}}$ | power of ANN message |
| $p_I$ | interference power |
| $p_{\mathrm{known}}$ | sum of interference power from known sources |
| $p_{\mathrm{Path}}$ | path probability |
| $p_r$ | receive power |
| $p_t$ | transmit power |
| $p_{\mathrm{unknown}}$ | sum of interference power from unknown sources |
| $p_z$ | noise power |
| $\mathcal{P}$ | set of nodes constituting a partition |
| $P_{L,n_i,n_j}$ | path loss between node $n_i$ and node $n_j$ in dB |
| $P_r$ | receive power in dB |

| | |
|---|---|
| $P_{r,\min}$ | minimum required receive power in dB |
| $P_t$ | transmit power in dB |
| $r_i,\ \boldsymbol{r}_i,\ r(t)$ | received signal |
| $R$ | radius of UCA |
| $R_C$ | control packet transmission bit rate |
| $R_D$ | data transmission bit rate |
| $\widehat{R}_{\mathrm{MUD-MAC}}$ | throughput bound of MUD-MAC |
| $\widehat{R}_{802.11}$ | throughput bound of IEEE 802.11 |
| $s(t)$ | transmitted signal |
| $s_{n_i,n_j}$ | mesh distance between node $n_i$ and node $n_j$ |
| $\mathbf{S}$ | mesh distance matrix |
| $\mathrm{SINR}_{\mathrm{eff}}$ | effective SINR |
| $\mathrm{SINR}_{\mathrm{exp}}$ | expected SINR |
| $s_i^{(k)},\ s^{(k)}(t)$ | transmit signal of node $n_k$ |
| $\tilde{s}_i^{(k)},\ \widetilde{s}^{(k)}(t)$ | estimated transmit signal of node $n_k$ |
| $t_{\mathrm{ACK}}$ | ACK duration |
| $t_{\mathrm{ANN}}$ | ANN duration |
| $t_{\mathrm{CTS}}$ | CTS duration |
| $t_{\mathrm{DATA}}$ | data packet duration |
| $t_{\mathrm{DIFS}}$ | DIFS duration |
| $t_{\mathrm{IFS}}$ | IFS duration |
| $t_{\mathrm{OBJ}}$ | OBJ duration |
| $t_{\mathrm{RTS}}$ | RTS duration |
| $t_{\mathrm{SIFS}}$ | SIFS duration |
| $t_{\mathrm{sweep}}$ | sweep duration |
| $T$ | time interval |
| $v,\ \widetilde{v}$ | variables for mobility detection |
| $x$ | node coordinate |
| $\mathbf{x}$ | antenna weight vector / partition membership indicator vector |
| $\overline{\mathbf{x}}$ | equivalent partition membership indicator vector |
| $y$ | node coordinate |
| $\boldsymbol{z}_i,\ z(t)$ | noise signal |
| $\alpha$ | path loss exponent / main lobe aperture |
| $\beta_{n_i,n_j}$ | absolute angular direction from node $n_i$ to node $n_j$ |
| $\delta$ | node degree |
| $\delta_c$ | connectivity criterion |
| $\Delta$ | phase difference |

| | |
|---|---|
| $\epsilon(g_a)$ | error of approximate pattern with respect to quiescent pattern |
| $\phi$ | elevation angle |
| $\gamma_{n_i}$ | main lobe direction of node $n_i$ |
| $\lambda$ | carrier wavelength |
| $\Pi_p$ | partitioning of an network into $p$ partitions |
| $\theta$ | azimuthal angle |
| $\Omega$ | number of topologies |
| $\xi$ | state information determining partition membership |
| $\widetilde{\xi}$ | auxiliary state information |
| $\xi_r$ | received state information |
| $\widetilde{\xi}_r$ | received auxiliary state information |

## Notation

| | |
|---|---|
| $E\{\bullet\}$ | Expected value of a random variable |
| $iff$ | if and only if |
| $p_i(x)$ | histogram value for $x = i$ of a random variable $X$ |
| $P(X = x)$ | probability mass function of a random variable $X$ |
| $\|\bullet\|$ | Euclidean vector norm |

