Context-Enabled Optimization of Energy-Autarkic Networks for Carrier-Grade Wireless Backhauling

Kontextunterstützte Optimierung
energieautarker, funkbasierter
Punkt-zu-Punkt-Backhaulnetze

von
Dipl.-Wirtsch.-Ing. Christian Mannweiler
geboren in Idar-Oberstein

Vom Fachbereich Elektrotechnik und Informationstechnik
der Technischen Universität Kaiserslautern
zur Verleihung des akademischen Grades
Doktors der Ingenieurwissenschaften (Dr.-Ing.)
genehmigte Dissertation

D 386

Datum der mündlichen Prüfung:	15. Dezember 2014
Dekan des Fachbereichs:	Prof. Dr.-Ing. Hans D. Schotten
Prüfungsvorsitzender:	Prof. Dr.-Ing. Andreas König
1. Berichterstatter:	Prof. Dr.-Ing. Hans D. Schotten
2. Berichterstatter:	Prof. Dr.-Ing. Ralf Tönjes

Zusammenfassung

In der heutigen Informations- und Wissensgesellschaft mit nahezu allgegenwärtigem Zugang zum Internet findet eine Vielzahl der Aktivitäten des täglichen Lebens vollständig oder teilweise „online“ statt. Dies gilt sowohl für berufliche als auch für private Zwecke. Telekommunikationsunternehmen, aber auch Regierungen, haben dies erkannt und entsprechend hohe Summen in eine zuverlässige und sichere Kommunikationsinfrastruktur investiert. In vielen ländlichen Gebieten der Industrie- und Entwicklungsländer jedoch, welche im besonderen Maße bereits unter den demographischen Entwicklungen leiden, bedroht die mangelnde Verfügbarkeit eines breitbandigen Internetzugangs zusätzlich die nachhaltige Entwicklung und Teilhabe in den Bereichen Wirtschaft, Soziales, Politik und Kultur. Aktuelle Technologien, wie z.B. LTE Advanced (LTE-A), Very High Speed Digital Subscriber Line (VDSL) oder Glasfaser, sind insbesondere für den Einsatz in dicht besiedelten Ballungsräumen konzipiert und entwickelt worden. Zwar verfügen sie somit über die notwendigen Kapazitäten, allerdings sind die mit ihrem Einsatz verbundenen Investitions- und operativen Kosten so hoch, dass sie in ländlichen, oftmals dünn besiedelten Gebieten für Betreiber nicht wirtschaftlich sinnvoll einzusetzen sind. Mithin wird hierdurch das Problem der *Digitalen Kluft* nur noch verschärft.

Angesichts dieser schwerwiegenden Defizite betrachtet diese Arbeit eine neue Klasse von Kommunikationsnetzen: energieautarke koordinierte Punkt-zu-Punkt-Funknetzwerke für das so genannte *Backhaul*. Dieses bezeichnet jenen Teil eines Kommunikationsnetzes, der Zugangspunkte (z.B. Basisstationen) mit dem Kernnetz (engl. *core network, backbone*) verbindet. Diese neue Klasse lässt sich durch drei wesentliche Eigenschaften charakterisieren: kosteneffiziente Radiotransceiver-Hardware, ein größtenteils selbstorganisierendes Netzwerkmanagement sowie eine autarke Energieversorgung aus erneuerbaren Quellen. Mit diesem neuen Ansatz soll es gelingen, kostengünstig Leistungsmerkmale wie hohe Zuverlässigkeit und Verfügbarkeit mit reduziertem Wartungsaufwand sowie unabhängiger Energieversorgung zu verbinden. Die Schwächen klassischer, so genannter vermaschter Funknetze (engl. *Wireless Mesh Network (WMN)*), z.B. Betrieb in nicht-lizenzierten Frequenzbändern, unzuverlässiges Routing, limitierte

Netzwerkmanagement- und Steuerungsfunktionen, niedrige Datenraten sowie hohe und stark schwankende Latenzzeiten, sollen durch eine konzeptionelle und technische Anpassung, Erweiterung und Verbesserung existierender Technologien und Protokolle behoben werden. So ermöglicht beispielsweise der Einsatz von Multiradio-Multikanal-Knoten eine Verwendung orthogonaler Funkkanäle. Dies vermindert Interferenzen und erhöht die Gesamtkapazität des Systems. Des Weiteren tragen intelligente Steuerungsalgorithmen aus dem Bereich selbstorganisierender Netze (engl. *Self-Organizing Networks (SON)*) auf kostengünstige Weise zu einer erheblichen Verbesserung des Netzwerkmanagements bei.

Um diesem Ansatz zum Durchbruch zu verhelfen, werden in dieser Arbeit einige wichtige Konzepte, Algorithmen und Evaluierungsansätze entwickelt und bewertet. In Anlehnung an die Theorie cyber-physischer Systeme (engl. *Cyber-Physical System (CPS)*) wird das Backhaulnetz zunächst als Multi-Domänen-System modelliert. Der resultierende Netzwerkgraph vereinigt hierbei auf neue Weise physische, netzwerk- und kontextbezogene Modellierungsansätze. Die Netzwerk-Domäne integriert hierbei Stand-der-Technik Modelle für Funknetze auf Systemebene, während die physische Domäne insbesondere die Modelle für Nutzermobilität, Energieverbrauch und -erzeugung beinhaltet. Die Kontext-Domäne umfasst ein offenes, skalierbares und erweiterbares Kontextmanagementsystem, dessen Architektur auf einer Produzenten-Konsumenten-Rollenverteilung sowie entsprechenden Schnittstellen und Protokollen basiert. Dabei realisiert das Kontext-Management-System die Akquisition, Verarbeitung, (semantische) Abstrahierung und Verteilung von Kontextdaten. Mit Hilfe des o.g., umfassenden Netzwerkgraphen wird eine - zunächst noch - allgemeine Problemstellung formuliert und entsprechende Lösungsalgorithmen entworfen. Insbesondere die Problematik der Topologiesteuerung (engl. *Topology Control (TC)*) wird hierbei adressiert, da sich wesentliche Eigenschaften eines Graphen (u.a. maximaler Fluss, Kantenzusammenhang, durchschnittlicher Knotengrad) über dessen Topologie steuern lassen. Die im Rahmen dieser Arbeit entwickelten TC-Algorithmen betreffen sowohl die Phase des Netzaufbaus als auch die des Betriebs. Wesentliche Eingabeparameter der Algorithmen sind hierbei u.a. der Standort und die Anzahl der Funkschnittstellen eines Backhaul-Knotens. Es wird gezeigt, dass es sich bei der Topologieoptimierung um ein NP-schweres (nicht in polynomieller Zeit lösbares) Problem aus der Klasse der quadratischen Zuordnungsprobleme (engl. *Quadratic Assignment Problem (QAP)*), im speziellen der Klasse der k-

(kanten)zusammenhängenden Graphen mit minimalen Kosten, handelt und wie optimale Lösungen angenähert werden können. Die Qualität heuristischer Lösungen hat hierbei maßgeblichen Einfluss auf die Routingeffizienz, die Redundanzgrade des Netzwerkgraphen, die Zuverlässigkeit der Dienstbereitstellung sowie den Energieverbrauch.

Für die Topologieoptimierung in der Phase des Netzaufbaus werden im Rahmen dieser Arbeit zwei neue Algorithmen entwickelt und evaluiert. Die so genannte maximale Pfadredundanz-Topologieoptimierung (engl. *maximum path redundancy TC*) hat zum Ziel, den Kantenzusammenhang sowie den Quellen-Senken-Fluss (Kapazität) des Netzwerkgraphen zu maximieren. Die ringbasierte Topologieoptimierung (engl. *cycle-based TC*) hingegen konzentriert sich auf den effizienten Einsatz von Funkschnittstellen, indem "Hop-Countßwischen Knotenpaaren gegen Kantenzusammenhang des Graphen abgewogen wird. Um einen verbundenen Graphen zu gewährleisten, wird zunächst eine Ring-Topologie aufgebaut, d.h. der minimale Kantenzusammenhang beträ gt zwei. In Bezug auf die Topologieoptimierung während des Betriebs wird ein neuer Algorithmus zur Optimierung der Backhaul-Topologie (engl. *Backhaul Topology Optimization (BTO)*) vorgestellt. Dieser umfasst ein parametrisierbares, nicht-lineares Optimierungsproblem, welches, je nach Variante, eine multikriterielle Zielfunktion und eine ausgewählte Menge von Nebenbedingungen enthält. Beispielhafte Zielgrößen sind Netzwerkkapazität, Energieverbrauch oder Fairness der Ressourcenallokation. Durch die periodische Lösung des Optimierungsproblems wird zu jedem Ausführungszeitpunkt die optimale Graphtopologie berechnet, d.h. ein k-(kanten)zusammenhängender Graph mit minimalen Kosten wird erzeugt.