## Abbreviations

| | |
|---|---|
| **ACK** | Acknowledgment |
| **ADC** | Analog to Digital Conversion/Converter |
| **AWGN** | Additive White Gaussian Noise |
| **CCDF** | Complementary Cumulative Distribution Function |
| **CDMA** | Code Division Multiple Access |
| **CTS** | Clear to Send |
| **DCF** | Distributed Coordination Function |
| **DIFS** | Distributed Coordination Function Interframe Space |
| **DNAV** | Directional Network Allocation Vector |
| **DOA** | Direction of Arrival |
| **DOF** | Degree of Freedom |
| **DQPSK** | Differential Quadrature Phase Shift Keying |
| **DSSS** | Direct Sequence Spread Spectrum |
| **FEC** | Forward Error Correction |
| **FIFO** | First In First Out |

| | |
|---|---|
| **GPS** | Global Positioning System |
| **GSM** | Global System for Mobile Communications |
| **ID** | Identifier |
| **IDMA** | Interleave Division Multiple Access |
| **IEEE** | Institute of Electrical and Electronics Engineers |
| **IFS** | Interframe Space |
| **ILP** | Integer Linear Program |
| **ISI** | Inter-Symbol Interference |
| **LOS** | Line of Sight |
| **MAC** | Medium Access Control |
| **MAI** | Multiple Access Interference |
| **MIMO** | Multiple Input Multiple Output |
| **MIP** | Mixed Integer Linear Program |
| **MNDB** | Maximum Node Degree Beamforming |
| **MUD** | Multi-User Detection |
| **NAV** | Network Allocation Vector |
| **NLOS** | Non-Line of Sight |
| **OSI** | Open Systems Interconnection |
| **PER** | Packet Error Rate |
| **QoS** | Quality of Service |
| **QPSK** | Quadrature Phase Shift Keying |
| **RAM** | Random Access Memory |
| **RDB** | Random Direction Beamforming |
| **RTR** | Ready to Receive |
| **RTS** | Request to Send |
| **SDMA** | Space Division Multiple Access |
| **SIFS** | Short Interframe Space |
| **SINR** | Signal to Interference and Noise Ratio |
| **SNR** | Signal to Noise Ratio |
| **TNDB** | Two-hop Node Degree Beamforming |
| **UCA** | Uniform Circular Array |
| **ULA** | Uniform Linear Array |
| **UMTS** | Universal Mobile Telecommunications System |
| **WLAN** | Wireless Local Area Network |

# Bibliography

## Publications by the Author


[EEH+07] J. Eberspächer, S. Eichler, C. Hartmann, S. Meister, R. Nagel, R. Vilzmann, and H. M. Zimmermann. Wireless multi-hop networks: Classification, paradigms and constraints. Technical report, Institute of Communication Networks, Technische Universität München, 2007. 4, 5

[HKV08] C. Hartmann, M. Kiese, and R. Vilzmann. Optimizing the node degree in wireless multihop networks with single-lobe beamforming. In *Proc. INFORMS Telecommunications Conf.*, College Park, MD, USA, March 2008. 14

[KVM+07] K. Kusume, R. Vilzmann, A. Müller, C. Hartmann, and G. Bauch. A multiuser detection perspective on medium access control in ad hoc networks. In *Proc. IEEE Global Communications Conf. (GLOBECOM)*, pages 801–806, November 2007. 74, 77, 80

[KVM+09] K. Kusume, R. Vilzmann, A. Müller, C. Hartmann, and G. Bauch. Medium access in spread spectrum ad hoc networks with multiuser detection. *EURASIP Journal on Advances in Signal Processing (to appear)*, 2009. 74, 77, 80, 81, 103

[VB05] R. Vilzmann and C. Bettstetter. A survey on MAC protocols for ad hoc networks with directional antennas. In *Proc. EUNICE Open European Summer School*, July 2005. 105

[VBH05] R. Vilzmann, C. Bettstetter, and C. Hartmann. On the impact of beamforming on interference in wireless mesh networks. In *Proc. IEEE Workshop on Wireless Mesh Networks (WiMesh)*, Santa Clara, USA, September 2005. 14

[VBH06] R. Vilzmann, C. Bettstetter, and C. Hartmann. BeamMAC: A new paradigm for medium access in wireless networks. *Int. Journal of Electronics and Communications (AEÜ)*, 60(1):3–7, January 2006. 105

[VBMH05] R. Vilzmann, C. Bettstetter, D. Medina, and C. Hartmann. Hop distances and flooding in wireless multihop networks with randomized beamforming. In *Proc. ACM Int. Conf. on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, Montreal, Canada, October 2005. 14

[VKHB07] R. Vilzmann, K. Kusume, C. Hartmann, and G. Bauch. A MAC perspective on multiuser detection in ad hoc networks. In *Proc. Int. Workshop on Cross Layer Design (IWCLD)*, pages 109–112, September 2007. 74

[VWAH06] R. Vilzmann, J. Widmer, I. Aad, and C. Hartmann. Low-complexity

beamforming techniques for wireless multihop networks. In *Proc. IEEE Conf. on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, Reston, USA, September 2006. 14

[VWAH07] R. Vilzmann, J. Widmer, I. Aad, and C. Hartmann. Interferer nulling based on neighborhood communication patterns. In *Proc. Tyrrenhian Int. Workshop on Digital Communication (TIWDC)*, Ischia Island, Napoli, Italy, September 2007. 14

# Other References

[3G 02] 3G TS 25.213. 3rd generation partnership project; technical specification group radio access network; spreading and modulation (FDD) (release 4), June 2002. 87

[Abr70] N. Abramson. The aloha system – another alternative for computer communications. In *Proc. Joint Computer Conf. (AFIPS)*, 1970. 68

[AKKD01] S. Agarwal, R. H. Katz, S. V. Krishnamurthy, and S. K. Dao. Distributed power control in ad-hoc wireless networks. In *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, San Diego, USA, September 2001. 75

[AKM04] A. Arora, M. Krunz, and A. Muqattash. Directional medium access protocol (DMAP) with power control for wireless ad hoc networks. In *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Dallas, USA, November 2004. 13

[App76] S. P. Applebaum. Adaptive arrays. *IEEE Transactions on Antennas and Propagation*, 24:585–598, September 1976. 19

[AWH07] J. G. Andrews, S. Weber, and M. Haenggi. Ad hoc networks: To spread or not to spread? *IEEE Communications Magazine*, 45(12):84–91, December 2007. 82, 100

[Bal97] C. A. Balanis. *Antenna Theory*. Wiley, 2nd edition, 1997. 15

[BB76] J. Buck and E. Buck. Synchronous fireflies. *Scientific American*, 234:74–85, 1976.

[BCJR74] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Transactions on Information Theory*, 20(2):284–287, March 1974. 80

[BE03] C. Bettstetter and J. Eberspächer. Hop distances in homogeneous ad hoc networks. In *Proc. IEEE Vehicular Technology Conf. (VTC)*, Jeju, Korea, April 2003. 33

[Bet03a] C. Bettstetter. *Mobility Modeling, Connectivity, and Adaptive Clustering in Ad Hoc Networks*. PhD thesis, Technische Universität München, Germany, October 2003. 8, 60

[Bet04c] C. Bettstetter. On the connectivity of ad hoc networks. *The Computer Journal*, 47(4):432–447, July 2004. Oxford University Press. 71

[BG96] C. Berrou and A. Glavieux. Near optimum error correcting coding and decoding: Turbo-codes. *IEEE Transactions on Communications*, 44(10):1261–1271, October 1996. 80

[BGLA02] L. Bao and J. J. Garcia-Luna-Aceves. Transmission scheduling in ad hoc networks with directional antennas. In *Proc. ACM Int. Conf. on Mobile Computing and Networking (MobiCom)*, Atlanta, USA, September 2002. ACM Press. 112

[BHM05] C. Bettstetter, C. Hartmann, and C. Moser. How does randomized beamforming improve the connectivity of ad hoc networks? In *Proc. IEEE Int. Conf. on Communications (ICC)*, Seoul, Korea, May 2005. 33, 42, 71, 72

[Bia00] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18:535–547, March 2000. 97

[BT97] D. Bertsimas and J. N. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific Optimization and Computation Series. Athena Scientific, Cambridge, USA, 1st edition, January 1997. 23

[CLZ06] P. Casari, M. Levorato, and M. Zorzi. DSMA: An access method for MIMO ad hoc networks based on distributed scheduling. In *Proc. ACM Int. Conf. on Wireless Communications and Mobile Computing (IWCMC)*, Vancouver, Canada, July 2006. 102, 131

[CR93] S. Chopra and M. R. Rao. The partition problem. *Mathematical Programming*, 59(1):87–115, 1993. 52

[CV03a] R. R. Choudhury and N. H. Vaidya. Deafness: A MAC problem in ad hoc networks when using directional antennas. Technical report, University of Illinois at Urbana-Champaign, July 2003. 110

[CYRV02] R. R. Choudhury, X. Yang, R. Ramanathan, and N. H. Vaidya. Using directional antennas for medium access control in ad hoc networks. In *Proc. ACM Int. Conf. on Mobile Computing and Networking (MobiCom)*, Atlanta, USA, September 2002. ACM Press. 109

[CYRV06] R. R. Choudhury, X. Yang, R. Ramanathan, and N. H. Vaidya. On designing mac protocols for wireless networks using directional antennas. *IEEE Transactions on Mobile Computing*, 5(5), May 2006. 13