Weiterhin werden die entwickelten Konzepte und Algorithmen einer analytischen und quantitativen (simulationsbasierten) Evaluierung unterzogen und in diesem Rahmen mit Referenz-Algorithmen aus den Bereichen Topologieoptimierung sowie energieautarker Systembetrieb verglichen. Dies erfordert zunächst den Entwurf einer Evaluierungsmethodik, die auch ein so genanntes *Balanced Scorecard*-Konzept zur Performanzbewertung mittels verschiedener Schlüsselindikatoren (engl. *Key Performance Indicator (KPI)*) umfasst und im Rahmen eines Simulators auf Systemebene umgesetzt wird. Der Simulator zeichnet sich insbesondere durch die realitätsnahe Modellierung der Multi-Domänen-Abhängigkeiten aus und integriert dabei erstmals Modelle zum Energie-Verbrauch und zur -Erzeugung in funkbasierten Punkt-zu-Punkt-Netzen sowie die darauf aufbauenden kontextaktivier-

ten Netzwerkmanagement-Algorithmen. Die Evaluierung wird anhand zweier typischer Einsatzstandorte für koordinierte Punkt-zu-Punkt-Funknetze vorgenommen. Dies ist zum Einen die Anbindung abgelegener Gemeinden in der Alpenregion, zum Anderen die Versorgung ländlicher Gebiete in Tansania. Ersterer zeichnet sich durch die schwierigen topographischen Voraussetzungen, z.B. in Alpentälern, aus, letzterer eignet sich auf Grund der äquatornahen Lage besonders für einen Betrieb mittels autarker Energieversorgung.

Im analytischen Teil der Evaluierung werden zunächst Ober- und Untergrenzen für den Energieverbrauch eines Backhaul-Knotens bestimmt und davon minimale und maximale Betriebszeit bei reinem Akkubetrieb abgeleitet. Unter der Annahme einer Energieversorgung mittels Photovoltaikmodul werden anschließend szenariospezifische Mindestgrößen für Pufferakku sowie Größe der Modulfläche berechnet, die u.a. die ortstypischen Wetter- und Globalstrahlungsbedingungen sowie die jahreszeitlichen Schwankungen des Sonnenstandes berücksichtigen. Weiterhin wird durch die Berechnung des Approximationslevels γ die Güte der konzipierten TC-Algorithmen nachgewiesen. Sowohl die maximale Pfadredundanz- als auch die ringbasierte Topologieoptimierung erreichen mindestens $\frac{2}{d_{\max}}$ des Zielfunktionswertes der optimalen Lösung ($d_{\max}$ beschreibt den maximalen Grad eines Backhaulknotens, d.h. die Anzahl der verfügbaren Funkschnittstellen).

Im quantitativen, simulationsbasierten Teil der Auswertung wird gezeigt, dass der BTO-Algorithmus zur Optimierung der Netztopologie während des Betriebs die Energieeffizienz steigert und mittels geeigneter Rekonfiguration der Topologie den Energieverbrauch gleichmäßiger über die zur Verfügung stehenden Energiereserven der Knoten verteilen kann. Dies hat zur Folge, dass verglichen mit dem Referenzfall durchschnittlich knapp 20% mehr Backhaul-Knoten verfügbar sind. Bei reinem Akkubetrieb wird die Zeit bis zur Komplettentladung der Pufferakkus im Durchschnitt um fast 22% verlängert. Somit können wesentlich längere Zeitspannen ohne Sonneneinstrahlung überbrückt werden. Weiterhin kann gezeigt werden, dass der Einsatz der BTO-Algorithmen einen positiven Effekt auf die Dienstgüte (engl. *Quality of Service (QoS)*) hat. Der durchschnittliche Durchsatz je Nutzer kann leicht (ca. 3%) erhöht werden und die Fairness der Nutzer-Ressourcenallokation verbessert sich um 10% bis 18%, abhängig vom betrachteten Szenario sowie von der verwendeten Metrik. Schließlich werden mittels einer Sensitivitätsanalyse kontextspezifische Dimensionierungsempfehlungen für die Planung eines Backhaulnetzes erarbeitet. Konkret

werden sinnvolle Intervalle für Kenngrößen wie nutzerbezogene Backhaul-Knotendichte, Anzahl der Gateways je Nutzer und Anteil der Gateway-Knoten an allen Backhaul-Knoten bestimmt.

Zusammenfassend wurde im Rahmen dieser Arbeit die Eignung von koordinierten funkbasierten Punkt-zu-Punkt-Backhaul-Netzen für entwicklungspolitisch relevante Szenarien untermauert. Der Einsatz der entwickelten Konzepte und Algorithmen trägt wesentlich dazu bei, mit Hilfe kostengünstiger Funktechnologien eine den Betreiber-Anforderungen entsprechende (engl. *carrier-grade*) Breitbandversorgung für dünn besiedelte ländliche Regionen umzusetzen. Für Gebiete, die über eine ausreichend hohe solare Einstrahlungsdauer oder über andere regenerative und dezentrale Energiequellen verfügen (z.B. sub-saharisches Afrika, Südeuropa, Südstaaten der USA), stellt die vorliegende Arbeit zusätzlich Verfahren für ein zuverlässiges Netzwerkmanagement bei autarker Energieversorgung bereit. Wichtige Beiträge dieser Arbeit werden in einem Testnetz, welches im Rahmen des Forschungsprojektes *SolarMesh – Energieeffiziente, autonome großflächige Sprach- und Datenfunknetze* aufgebaut wurde, unter realen Bedingungen und in einem operativen Umfeld getestet.

Schlagwörter – zuverlässige, hochverfügbare Punkt-zu-Punkt-Funknetze; autonomes Netzwerkmanagement; Backhaulnetze in ländlichen Regionen; Multi-Domänen-Modellierung und Evaluierungsmethodik; Graphentheorie; kontextsensitive Topologieoptimierung; multikriterielle Optimierung; Energieversorgung aus erneuerbaren Quellen; autarke Energieversorgung

Abstract

In today's society of virtually ubiquitous access to the Internet, a tremendous amount of daily activities, both private and business, have shifted to the online world. Acknowledging the paramount role the Internet plays in our daily life, governments and telecommunications companies have made huge investments to deploy reliable and secure broadband metropolitan (MAN) and wide area networks (WAN). However, in rural areas around the world, particularly in developing countries, but also in remote regions of industrialized nations, the lack of broadband networks perils the social, economic, and political inclusion of people and organizations in these regions. While existing networking technologies for broadband access and WAN connectivity, among them LTE-A, VDSL, optical fiber, or microwave radio relaying solutions, meet the major technical requirements, in particular the provisioning of high data rates, they largely fail at proving their economic viability when deployed in sparsely populated areas. Essentially, current technologies are designed for operation in densely populated metropolitan areas, making them inapt for deployment in rural regions. As an alarming consequence, we can observe a growing digital divide.

In the light of these serious deficiencies, this work establishes the novel category of coordinated Wireless Backhaul Networks (WBNs) for energy-autarkic point-to-point radio backhauling. This networking concept is based on three major building blocks: cost-efficient radio transceiver hardware, a self-organizing network operations framework, and power supply from renewable energy sources. The aim of this novel backhauling approach is to combine carrier-grade network performance with reduced maintenance effort as well as independent and self-sufficient power supply. In order to facilitate the success prospects of this concept, the thesis comprises a set of important contributions.

Formal, multi-domain system model and evaluation methodology
Adapted from the theory of cyber-physical systems, the author devises a multi-domain evaluation methodology and a system-level simulation framework for energy-autarkic coordinated WBNs, including a novel *balanced scorecard* concept. Coordinated WBNs are modeled using specific graph types that extend existing approaches by incorporating network, physical, and contextual domain aspects. While the network domain covers conventional system-level network models, the physical domain features a novel,

comprehensive energy supply and consumption model. The context domain incorporates algorithms to collect, analyze, and exploit context information for autonomous, energy-efficient, and self-organized network operation. The resulting integrated and extensible graph model allows for the definition formal problem statements and the application and evaluation of a broad set of network optimization algorithms.

Context-enabled network management algorithms

The thesis specifically addresses the topic of Topology Control (TC) in point-to-point radio networks and how it can be exploited for network management purposes. Given a set of network nodes equipped with multiple radio transceivers and known locations, TC continuously optimizes the setup and configuration of radio links between network nodes, thus supporting initial network deployment, network operation, as well as topology re-configuration. In particular, the author shows that TC in WBNs belongs to the class of NP-hard QAPs and that it has significant impact in operational practice, e.g., on routing efficiency, network redundancy levels, service reliability, and energy consumption. Two novel algorithms focusing on maximizing edge connectivity of network graphs are developed with respect to topology design decisions in the network deployment phase: Maximum Path Redundancy (MPR) and cycle-based TC. Regarding TC during network operation, the author designs different variants of a Backhaul Topology Optimization (BTO) algorithm considering important constraints and an objective function trading off network capacity against energy consumption and fairness of resource allocation. Each of these variants results in specific minimum cost k-edge-connected spanning subgraph solutions.

Performance evaluation of algorithms

Finally, this work carries out an analytical benchmarking and a numerical performance analysis of the introduced concepts and algorithms. The author analytically derives minimum performance levels of the the developed TC algorithms. This is achieved by computing analytical approximation levels that indicate to what extent algorithms approach optimum results. In case of MPR and cycle-based TC, the analysis shows that their approximation level always exceeds a minimum that can be expressed as a function of the number of radio interfaces available at a backhaul node, which is assumed to lie between two and four. Moreover, the derivation of upper and lower bounds for power consumption of a node determines required battery capacities and photovoltaic panel sizes for energy-autarkic

operation. For the analyzed scenarios of remote Alpine communities and rural Tanzania, the evaluation shows that the algorithms improve energy efficiency and more evenly balance energy consumption across backhaul nodes, thus increasing the number of available backhaul nodes by approximately 20% compared to state-of-the-art TC algorithms. On the temporal scale, the lifetime of backhaul nodes is extended by 21%, thus accommodating longer periods of lower solar irradiation levels. Further, the author derives scenario-specific minimum solar irradiation levels as well as corresponding solar panel and battery sizes that are required for a continuous operation of backhaul nodes solely relying on solar energy supply.