[DA93] A. Dolan and J. Aldous. *Networks and Algorithms - An Introductory Approach*. John Wiley & Sons, 1993. 21

[Dan63] G. B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, USA, 1963. 23, 71

[DGNS98] E. Dahlman, B. Gudmundson, M. Nilsson, and J. Sköld. UMTS/IMT-2000 based on wideband CDMA. *IEEE Communications Magazine*, 36(9):70–80, September 1998. 74

[DHS01] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley-Interscience, 2nd edition, 2001. 48, 49

[DM69] C. J. Drane and J. F. McIlvenna. Gain maximization and controlled null placement simultaneously achieved in aerial array patterns. Technical Report

Rep. AFCRL-69-0257, Air Force Cambridge Res. Labs, Bedford, USA, June 1969. 19

[DP97] D. Divsalar and F. Pollara. Hybrid concatenated codes and iterative decoding. Technical Report TDA Progress Report 42-130, California Institute of Technology, Pasadena, California, 1997. 80

[ECS+07] Y. Eisenberg, K. Conner, M. Sherman, J. Niedzwiecki, and R. Brothers. Mud enabled media access control for high capacity, low-latency spread spectrum communications. In *Proc. Military Communications Conf. (MILCOM)*, Orlando, USA, October 2007. 102

[EE04] T. ElBatt and A. Ephremides. Joint scheduling and power control for wireless ad hoc networks. *IEEE Transactions on Wireless Communications*, 3(1):74–85, January 2004. 75

[EGE02] J. Elson, L. Girod, and D. Estrin. Fine-grained network time synchronization using reference broadcasts. In *Proc. Symp. on Operating Systems Design and Implementation (OSDI)*, Boston, USA, December 2002. 11, 12

[EVB01b] J. Eberspächer, H.-J. Vögel, and C. Bettstetter. *GSM — Switching, Services, and Protocols*. Wiley, 2nd edition, March 2001. 1

[Gal85] R. G. Gallager. A perspective on multiaccess channels. *IEEE Transactions on Information Theory*, 31(2):124–142, March 1985. 75

[GK00] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000. 127, 128

[GKS03] S. Ganeriwal, R. Kumar, and M. B. Srivastava. Timing-sync protocol for sensor networks. In *Proc. Int. Conf. on Embedded Networked Sensor Systems (SenSys)*, Los Angeles, USA, November 2003. 11, 12

[GL00] A. C. V. Gummalla and J. O. Limb. Wireless medium access control protocols. *IEEE Communications Surveys & Tutorials*, 3(2), 2000. 84

[GM95] M. Gondran and M. Minoux. *Graphs and Algorithms*. Wiley-Interscience Series in Discrete Mathematics, 1995. 21

[God97] L. C. Godara. Application of antenna arrays to mobile communications, part II: Beam-forming and direction-of-arrival considerations. *Proceedings of the IEEE*, 85(8):1195–1245, August 1997. 108

[GSD03] K.H. Grace, J.A. Stine, and R.C. Durst. An approach for modestly directional communications in mobile ad hoc networks. In *Proc. IEEE Int. Conf. on Computer Communications and Networks (ICCCN)*, Dallas, USA, October 2003. 113

[GW89] M. Grötschel and Y. Wakabayashi. A cutting plane algorithm for a clustering problem. *Mathematical Programming*, 45(1):59–96, 1989. 55

[HA07] A. Hasan and J. G. Andrews. The guard zone in wireless ad hoc networks. *IEEE Transactions on Wireless Communications*, 6(3):897–906, March 2007. 102

[HD02] Z. J. Haas and J. Deng. Dual busy tone multiple access (dbtma)-a multiple ac-

cess control scheme for ad hoc networks. *IEEE Transactions on Communications*, 50(6):975–985, June 2002. 85, 111

[HHH03] C. Hu, Y. Hong, and J. Hou. On mitigating the broadcast storm problem with directional antennas. In *Proc. IEEE Int. Conf. on Communications (ICC)*, Anchorage, USA, May 2003. 13

[HOP96] J. Hagenauer, E. Offer, and L. Papke. Iterative decoding of binary block and convolutional codes. *IEEE Transactions on Information Theory*, 42(2):429–445, March 1996. 80

[HS05a] Y. W. Hong and A. Scaglione. A scalable synchronization protocol for large scale sensor networks and its applications. *IEEE Journal on Selected Areas in Communications*, 23(5):1085–1099, May 2005. 11, 12

[HS05b] A. S. Hu and S. D. Servetto. Algorithmic aspects of the time synchronization problem in large-scale sensor networks. *Mobile Networks and Applications*, 10(4):491–503, 2005. 11, 12