In summary, this thesis provides a set of methodological and algorithmic contributions for modeling, optimizing, and evaluating the performance of coordinated WBNs. Besides supporting a successful implementation and deployment of this novel concept for backhaul networking, these contributions help to counteract the growing digital divide people in sparsely populated and remote area experience.

Keywords – carrier-grade point-to-point radio networks; coordinated backhaul networks in rural areas; autonomous networking, multi-domain modeling and evaluation methodology; graph theory; context-aware topology control; multi-criteria optimization; power supply from renewable sources; self-sufficient energy supply

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

The global telecommunications market currently finds itself in a phase of groundbreaking changes. Besides the geographic expansion into virtually every country, the growth of data traffic and the amount of Internet-capable devices, particularly mobile devices, around the globe has reached unprecedented levels. In fact, in many regions, networks are not able to keep pace with these tremendous device populations. More precisely, the growth in Internet traffic is attributed to two factors. On the one hand, by 2017, there will be nearly 1.4 mobile devices per capita, which is equivalent to approximately 10 billion devices, whereas there was approximately one device per capita (i.e., seven billion) in 2013 [Cis13]. On the other hand, the traffic incurred per device will dramatically increase due to more bandwidth-intensive applications, such as video streaming, as well as a shift of business and leisure activities to networked devices. Therefore, mid-term forecasts expect overall global Internet traffic to reach a volume of 1.3 zettabytes ($1.3 \cdot 10^{27}$ bytes) by the end of 2016 [Cis12]. This equals a more than 3.5-fold increase when compared to 2011. Peak traffic volume, which is relevant for network dimensioning, will even increase fivefold within the same time span. In parallel, the current shift from wired to wireless devices will continue. While traffic originating from wired devices accounted for 55% of total data volume in 2011, the share of traffic from wireless devices will prevail by 2016, reaching 61% of total volume [Cis13].

In long-term forecasts, leading network manufacturers estimate that today's dominating scenarios of human-centric communication will, by 2020, be complemented by a tremendous increase in the numbers of communicating machines, resulting in 50 billion wireless devices (i.e., more than 6.5

devices per capita) connected to the Internet [Eri11]. For highly-developed ICT societies, such as Western Europe, predictions estimate that traffic to be transported over cellular networks will increase by a factor of approximately 70 until 2020 [Eva11]. Other forecasts by telecommunications companies expect an even more massive traffic growth, projecting global wireless traffic volumes to increase by a factor of 1000 until 2020 when compared to 2010 figures, primarily driven by increased usage of mobile multimedia services. The impressive dimensions of these figures give a rough idea about the challenges future network development is exposed to and the comprehensiveness future technologies have to exhibit in terms of scalability, efficiency, and versatility [Bra13].

From the network perspective, the availability of broadband capacity in access and wide area networks has meanwhile reached satisfying levels in metropolitan and urban areas, thus accommodating the tremendous surge in Internet traffic. However, in suburban and rural areas, both private users and companies, in particular Small or Medium-sized Enterprises (SMEs), suffer from limited bandwidth in wide area networks (Wide Area Networks (WANs)). This does not only compromise the success of established businesses, but also deters other firms from setting up operations in these regions. At the same time, from the network operators' perspective, it is not very attractive to invest in modern Information and Communication Technology (ICT) backhaul infrastructure since important key performance indicators (KPIs, such as, Average Revenue Per User (ARPU), number of users, and overall contribution margin) and thus profit margins are comparatively low in rural and sparsely populated areas. Consequently, this dilemma, besides manifesting the digital divide between rural and urban areas, also contributes to the gap in other socio-economic aspects. More specifically, long-term implications include loss of attractive and sustainable jobs, degradation of purchasing power, migration of young and well-educated people to urban areas (*brain drain*), drop of tax revenues of communities, decline of public and private infrastructure and buildings as well as a shortfall in cultural activities. While insufficient Internet connectivity does not serve as the sole explanation for such trends, it certainly plays a major role.

In parallel, developing countries face similar challenges with respect to surges in traffic volume, device population, and resulting requirements for broadband WANs. The Middle East and Africa will have the strongest mobile data traffic growth of any region at 77% CAGR until 2017 [Cis13].

Apart from obvious research and business opportunities (developing communication networks for rural Africa, tapping a large unsatisfied demand), there is the aspect of economic cooperation and development. Developing and deploying ICT infrastructure in rural regions of developing countries will improve the life of billions of people who do not have satisfactory access to the Internet or other telecommunication services [SSJ11]. Due to the lack of viable alternatives, communication networks are the only realistic means for large scale access to information. Hence, investments into these infrastructures will have direct consequences on economic growth, the quality of medical services, education, inclusion into everyday life, and diversity of opinion [MLKS12].

From today's perspective, the availability of broadband infrastructure in rural areas and emerging countries for providing Internet services is far from satisfactory. This is attributed to a set of causes. First, probably being the most important structural problem, the unfavorable comparison of cost of state-of-the-art technology (e.g., cellular networks or optical fiber) on the one hand and revenue prospects in terms of ARPU on the other prevents operators from committing considerable investments in non-urban areas. This phenomenon cannot only be observed in Germany but throughout entire Europe. Beyond these initial investments, on-site network maintenance is more expensive in rural areas since equipment density is low, i.e., a single field engineer cannot cover as many nodes as his counterpart in urban environments and also has to travel longer distances. In developing countries, the general lack of qualified maintenance personnel is an additional problem. To mitigate these issues, built-in self-organization functionalities for extended autonomous operation and reduced system complexity are fundamental requirements that latest technology and hardware do not cover at reasonable costs. But then, more cost-efficient technologies, such as wide area mesh networking based on technologies operating in Industrial, Scientific, and Medical (ISM) bands, do currently not support carrier-grade performance of a backhaul network, i.e., no guarantees can be given for important Quality of Service (QoS) parameters, e.g., data rate, latency, jitter, Packet Error Rate (PER), Bit Error Rate (BER), etc.

Besides these conventional network performance considerations, backhauling networks have to cope with another challenge. Due to remote deployment sites, power supply becomes a major concern when planning network topology. In industrialized countries, backhaul networks often have to cross remote or topographically challenging regions where no power grid

is available. Emerging countries, though, face even more fundamental restrictions. Here, availability and reliability of power supply is very fluctuating, even in urban areas. At times, power outages can last as long as several hours, making back-up solutions absolutely mandatory. Both problems can be addressed by alternative energy supply, e.g., from sun or wind. Since these sources are not available continuously, either, on-site battery capacities have to be dimensioned to overcome phases of darkness or lull. Moreover, network performance and energy consumption need to be adapted to availability of energy reserves.

In summary, three major technological deficiencies of today's wireless backhaul networking motivate the work in this thesis: cost-efficient carrier-grade radio technology, self-organized network operation, and autarkic, i.e., self-sufficient energy supply.

1.2 Objectives

In light of the aforementioned challenges, a cost- and energy-efficient, yet carrier-grade backhaul infrastructure serves as an important component of a viable system solution for affordable broadband communications in rural and sub-urban areas in both industrialized and emerging countries. However, there is no established system technology with sufficient operational deployment history. Hence, there is a significant need for analytical research efforts covering the evaluation of deployment options and operation schemes for coordinated wireless backhaul networks. This work addresses a selected range of relevant topics, as depicted in the following.

In broad terms, this thesis aims at provisioning the modeling and evaluation framework that supports the successful introduction and deployment of cost-efficient, carrier-grade wireless backhaul networks. Starting from the identified drawbacks of current backhaul networking solutions, the first objective of this thesis is to propose and characterize a new network category labeled *Coordinated Wireless Backhaul Networks* (WBN) that can reduce the existing reluctance of operators to deploy broadband networks in rural areas. Second, the author aims at developing a comprehensive modeling and evaluation framework for this network category. Third, different context-enabled network optimization algorithms specifically addressing the constraints of fluctuating yet self-sufficient energy supply and carrier-grade network performance shall be developed. Fourth, these algorithms are analyzed both analytically as well as quantitatively using the designed

modeling and evaluation framework. In practical terms, the following list depicts some of the questions the thesis shall answer:

1. Principal deployment and operation issues
    - What are the defining characteristics of coordinated wireless backhaul networks?
    - What are the limitations of current implementations and deployments?
    - Which radio technology can fulfill the strict cost limits?
    - Do energy-autarkic wireless backhaul networks (WBNs) have the potential to serve as a viable solution for broadband connectivity in rural areas?
2. Autonomous, coordinated networking
    - How does context information enhance optimization algorithms to achieve a carrier-grade network performance in light of the given energy restrictions?
    - What are concrete Topology Control (TC) optimization algorithms that attenuate the impact of these restrictions?
    - How do these algorithms influence QoS and network performance metrics?
    - How can the trade-off between network performance and energy efficiency be resolved in a satisfactory manner?
3. Autarkic energy supply
    - What are the specific constraints that are introduced by power supply from renewable sources, such as sun or wind?
    - What are the factors to properly dimension the size of Photo-Voltaic (PV) modules and the capacity of buffer batteries in solar-powered backhaul networks?
    - What are the minimum requirements in terms of average solar irradiation?
    - Which are the geographical locations (geographical latitudes) and weather conditions meeting these minimum requirements?

1.3 Approach and Outline

Given the problem description and the corresponding thesis objectives, as depicted in previous sections, the structure and methodological approach developed for this work is illustrated in Figure 1.1. Chapter 2 establishes the category of coordinated WBNs by elucidating the defining characteristics, the development history, and typical deployment scenarios of this network category. The chapter motivates the two main deployment scenarios considered in this thesis: rural communities in Tanzania and remote villages in the Alps. The following core of this thesis consists of four parts. First, the author develops a preliminary generic, multi-objective problem statement as well as a novel modeling and evaluation approach in Chapter 3. This facilitates the development of several specific formal problem statements which include, among others, topology control and energy-autarkic operation. The initial, general structure of the analyzed optimization problems, which belong to the class of Quadratic Assignment Problem (QAP), is shown in Problem 1.1, see below.

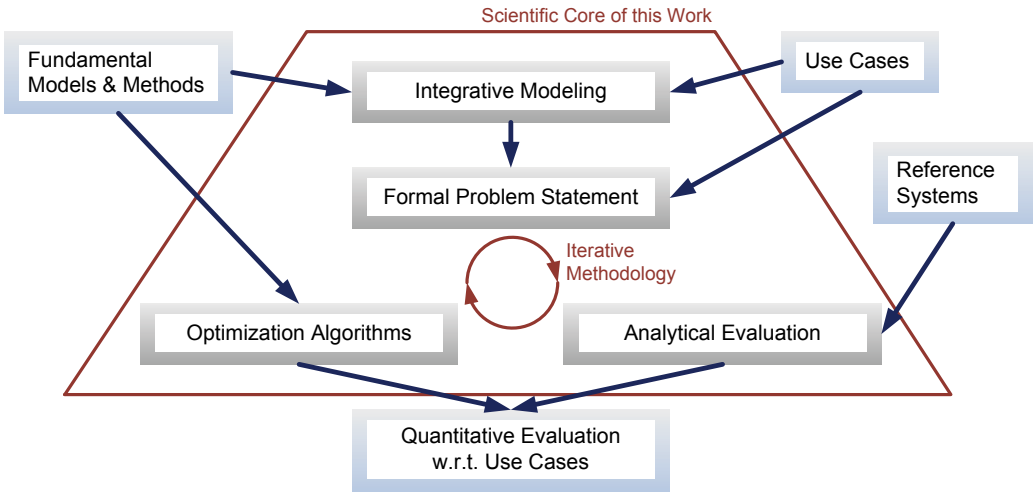


Figure 1.1: Methodological approach of this thesis

Chapter 4 depicts state-of-the-art approaches for modeling point-to-point radio networking and motivates the differentiation into the physical, network, and context domain. The author analyzes the suitability of the models in the context of this work and, if necessary, motivate and implement modifications and extensions. In parallel, the models allow for continuously refining the problem statement. In a subsequent step, they are integrated to

yield a comprehensive multi-domain model of a coordinated WBN. Chapter 5 presents an overview of symbolic and numerical optimization techniques for solving multi-objective linear and non-linear problems. A set of criteria is used to evaluate their efficiency with respect to finding or approximating optimal solutions for QAPs. In Chapter 6, the preceding analysis serves as a basis for developing various novel optimization algorithms that address the specific problem of context-aware TC for WBN deployment and operation. By means of a symbolic analysis, Chapter 7 derives performance boundaries of numerical algorithms, which in turn serve as a guidance for further improving both problem statement as well as the employed optimization methods. Eventually, Chapter 8 returns to the specific deployment scenarios, applies the developed algorithms in a particular context, and thus quantitatively evaluate their effectiveness with respect to a context-enabled optimization of coordinated WBNs. Further, numerical results are used to verify analytical bounds and evaluate their quality. A summary of the thesis, highlighting important novelties and contributions, as well as an outlook on follow-up research work concludes this thesis in Chapter 9. Finally, for the sake of completeness, Appendices A through D depict state-of-the-art concepts from graph theory, routing protocols, context management functions, and optimization algorithms that are relevant in the scope of this thesis.

Problem 1.1: Initial, generic problem statement

$$\text{maximize} \quad Q(\mathbf{x}) + E(\mathbf{x}) + F(\mathbf{x}) \tag{1.1a}$$

$$\text{s.t.} \quad \mathbf{h}(\mathbf{x}) = \mathbf{0}, \qquad \mathbf{h} \in \mathbb{R}^{p_h} \tag{1.1b}$$

$$\mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \qquad \mathbf{g} \in \mathbb{R}^{p_g} \tag{1.1c}$$

In the scalar objective function (Equation (1.1a)), Q, E, and $\mathbf{x} \in \mathbb{R}^m$ generically denote the collection of QoS-related KPIs, energy efficiency-related KPIs, and the vector of input variables, respectively. F serves as a metric for the fairness of resource allocation among users. The optimization is subject to p_h equality constraints $\mathbf{h}(\mathbf{x})$ (Equation (1.1b)) and p_g inequality constraints $\mathbf{g}(\mathbf{x})$ (Equation (1.1c)).

1.4 Contributions

Integrative multi-domain modeling of energy-autarkic coordinated WBNs for broadband communications, context-aware performance optimization, as well as analytical and quantitative evaluation of the developed algorithms form the centerpiece of this work. Corresponding contributions fall into two distinct areas: contributions in the conceptual and methodological area on the one hand and contributions in algorithm design and evaluation on the other. With respect to the former, the thesis develops and refines a general problem statement and designs an evaluation methodology for modeling WBNs and for evaluating proposed optimization algorithms. This methodology is implemented in a novel multi-domain simulation framework. Furthermore, the analytical and quantitative evaluations illustrate both the flexibility of the evaluation framework as well as the superiority of the developed algorithms when compared to state-of-the-art approaches. More specifically, this thesis has realized the following contributions.

1. **Formal, multi-domain system model and evaluation methodology**

   This work

   - defines and specifies the scope of physical, network, and context domains in WBNs to efficiently model real-world networking infrastructures [MLK+12],
   - develops a formal, multi-domain system-level modeling approach for WBNs, including the autarkic energy supply [MLKS12],
   - develops an evaluation methodology and an corresponding simulation framework, and
   - constructs and refines a formal statement of several WBN-related optimization problems, among them minimum cost k-edge-connected spanning subgraph problem and Quadratic Assignment Problem (QAP) for optimal graph topology [MCS13d, MCS13c].

2. **Architectural concepts for context management**

   This work

   - identifies requirements and functionalities for open, scalable, and extensible context management [MRS+10, SKMS11],
   - develops centralized and distributed producer-broker-consumer architecture concepts [MAS+10, SMKa12, MSK+13],

- develops data (representation) models for context abstraction techniques [MKSS09a, MKSS09b, MSMS10, SWM+11], and
- integrates context-based security and privacy functions into the context management framework [MKSS11b].

3. **Context-enabled network management algorithms**

   This work

   - designs a concept for context-enhanced Radio Resource Management (RRM) in heterogeneous network environments [MKSS10, KMSS10a, CMK+10],
   - implements configurable user classification schemes for efficient multicasting [MKSS11a, MSKS11],
   - develops network deployment algorithms addressing different variations of the minimum cost k-edge-connected spanning subgraph problem [CMS13],
   - develops and successfully applies multi-objective Backhaul Topology Optimization (BTO) algorithms controlling and improving network capacity, network connectedness, resource allocation, lifetime management of nodes and the entire network, and energy balance control during network operation [MCS13d, MCS13c].

4. **Performance evaluation of algorithms**

   This work

   - integrates multi-dimensional performance metrics into a balanced scorecard for WBN performance analysis [MCS13a],
   - analytically verifies performance gains realized by the developed optimization algorithms, e.g., energy efficiency improvements in BTO-operated WBNs [MCS13a],
   - derives location-specific minimum requirements for power supply equipment in case of autarkic energy generation, in particular battery capacity and PV panel size [MS14a],
   - quantitatively evaluates the improvements of quantities, such as node failure rates, average user throughput, average node degree, network graph connectedness, fairness of resource allocation [MCS13c, MCS13b],
   - performs a sensitivity analysis of important quantities with respect to variations of input parameters [MCKS13],

- confirms the sufficient reliability and satisfactory QoS levels of energy-autarkic WBNs for selected deployment scenarios with unreliable (non-existing) energy supply from power grid [MS14b, MCS14].

Before outlining the formal problem statement and the evaluation methodology (Chapter 3), the following chapter will give a compact overview on the evolution, characteristics, advantages, and current limitations of WBNs.

Chapter 2

Wireless Backhaul Networks for Mobile Broadband

The success of wireless communication systems originates from the compliance to standards assuring that the same hardware (particularly, user terminals and base stations) can be used in a truly ubiquitous manner, independent of time and location. Technical standards, such as those prepared by the 3rd Generation Partnership Project (3GPP), an association of six standard development organizations, cover the technical specification of entire cellular systems ranging from physical layer aspects to medium access schemes or system management policies. Over time and triggered by ever-increasing requirements, the technical development of public wireless network systems has undergone tremendous changes and given rise to innovative technologies on different layers. Accordingly, this has resulted in a vast heterogeneity with respect to wireless network classes suited for different deployment and operating conditions, the most prominent class certainly being cellular communication networks. However, other wireless networks, e.g., Wireless Local Area Networks (WLANs) standardized by IEEE 802.11, also show impressive track records.