[HSSJ02] Z. Huang, C. C. Shen, C. Srisathapornphat, and C. Jaikaeo. A busy-tone based directional mac protocol for ad hoc networks. In *Proc. Military Communications Conf. (MILCOM)*, Anaheim, USA, October 2002. 111

[HYA03] A. Hasan, K. Yang, and J. G. Andrews. Clustered CDMA ad hoc networks without closed-loop power control. In *Proc. Military Communications Conf. (MILCOM)*, Boston, USA, October 2003. 75

[IEE97] IEEE Std 802.11-1997. Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, June 1997. 89, 94

[JLB04] R. Jurdak, C. V. Lopes, and P. Baldi. A survey, classification and comparative analysis of medium access control protocols for ad hoc networks. *IEEE Communications Surveys & Tutorials*, 6(1):2–16, 2004. 84

[Joh02] S. Johnson. *Emergence: The Connected Lives of Ants, Brains, Cities and Software*. Penguin, 2002. 55

[KB05] K. Kusume and G. Bauch. CDMA and IDMA: Iterative multiuser detections for near-far asynchronous communications. In *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Berlin, Germany, September 2005. 78, 82

[KB06a] K. Kusume and G. Bauch. Cyclically shifted multiple interleavers. In *Proc. IEEE Global Communications Conf. (GLOBECOM)*, San Francisco, USA, November 2006. 85, 87

[KB06b] K. Kusume and G. Bauch. A simple complexity reduction strategy for interleave division multiple access. In *Proc. IEEE Vehicular Technology Conf. (VTC)*, Montreal, Canada, September 2006. 80

[KB06c] K. Kusume and G. Bauch. Some aspects of interleave division multiple access in ad hoc networks. In *Proc. Int. Symp. on Turbo Codes and Related Topics*, Munich, Germany, April 2006. 80

[KDUB07] K. Kusume, G. Dietl, W. Utschick, and G. Bauch. Performance of interleave division multiple access based on minimum mean square error detection. In

*Proc. IEEE Int. Conf. on Communications (ICC)*, Glasgow, Scottland, June 2007. 80

[KJT03] T. Korakis, G. Jakllari, and L. Tassiulas. A mac protocol for full exploitation of directional antennas in ad-hoc wireless networks. In *Proc. ACM Int. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Annapolis, USA, June 2003. ACM Press. 111

[KK05] V. Kawadia and P.R. Kumar. Principles and protocols for power control in wireless ad hoc networks. *IEEE Journal on Selected Areas in Communications*, 23(1):76–88, January 2005. 75

[Kos06] H. Koskinen. Analytical study of connectivity in wireless multihop networks utilizing beamforming. In *Proc. ACM Int. Conf. on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, Terromolinos, Spain, October 2006. 71, 72

[KSV00] Y. B. Ko, V. Shankarkumar, and N. H. Vaidya. Medium access control protocols using directional antennas in ad hoc networks. In *Proc. IEEE Conf. on Computer Communications (INFOCOM)*, Tel-Aviv, Israel, March 2000. 13, 109

[LCZ06] M. Levorato, P. Casari, and M. Zorzi. On the performance of access strategies for MIMO ad hoc networks. In *Proc. IEEE Global Communications Conf. (GLOBECOM)*, San Francisco, USA, November 2006. 102

[LFE07] C. Lu, F. H. Fitzek, and P. C. Eggers. Capacity enhancement by terminal originated beamforming for wireless local area networks. *Wireless Personal Communications*, 43(2):573–587, 2007. 131

[LHL+08] R. Learned, B. Hombs, M. Lande, J. Tranquilli, L. Russo, J. Farkas, J. Niedzwieck, Y. Eisenberg, K. Conner, M. Sherman, L. R. Brothers, B. Pierce, and J. DeBardelaben. Interference multiple access wireless network demonstration enabled by real-time multiuser detection. In *Proc. IEEE Radio and Wireless Symp. (RWS)*, Orlando, USA, January 2008. 102

[LL96] J. Litva and T. K. Lo. *Digital Beamforming in Wireless Communications*. Artech House, Inc., Norwood, USA, 1996. 15, 16, 17

[LR97] J. D. Laster and J. H. Reed. Interference rejection in digital wireless communications. *IEEE Signal Processing Magazine*, 14(3):37–62, May 1997. 76

[LTR+02] D. Lal, R. Toshniwal, R. Radhakrishnan, D. Agrawal, and J. Caffery. A novel MAC layer protocol for space division multiple access in wireless ad hoc networks. In *Proc. IEEE Int. Conf. on Computer Communications and Networks (ICCCN)*, Miami, USA, October 2002. 110