Within the scope of this thesis, the author defines and deals with a new category of wireless networks that is denoted *coordinated Wireless Backhaul Network (WBN)* for broadband wide-area networking. Since coordinated WBNs constitute a comparatively new class of networks, to date, research in this domain has been limited and many problems (including those described in Section 1.1) remain to be described, analyzed, and understood to a sufficient extent, for the sake of both a complete theoretic system model as well as application in the field. Additionally, WBN frequently have to cope with limited power resources, similar to battery-driven nodes in a

Mobile Ad-hoc NETwork (MANET). Hence, autarkic energy supply, e.g., from PV modules, becomes a common work-around. However, it adds more complexity to system design and operation. Figure 2.1 depicts some deployment options for WBNs, which are further detailed in Section 2.2. The following chapter will lay the foundations with respect to the specifics of WBNs (elucidating why they need to be considered as a separate class of networks), selected challenges occurring during their deployment and operation, as well as appropriate solutions proposed in this thesis.

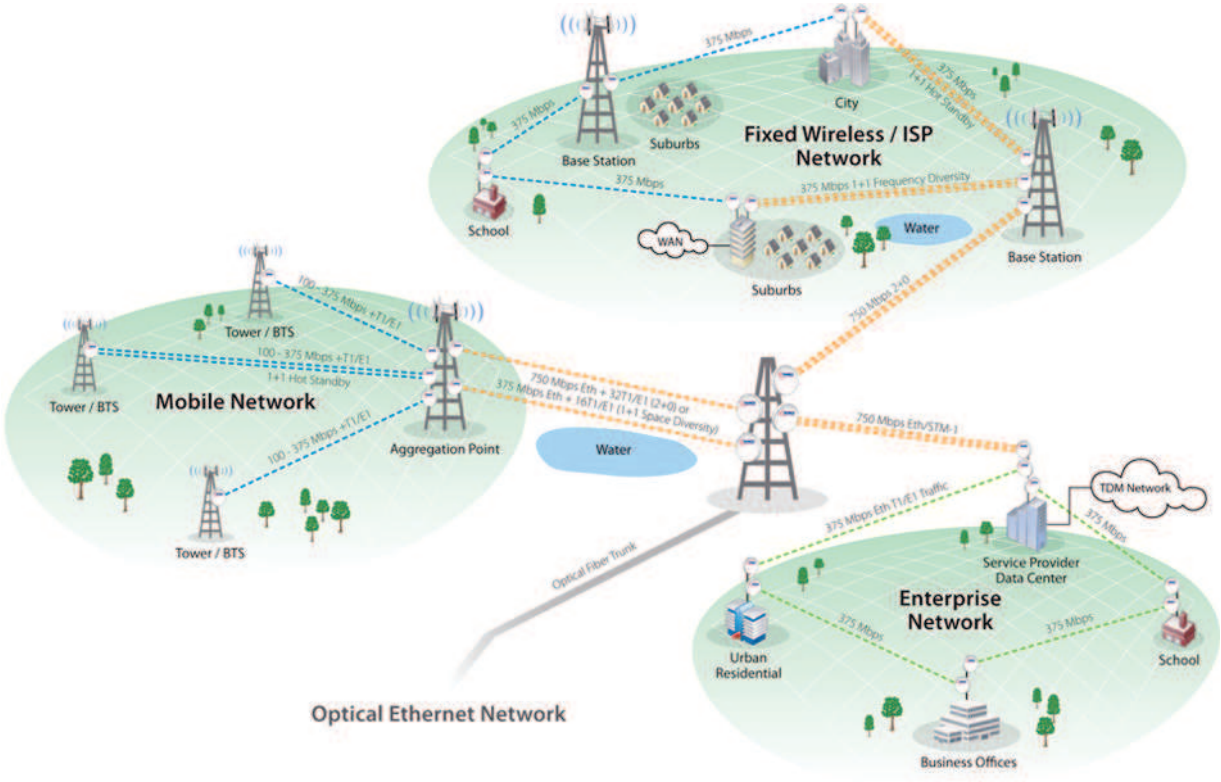


Figure 2.1: Deployment options for wireless backhaul networks [Tra13]

2.1 Evolution of Coordinated Wireless Backhaul Networks

Coordinated WBNs are the latest development in a line of network concepts and technologies exhibiting a completely flat hierarchy. Figure 2.2 depicts the gradual evolution from simple multi-hop networks to managed WBNs for mobile broadband. The class of Wireless Sensor Networks (WSNs) is usually considered to form the origin of this evolution. While they suffer from several limitations, among them low data rates, potentially high la-

tencies, unreliable connectivity, and permanent node failures due to limited energy resources (usually low-capacity batteries), their main benefits consist of low investment and operational costs. Typical deployments consist of several (up to thousands of) meshed, single radio nodes (frequently based on IEEE 802.15.4 technology). A moderate evolutionary step comprised the advent of mobile ad-hoc networks (MANETs) that are composed of more powerful (mobile) nodes and devices (often single radio, IEEE 802.11), e.g., with increased processing power, energy and storage resources, as well as higher transmission bandwidth. The more general category of wireless mesh networks (WMN) consists of (frequently stationary) nodes with a single radio transceiver, while subsequent developments are often comprised of nodes with multiple radio transceivers, which are still comparatively simple to organize and manage. However, due to using the same channel (i.e., radio frequency) across the entire network, they are prone to interference problems, e.g., the *hidden node* issue [LBC+01]. Therefore, multi-radio multi-channel WMNs were the next logical evolution. They also form the basis for the class of coordinated WBNs as considered in this thesis. In contrast to earlier mesh networks, such multi-radio multi-channel WMNs have an increased capacity, since they employ network management methods, e.g., orthogonal channel assignments.

Indeed, these advantages come at the cost of increased management complexity with respect to, among others, node discovery, channel assignment, channel separation, or link calibration. Not being designed for such network types, standard network protocols (particularly those of Institute of Electrical and Electronics Engineers (IEEE) radio technologies) either are subject to significant adaptations or require re-design from scratch. Furthermore, capacity management procedures, e.g., monitoring and allocation of radio resources, as well as mechanisms ensuring minimum Quality of Service (QoS) need to be incorporated into the system architecture [BBC+08].

Exhibiting these favorable characteristics, while at the same time preserving a low to moderate cost level due to their close relationship to WMNs, energy-autarkic, coordinated WBNs increasingly succeed as a wireless deployment solution for rural and emerging regions. In combination with suited access nodes, they provide connectivity for user terminals and, at the same time, serve as a high-capacity wireless point-to-point network [JS03] for relaying traffic from the backbone (core) network to the access nodes or vice versa, cf. Figure 2.4.