[MA05] S. Mukherjee and D. Avidor. On the probability distribution of the minimal number of hops between any pair of nodes in a bounded wireless ad-hoc network subject to fading. In *Proc. Int. Workshop Wireless Ad-Hoc Networks (IWWAN)*, London, UK, May 2005. 33

[MC05] R. Müller and L. Cottatellucci. *Smart Antennas – State of the art*, chapter Joint antenna combining and multiuser detection, pages 77–96. Hindawi Publishing Corp., 2005. 126

[MKR03] A. Muqattash, M. Krunz, and W. E. Ryan. Solving the near-far problem in CDMA-based ad hoc networks. *Ad Hoc Networks*, 1(4):435–453, 2003. 75

[MKSL04] M. Maróti, B. Kusy, G. Simon, and Á. Lédeczi. The flooding time synchronization protocol. In *Proc. Int. Conf. on Embedded Networked Sensor Systems (SenSys)*, Baltimore, USA, November 2004. 11, 12

[MN98] M. Matsumoto and T. Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 8(1):3–30, 1998. 139

[MP02] R. H. Mahadevappa and J G. Proakis. Mitigating multiple access interference and intersymbol interference in uncoded CDMA systems with chip-level interleaving. *IEEE Transactions on Wireless Communications*, 1(4):781–792, October 2002. 80, 93

[MS90] R. E. Mirollo and S. H. Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics*, 50(6):1645–1662, 1990.

[MT01] G. Mergen and L. Tong. Receiver controlled medium access in multihop ad hoc networks with multipacket reception. In *Proc. Military Communications Conf. (MILCOM)*, Washington, USA, October 2001. 75

[NC06] N. Nie and C. Comaniciu. Energy efficient aodv routing in cdma ad hoc networks using beamforming. *EURASIP Journal on Wireless Communications and Networking*, 2006(2):14–14, 2006. 13

[NMMH00] A. Nasipuri, J. Mandava, H. Manchala, and R. E. Hiromoto. On-demand routing using directional antennas in mobile ad hoc networks. In *Proc. IEEE Int. Conf. on Computer Communications and Networks (ICCCN)*, Las Vegas, USA, October 2000. 13

[NTCS99] S. Y. Ni, Y. C. Tseng, Y. S. Chen, and J. P. Sheu. The broadcast storm problem in a mobile ad hoc network. In *Proc. ACM Int. Conf. on Mobile Computing and Networking (MobiCom)*, Seattle, USA, August 1999. ACM. 34

[NYYH00] A. Nasipuri, S. Ye, J. You, and R. E. Hiromoto. A MAC protocol for mobile ad hoc networks using directional antennas. In *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, Chicago, USA, September 2000. 110

[OWJR05] G. Orfanos, B. Walke, J. W. Jansen, and H. J. Reumerman. Interference aware adaptive MAC protocol for MC-CDMA ad hoc wireless LANs. In *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Berlin, Germany, September 2005. 102

[PB05] C. Prehofer and C. Bettstetter. Self-organization in communication networks: Principles and design paradigms. *IEEE Communications Magazine*, 43(7):78–85, July 2005. 55, 63

[Pen97] M. D. Penrose. The longest edge of the random minimal spanning tree. *Annals of Applied Probability*, 15(2):145–164, 1997. 22

[Pen99] M. D. Penrose. On k-connectivity for a geometric random graph. *Random Structures and Algorithms*, 15(2):145–164, 1999. 22

[PLWL06] L. Ping, L. Liu, K. Y. Wu, and W. K. Leung. Interleave-division multiple-access. *IEEE Transactions on Wireless Communications*, 5(4):938–947, April 2006. 75, 93

[Pro00] J. G. Proakis. *Digital Communications*. McGraw-Hill, 4th edition, 2000. 11

[PS03] C. Peraki and S. D. Servetto. On the maximum stable throughput problem in random networks with directional antennas. In *Proc. ACM Int. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Annapolis, USA, June 2003. 13, 127

[QZC05] X. Qian, B. Zheng, and J. Cui. Increasing throughput of CDMA-based ad hoc network by multiuser detection. In *Proc. Asia-Pacific Conf. on Communications*, Perth, Australia, October 2005. 75, 100

[QZY05] X. Qian, B. Zheng, and G. Yu. Throughput analysis for fully-connected ad hoc network joint with multiuser detection. *IEICE Transactions on Communications*, 88(9):3554–3562, September 2005. 75, 100