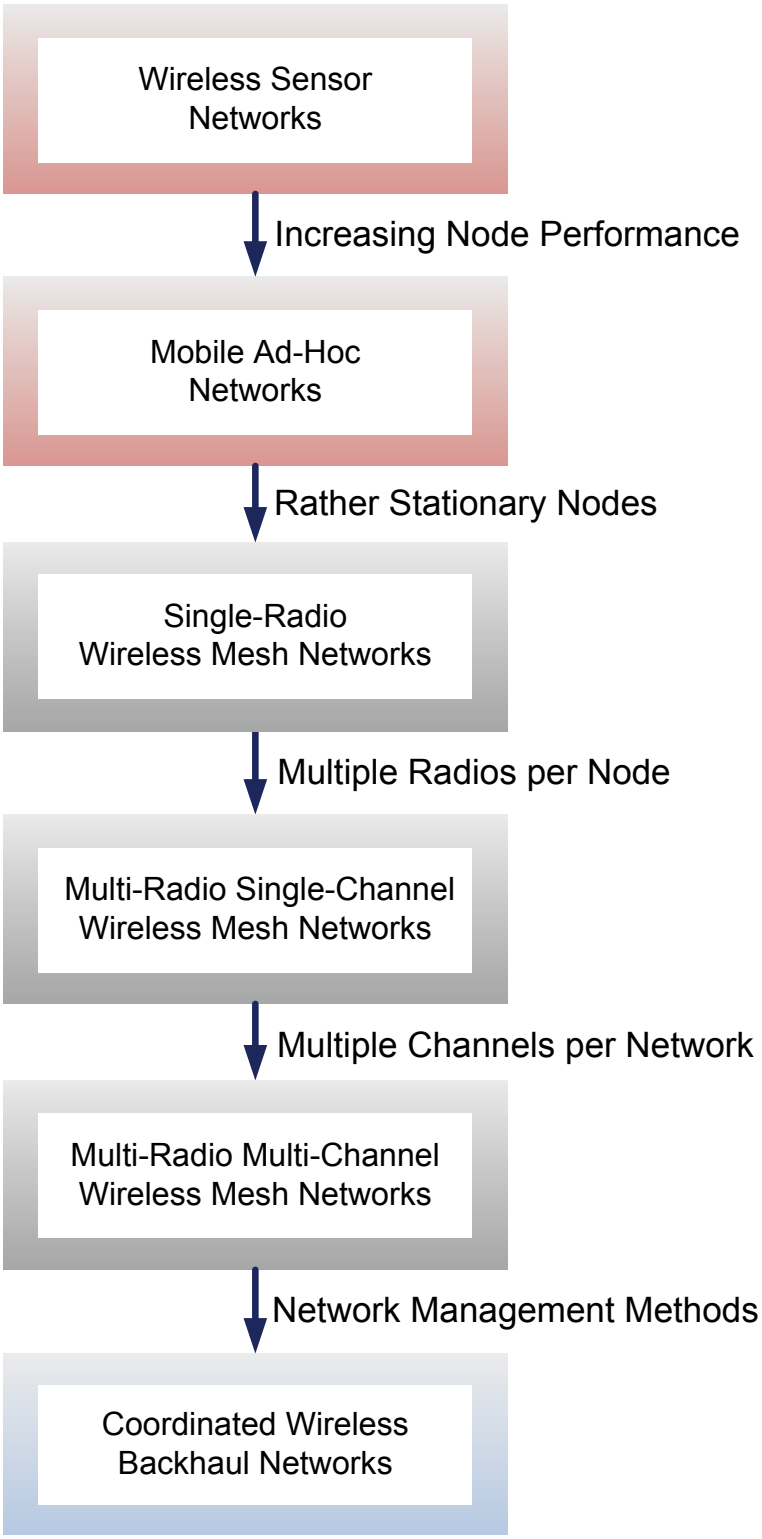


Figure 2.2: Evolution path to WBNs

2.2 Wireless Backhaul Networks: Notion, Classification, and Characteristics

WBNs form a new class of communication network infrastructure in many ways. Based on a set of selected dimensions, the category of WBNs can be defined, classified, and distinguished from other types of networks. This classification serves multiple purposes: Each category of network is suited for a defined scope of tasks. Accordingly, besides significantly differing requirements, individual networks require tailored technical solutions for both hardware and software. Moreover, financial investments, operating costs, and life cycles can vary tremendously.

Figure 2.3 depicts the relationship of WBNs with other network categories on the dimensions *hierarchy level* and *transmission medium*. In the network hierarchy dimension, backbone refers to a set of high-capacity channels (usually wired or optical fiber) connecting local, metropolitan, regional, or national networks and core routers for long-distance interconnections. At connection points, high-capacity switching centers perform data forwarding. In contrast, access nodes are the part of the network where user terminals and other end devices directly connect to, e.g., base stations in cellular communication systems, WLAN access points in WLAN networks, or the local loop (subscriber line) for Digital Subscriber Line (DSL) and Public Switched Telephone Network (PSTN), respectively. The term *backhaul network* refers those parts of the overall ICT infrastructure connecting access nodes with backbone networks. As such, the backhaul network, also referred to as the *middle mile* of a telecommunications network, is part of both Radio Access Networks (RANs) as well as Metropolitan Area Networks (MANs) and WANs, and provides relatively fast, high-capacity connections from an operator's core network to edge nodes. Covered distances of backhaul networks can range from a few kilometers to a few hundred kilometers. With regard to the employed transmission medium, fiber optic lines usually dominate; however, bonded copper lines or Line Of Sight (LOS) microwave relay links are also common, e.g., in mobile communication systems for connecting a remote base station [Sal13]. Recently, more complex wireless backhaul solutions based on cost-efficient mesh technologies have gained attraction. Figure 2.4 illustrates the discussed separation into access nodes, backhaul, and backbone (core) network.

The second dimension (transmission mode) distinguishes between *wired* and *wireless* networks, where the wireless category is further split into

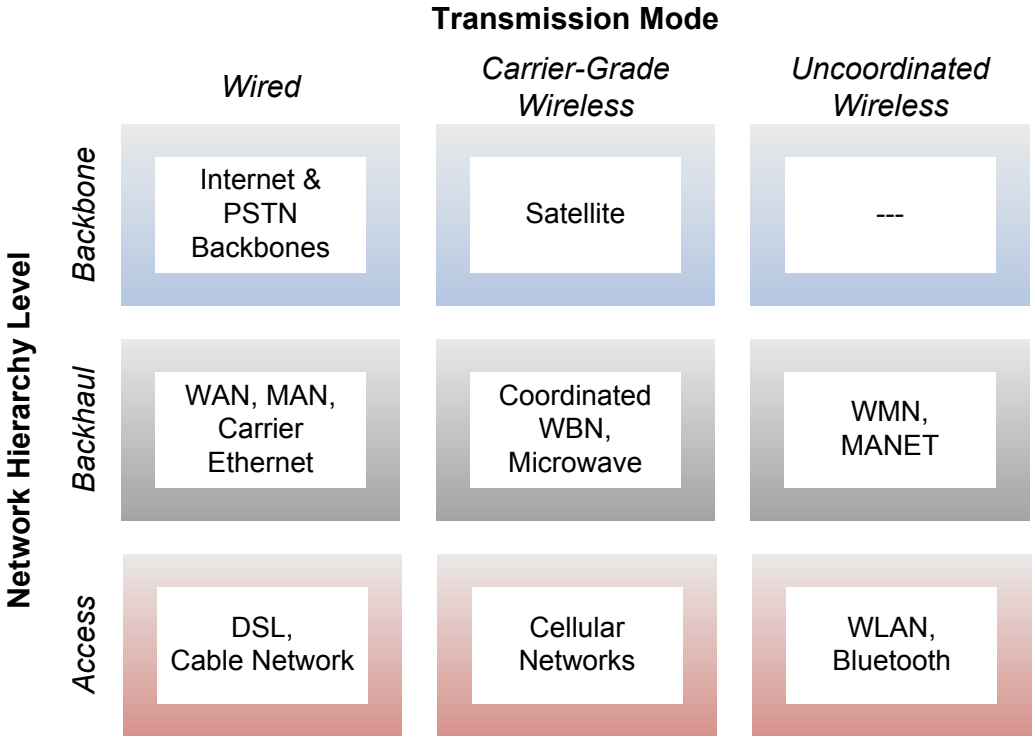


Figure 2.3: Two-dimensional classification of communication networks [MCS13b]

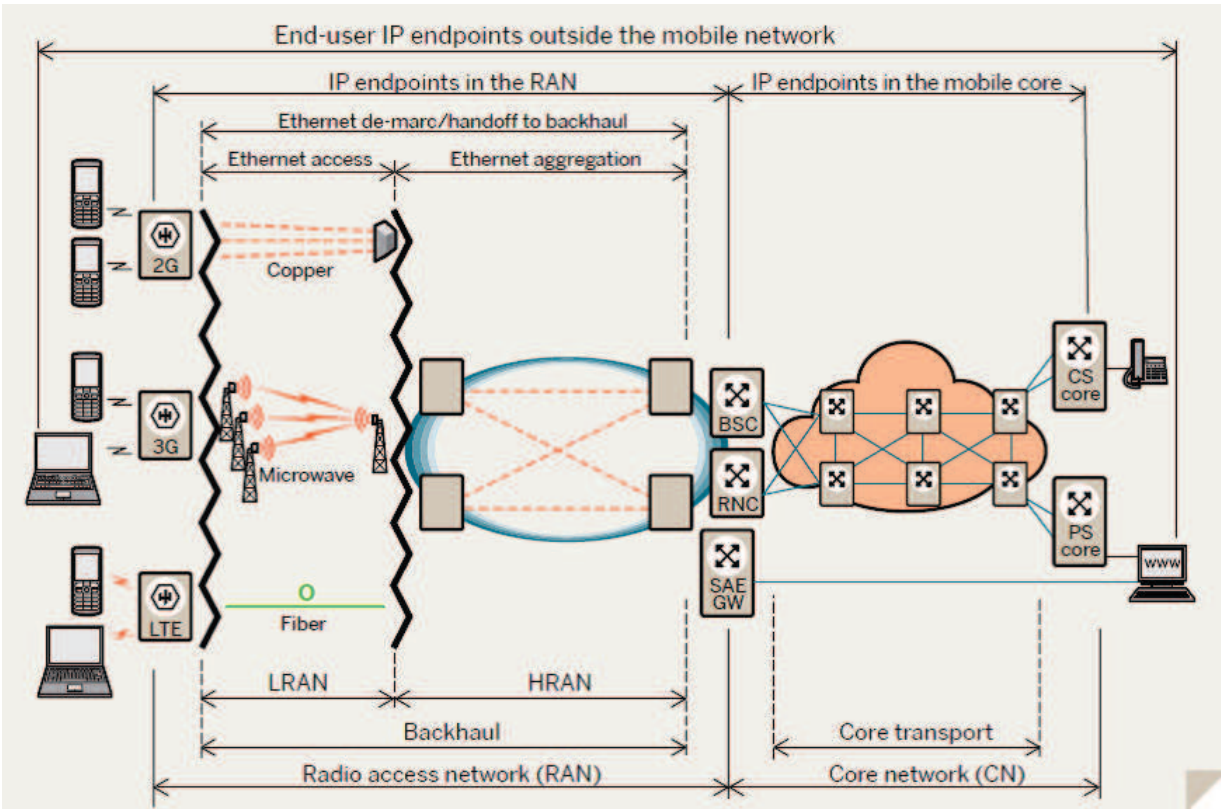


Figure 2.4: Boundaries between access, backhaul, and core network [Chu08]

carrier-grade and *uncoordinated*, mainly referring to spectrum usage and medium access. For example, technologies such as WLAN use an uncoordinated, randomized scheme for medium access (Carrier Sense Multiple Access/Collision Detection (CSMA/CD) or Carrier Sense Multiple Access/-Collision Avoidance (CSMA/CA), respectively), whereas cellular technologies use coordinated schemes (such as, Orthogonal Frequency Division Multiple Access (OFDMA) for Long Term Evolution (LTE) or Code Division Multiple Access (CDMA) for High-Speed Packet Access (HSPA)) and apply coordinated spectrum reuse approaches.

Resulting from the outlined classification, the author will assume the following notion to describe a coordinated WBN (adapted from [Chu08], [FC11], and [Ava10]):

> In a hierarchical telecommunications network, the WBN comprises the intermediate links between the core, or backbone, of the network and the small sub-networks at the edge, using a cost-efficient wireless transmission technology (such as IEEE 802.11). Coordinated WBNs are an evolution of WMNs. They usually consist of stationary nodes and exhibit characteristics of carrier-grade networks. For example, at network level, resilience and fail-safe operations can be guaranteed to sufficient levels, and at service level, key QoS parameters, such as, bit rates, delay, and jitter are kept within acceptable boundaries.

In summary, even though utilizing low-cost solutions, e.g., IEEE 802.11 as preferred radio technology, coordinated WBNs need to exhibit so-called *carrier-grade* [PB10] characteristics, comprising the following general attributes [Met11], [BRR+13]:

- Reliability, i.e., rapidly detect and recover from node, link, or service failures assuring very high service availability to the end user,
- scalability with respect to provisioned bandwidth, service delivery, and geographical coverage,
- guaranteed provisioning of specified QoS classes to service providers and customers,
- support of standardized services (voice, video, data),
- advanced service management, i.e., capability to monitor, diagnose, and centrally manage the network using simple, efficient, and standardized processes.

As an example, multiple player services in telecommunications require predictable service execution and delivery from both operator and customer perspective. Particularly, a sufficient level of reliability in end-to-end performance with respect to Quality of Experience (QoE) (particularly data rates, latency, and jitter) has to be provided anytime and anywhere. Deployment, operation, and maintenance of a WBN has to be simple, robust, and efficient. This includes equipment, technical processes (e.g., network design, network planning, network control, network configuration, optimization, and healing), as well as managerial and administrative processes (e.g., Authentication, Authorization, and Accounting (AAA), security, billing and charging, user management, and implementation of Service Level Agreements (SLAs)) [PB10]. Finally, the network has to be cost- and energy-efficient to keep capital expenditures (CAPEX) and overall OPerational EXpenditures (OPEX) within acceptable boundaries.

Prevalence and Deployment Options

For historical as well as performance reasons, the majority of backhaul networks use wired solutions, among them optical fiber or bonded copper. With the convergence of circuit-switched and packet-switched networks to all-IP next generation networks, carrier-grade Ethernet has become more prominent. Nowadays, while wired networks are still preferred, operators have acknowledged the advantages of WBNs in specific deployment scenarios and first reference implementations have successfully been launched, e.g., for connecting Alpine villages or rural areas of Sub-Saharan Africa, but also campus-style deployments for companies and universities [Aru12]. Figure 2.5 gives an overview of available backhaul technologies and their respective performance in terms of latency and capacity.

The deployment of cost-efficient wireless backhaul networks is beneficial due to multiple reasons, among them

- growing numbers of subscribers, particularly in developing countries,
- exploding demand for ubiquitous broadband access,
- migration to networks with higher capacities like Universal Mobile Telecommunications System (UMTS) (3G) and LTE-A (4G),
- provisioning of services to rural and other sparsely populated areas.

Demand for backhaul capacity continues to rise as consumers get accustomed to enjoying more data and video services on their mobile devices.

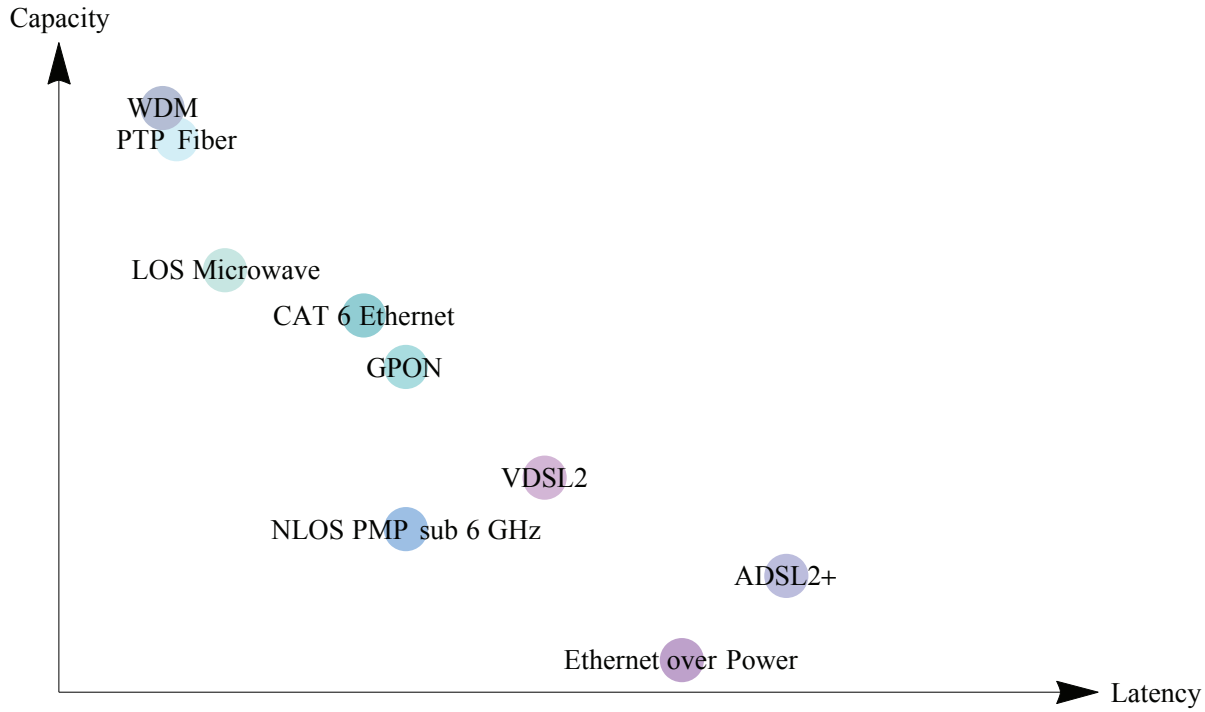


Figure 2.5: Capacity and latency characteristics of various backhaul transmission technologies [Eri12]

Surging deployments of 4th-generation LTE-A cellular networks are poised to boost that demand even further while at the same time promising to deliver an experience similar to wired networks. Radio and microwave links enjoy significant advantages over fiber for 4G backhaul. Superior flexibility, durability, and quicker time to deployment result in lower total cost of ownership (TCO). Even where fiber is already deployed, wireless point-to-point links are the preferred choice to substitute fiber in case of outages or during peak traffic times.

2.3 Summary

In Chapter 2, the notion of coordinated WBNs for mobile broadband has been introduced and defined. The evolution from MANETs to WMNs to coordinated WBNs, which brings together the advantages of cost-efficient radio technologies (e.g., IEEE 802.11) and carrier-grade network management procedures from mobile networks in a single, coherent deployment solution, has been described. The lack of coordinating SON capabilities, e.g., managing autarkic energy supply, in current WMN networks has been identified as one of the most significant limitations. Accordingly, algorithms for autonomous network management comprise a major research challenge

that is addressed in this work. This particularly includes the reliable operation of WBNs relying on renewable energy sources. Here, context-aware Topology Control (TC) algorithms, as further discussed in Chapter 6, can make significant contributions.

Chapter 3

Formal Problem Statement and Evaluation Methodology

The upcoming chapter outlines the chosen evaluation methodology, presents the integrated system model used for simulation-based evaluation, and, based on this, derives the formal problem statement. More specifically, the evaluation methodology is designed in a way to emphasize the multi-domain approach for analyzing context-aware optimization algorithms for coordinated wireless backhaul networks (WBN) and explicitly incorporates two feedback loops for iterative analysis. With respect to the modeling strategy, the models to be developed in Chapter 4 are integrated into a graph-based representation of a WBN in a new manner that allows to flexibly combine analytical tools from graph theory and network evaluation. Moreover, the utilization of a graph model facilitates a formal, efficient, and extensible depiction of the problem statement. Hence, beyond incorporating the aspects relevant in the scope of this work, particularly energy budget management in the light of autarkic supply and autonomous topology control, the problem statement possesses a structure that allows to include further network management aspects. Finally, this chapter introduces a new balanced scorecard to allow for a comprehensive, multi-domain evaluation of the developed optimization algorithms. The developed evaluation methodology is realized in Chapter 8.

3.1 Basic Concepts of Graph Theory

Allowing for a rigorous application of various analytical methods, concepts from graph theory are predestined for modeling point-to-point radio networks. In order to formally state optimization goals and analyze algorithm

performance in the context of WBN operation, some basic notions from graph theory are presented in the following [Die06]. Further concepts that are relevant in the scope of this thesis are depicted in Appendix A. In this section, a short overview is given in Table 3.1, presenting the mathematical concepts and their equivalents in the context of WBNs. Subsequently, a formal definition for each quantity is given.

Table 3.1: Concepts from graph theory and equivalents in a WBN

Graph-theoretical concept	**Equivalent in a WBN**
graph (set of vertices, set of undirected edges)	wireless backhaul network (nodes and duplex links)
path between vertices	route between nodes
connected graph	all nodes are part of the network
vertex degree	number of links at a node
edge weight	link capacity
edge connectivity of a graph	route redundancy level of the network
adjacency matrix	nodes within single-hop distance
incidence matrix	links originating at node

The most fundamental concept in graph theory is a graph itself, which is formed by vertices and edges. This work only considers undirected graphs, since it assumes bidirectional communication over every link of the WBN.