[Ram01] R. Ramanathan. On the performance of ad hoc networks with beamforming antennas. In *Proc. ACM Int. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Long Beach, USA, October 2001. 13, 23

[Rap02] T. S. Rappaport. *Wireless Communications: Principles and Practice, Second Edition*. Prentice-Hall, Inc., 2002. 9, 75

[Rho04] D.L. Rhodes. Efficient routing in ad hoc networks with directional antennas. In *Proc. Military Communications Conf. (MILCOM)*, Monterey, USA, October 2004. 13

[RHV97] P. Robertson, P. Hoeher, and E. Villebrun. Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo decoding. *European Transactions on Telecommunications (ETT)*, 8(2):119–125, March 1997. 80

[RM00] V. Rodoplu and T. H. Meng. Position based CDMA with multiuser detection (P-CDMA/MUD) for wireless ad hoc networks. In *Proc. IEEE Int. Symp. on Spread-Spectrum Techniques and Applications (ISSSTA)*, Parsippany, USA, September 2000. 75, 102

[Rob75] L. G. Roberts. ALOHA packet system with and without slots and capture. *ACM SIGCOMM Computer Communication Review*, 5(2):28–42, April 1975. 68

[RSB+03] S. Roy, D. Saha, S. Bandyopadhyay, T. Ueda, and S. Tanaka. A network-aware mac and routing protocol for effective load balancing in ad hoc wireless networks with directional antenna. In *Proc. ACM Int. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Annapolis, USA, June 2003. ACM. 13, 111, 112

[RSO+05] D. Raychaudhuri, I. Seskar, M. Ott, S. Ganu, K. Ramachandran, H. Kremo, R. Siracusa, H. Liu, and M. Singh. Overview of the ORBIT radio grid testbed for evaluation of next-generation wireless network protocols. In *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, New Orleans, USA, March 2005. 6

[RW99] S. Roy and H. Y. Wang. Performance of CDMA slotted ALOHA multiple

access with multiuser detection. In *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, New Orleans, USA, September 1999. 75, 100

[SdV04] J. A. Stine and G. de Veciana. A paradigm for quality-of-service in wireless ad hoc networks using synchronous signaling and node states. *IEEE Journal on Selected Areas in Communications*, 22(7):1301–1321, September 2004. 112, 131

[SE02] C. Sankaran and A. Ephremides. The use of multiuser detectors for multicasting in wireless ad hoc CDMA networks. *IEEE Transactions on Information Theory*, 48:2873–2887, November 2002. 75, 100

[SG02] B. Shrader and T. Giles. Scheduling and performance of multihop radio networks with multi user detection. In *Proc. Swedish Workshop on Wireless Ad-hoc Networks*, Stockholm, Sweden, May 2002. 75, 100

[SGZ01] M. Sanchez, T. Giles, and J. Zander. CSMA/CA with beam forming antennas in multihop packet radio. In *Proc. Swedish Workshop on Wireless Ad-hoc Networks*, Stockholm, Sweden, March 2001. 110

[SH04b] H. Schoeneich and P. A. Hoeher. A hybrid multiple access scheme delivering reliability information. In *Proc. Int. ITG Conf. on Source and Channel Coding (SCC)*, Erlangen, Germany, January 2004. 80, 93

[Sha48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27, July 1948. 127

[SJ04] A. K. Saha and D. B. Johnson. Routing improvement using directional antennas in mobile ad hoc networks. In *Proc. IEEE Global Communications Conf. (GLOBECOM)*, Dallas, USA, November 2004. 13

[SKS07] M. Scheffel, M. Kiese, and T. Stidsen. A clustering approach for scalable network design. In *Proc. Int. Network Optimization Conf. (INOC)*, Spa, Belgium, April 2007. 55

[SR03a] A. Spyropoulos and C. S. Raghavendra. Asympotic capacity bounds for ad-hoc networks revisited: the directional and smart antenna cases. In *Proc. IEEE Global Communications Conf. (GLOBECOM)*, San Francisco, USA, December 2003. 13, 128

[SR03b] A. Spyropoulos and C. S. Raghavendra. Capacity bounds for ad-hoc networks using directional antennas. In *Proc. IEEE Int. Conf. on Communications (ICC)*, Anchorage, USA, May 2003. 13, 128

[SS03] H. Singh and S. Singh. A MAC protocol based on adaptive beamforming for ad hoc networks. In *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Beijing, China, September 2003. 112

[SS04a] H. Singh and S. Singh. Smart-802.11b mac protocol for use with smart antennas. In *Proc. IEEE Int. Conf. on Communications (ICC)*, Paris, France, June 2004. 111, 112

[Ste82] H. Steyskal. Synthesis of antenna patterns with prescribed nulls. *IEEE Communications Magazine*, (2):273–279, 1982. 19, 20

[Ste03] M. E. Steenstrup. Neighbor discovery among mobile nodes equipped with

smart antennas. In *Proc. Scandinavian Workshop on Wireless Ad-hoc Networks*, Stockholm, Sweden, May 2003. 13

[Sti06] J. A. Stine. Exploiting smart antennas in wireless mesh networks using contention access. *IEEE Wireless Communications*, 13(2):38– 49, April 2006. 112, 113

[SVGD02] J. A. Stine, G. De Veciana, K. H. Grace, and R. C. Durst. Orchestrating spatial reuse in wireless ad hoc networks using synchronous collision resolution. *Journal of Interconnection Networks (JOIN)*, 3:167–198, 2002. 112, 113

[TAB06] A. Tyrrell, G. Auer, and C. Bettstetter. Synchronization inspired from nature for wireless meshed networks. In *Proc. Int. Conf. on Wireless Communications, Networking and Mobile Computing (WiCOM*, Wuhan City, China, September 2006. 11, 12

[TG03] S. Toumpis and A.J. Goldsmith. Capacity regions for wireless ad hoc networks. *IEEE Transactions on Wireless Communications*, 2(4):736–748, July 2003. 127

[TG06] S. Toumpis and A. J. Goldsmith. New media access protocols for wireless ad hoc networks based on cross-layer principles. *IEEE Transactions on Wireless Communications*, 5(8):2228–2241, August 2006. 131

[TKU05] H. Boche J. Fonollosa J. Bach Andersen T. Kaiser, A. Bourdoux and W. Utschick. *Smart Antennas – State of the Art*. Hindawi Publishing Corp., 2005. 15

[TMRB02] M. Takai, J. Martin, A. Ren, and R. Bagrodia. Directional virtual carrier sensing for directional antennas in mobile ad hoc networks. In *Proc. ACM Int. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Lausanne, Switzerland, June 2002. 110

[TNV04] L. Tong, V. Naware, and P. Venkitasubramaniam. Signal processing in random access. *IEEE Signal Processing Magazine*, 21(5):29–39, September 2004. 73, 75

[VE05] S. Vural and E. Ekici. Analysis of hop-distance relationship in spatially random sensor networks. In *Proc. ACM Int. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Urbana-Champaign, USA, May 2005. 33

[Ver03] S. Verdu. *Multiuser Detection*. Cambridge University Press, 2003. 76

[WNE02] J. E. Wieselthier, G. D. Nyuyen, and A. Ephremides. Energy-aware wireless networking with directional antennas: The case of session-based broadcasting and multicasting. *IEEE Transactions on Mobile Computing*, 1(3):176–191, July 2002. 13

[WP99] X. Wang and H. V. Poor. Iterative (Turbo) soft interference cancellation and decoding for coded CDMA. *IEEE Transactions on Communications*, 47(7):1046–1061, July 1999. 80, 93

[Yeh04] C. H. Yeh. Spread spectrum techniques for solving MAC-layer interference issues in mobile ad hoc networks. In *Proc. IEEE Vehicular Technology Conf. (VTC)*, Milan, Italy, May 2004. 102

[YPK03] S. Yi, Y. Pei, and S. Kalyanaraman. On the capacity improvement of ad hoc wireless networks using directional antennas. In *Proc. ACM Int. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Annapolis, USA, June 2003. 13, 128

[YPKAS07] S. Yi, Y. Pei, S. Kalyanaraman, and B. Azimi-Sadjadi. How is the capacity of ad hoc networks improved with directional antennas? *Wireless Networks*, 13(5):635–648, 2007. 128

[ZDJ07] X. Zhou, S. Durrani, and H. Jones. Analytical study of connectivity in wireless ad hoc networks with random beamforming. In *Proc. Int. Conf. on Signal Processing and Communication Systems (ICSPCS)*, Gold Coast, Australia, December 2007. 71

[ZJDS07] X. Zhou, H. Jones, S. Durrani, and A. Scott. Effects of beamforming on the connectivity of ad hoc networks. In *Proc. Australian Communications Theory Workshop (AusCTW)*, Adelaide, Australia, February 2007. 71

[ZZA+06] M. Zorzi, J. Zeidler, A. Anderson, B. Rao, J. Proakis, A. L. Swindlehurst, M. Jensen, and S. Krishnamurthy. Cross-layer issues in MAC protocol design for MIMO ad hoc networks. *IEEE Wireless Communications*, 13(4):62–76, 2006. 75