Definition 3.1 (*Graph*)

An undirected graph G is a pair $(\mathcal{V}, \mathcal{E})$ consisting of a vertex set $\mathcal{V} = \{v_1, v_2, \cdots, v_m\}$ and an edge set $\mathcal{E}$, $|\mathcal{E}| = n$. An edge e is uniquely described by a pair of vertices, $e_{i,j} = \{v_i, v_j\}$, $i, j = 1, \ldots, m$. In undirected graphs, edges have no direction associated with themselves, i.e., $e_{i,j} = e_{j,i}$.

Graphs have distinct topological shapes and therefore exhibit specific characteristics. Paths and cycles can be a graph themselves or be found within a graph.

Definition 3.2 (*Paths and Cycles*)

A graph on m vertices is called a path (of length $m-1$) if the vertices can be labeled as $v_1, \ldots, v_m$ in such a way that $\mathcal{E} = \{e_{i,i+1} \,|\, i = 1, \ldots, m-1\}$. A graph is called a cycle (of length m) if the vertices can be labeled as $v_1, \ldots, v_m$ in such a way that $\mathcal{E} = \{e_{i,i+1} \,|\, i = 1, \ldots, m-1\} \cup \{e_{m,1}\}$. A cycle is called odd (even) if m is odd (even).

The issue of reaching a graph vertex from any other vertex in a graph is described by the notion of connectivity. A connected graph is defined as follows.

Definition 3.3 (*Connected Graph*)

A graph G is connected if there is a path between each pair of vertices.

A graph edge is incident to two vertices. Their number of a node's incidental edges is described by the notion of vertex degree.

Definition 3.4 (*Vertex (Node) Degree*)

The degree $d_G(v)$ $(d(v))$ of a vertex (node) v is the number of edges it is incident to. A graph is regular if all vertices have the same degree, and k-regular if all vertices have degree k.

The connectivity of a graph is an important path redundancy metric. While a single path has no redundancy since there is no alternative route for any source-destination pair, a cycle graph always provides at least two possibilities for reaching a destination vertex.

Definition 3.5 (*Edge Connectivity*)

The minimum number of edges $\chi(G)$ whose deletion from a graph G disconnects G is called the edge connectivity. The edge connectivity of a disconnected graph is 0, while that of a connected graph is at least 1. Let $\chi(G)$ be the edge connectivity of a graph G and d_{min} the minimum node degree, then for any graph holds $\chi(G) \leq d_{min}$.

Adjacency and incidence matrices form convenient alternatives for a complete specification of a graph $G = (\mathcal{V}, \mathcal{E})$. The matrices describe the neighboring relations of vertices with each other and the relation of vertices and edges, respectively.

Definition 3.6 (*Adjacency and Incidence Matrices*)

The adjacency matrix of a graph $G = (\mathcal{V}, \mathcal{E})$ is the $m \times m$ $(0-1)$-matrix $\mathbf{A}(G) = (a_{i_1,i_2})$, $i_1, i_2 = 1, \ldots, m$, whose rows and columns are labeled by vertices, with entry $a_{i_1,i_2} = 1$ if and only if $e_{i_1,i_2} \in \mathcal{E}$. The incidence matrix is the $m \times n$ $\{0-1\}$-matrix $\Psi(G) = (\psi_{i,j})$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ whose rows are labeled by vertices and whose columns are labeled by edges, with entry $\psi_{i,j} = 1$ if and only if v_i is incident to e_j.

3.2 Formal Problem Statement

In point-to-point radio networks, Topology Control (TC) is a versatile instrument to effectively execute control over a wide range of parameters, among them network capacity and utilization, power control, routing, latency and jitter, or coverage. Chapter6 will elaborate on the technical details of TC techniques as well as the benefits and costs of applying them to WBNs. In short, TC algorithms can significantly contribute to performance improvements with respect to network coverage and capacity while at the same time extending network lifetime by reducing the load of energy-autarkic network nodes. In the efficient execution of TC, initial network topology and configuration, i.e., the set of network nodes, their coordinates, and the initial link configuration, form a cornerstone. Obviously, the initial configuration of a network can be realized according to several objectives, e.g. maximum capacity, maximum path redundancy by maximizing node degrees, minimizing energy consumption of nodes, maximizing coverage, etc. In the scope of this thesis, given a set of nodes, their coordinates, and maximum node degrees, the author considers the *initial* TC problem of finding the best (according to some objective function) set of links (edges) connecting all nodes of the network. More strictly, the chosen approach does not only strive for a connected graph, rather, the goal is to construct an initial graph with edge connectivity χ approaching maximum node degree $d_{\max}$, i.e., $d_{\max} - 1 \leq \chi \leq d_{\max}$. Equivalently, this problem can be considered from a different perspective, yielding an alternative problem statement. Given a set $\mathcal{V}$ of m ($m \geq 2$) backhaul nodes or vertices, the placement function $\mathbf{p} : \mathcal{V} \rightarrow P \subset \mathbb{R}^2$ assigns two-dimensional coordinates to each node. This function allows for defining a Euclidean distance between any two nodes, resulting in a symmetric, non-negative distance matrix $\mathbf{D}_{m \times m} = (\delta_{i_1,i_2})$. Further, an for each pair of nodes, an adjacency relation a_{i_1,i_2} can be defined as follows:

$$a_{i_1,i_2} = \begin{cases} 1 & \text{if nodes are connected,} \\ 0 & \text{if nodes are not connected.} \end{cases}$$

This results in a second symmetric, so-called adjacency matrix $\mathbf{A}_{m \times m} = (a_{i_1,i_2})$. Now, the basic problem, assuming a set of constraints, consists of assigning each node to a pair of coordinates with the goal of minimizing the inner product of the resulting distances and the corresponding pre-

set adjacency relations. Mathematically, this problem is referred to as a Quadratic Assignment Problem (QAP), which is a fundamental problem from the class of k-center problems in combinatorial optimization and a generalization to a wide range of other well-known problems, among them traveling salesman, linear arrangement, or fixed-size clustering.

In general, the QAP is defined as follows [Cel98]. Let the edge weight matrix $\mathbf{W} = (w_{i_1,i_2})$ and the distance matrix $\mathbf{D} = (\delta_{i_1,i_2})$, $i_1, i_2 = 1, 2, \ldots, m$, be two $m \times m$ symmetric, non-negative matrices. Find the bijective permutation (assignment) $\Pi : \{1, 2, \ldots, m\} \rightarrow \{1, 2, \ldots, m\}$ of nodes that minimizes

$$U(\Pi) = \sum_{i_1,i_2 \in \{1,2,\ldots,m\},\, i_1 \neq i_2} w_{i_1,i_2}\, \delta_{\Pi(i_1),\Pi(i_2)}, \tag{3.1}$$

which is also referred to as the inner product between the weight and distance matrices.

Except for a few cases with specific configurations of $\mathbf{W}$ and $\mathbf{D}$, this problem is generally proven to be Non-deterministic Polynomial time (NP)-hard, i.e., the optimal solution cannot be found in polynomial time [Cel98]. As such, several algorithms exist that aim at approximating the optimal solution of the problem as closely as possible. For example, Nagarajan et al. [NS09] present an $\mathcal{O}\left(\frac{1}{\sqrt{m}\log^2 m}\right)$-approximation algorithm. Using common scaling arguments, they reduce both matrices $\mathbf{W}$ and $\mathbf{D}$ so that they exclusively have {0,1}-entries. In that case, $\mathbf{W}$ and $\mathbf{D}$ can be considered as adjacency matrices of two undirected graphs and Equation (3.1) can be maximized by finding a permutation Π, such that the number of common edges in $\mathbf{W}$ and $\Pi(\mathbf{D}) = (\delta_{\Pi(i_1),\Pi(i_2)})$ (and hence the inner product) is maximized. Further approximation algorithms and their according approximation level have been obtained by Arkin et al. (*capacitated star graph packing* [AHRS04]) or Hassin et al. (*maximum clustering with given cluster sizes* [HR06]).

Transferring the general QAP problem statement to the specific case of *topology control* in WBNs as defined above, the adjacency matrix $\mathbf{A}$ replaces the weight matrix $\mathbf{W}$. Further, the permutation function Π determines the assignment of nodes to coordinates. Consequently, the goal is to find an assignment (permutation) of nodes to locations that minimizes

$$Q_{TC}(\Pi) = \sum_{i_1,i_2 \in \{1,2,\ldots,m\}, i_1 \neq i_2} a_{i_1,i_2}\, \delta_{\Pi(i_1),\Pi(i_2)}. \tag{3.2}$$

Conceptually, the distance matrix $\mathbf{D}$ exhibits significant versatility. Be-

sides simply implementing Euclidean distance between a pair of nodes, $\mathbf{D}$ can also incorporate more abstract distance concepts, e.g., required costs for connecting two nodes. Further, quantities with a negative correlation to distance, such as link capacity, are accounted for by specifying a function that appropriately maps the according quantity to a distance value (e.g., $\delta(c) = c^{-r}, r \in \mathbb{R}$). Apart from optimizing Equation (3.2), maximizing path redundancy in a communication network forms another important design goal. The notion of path redundancy, i.e., the availability of multiple, disjunctive paths from one node to another, is equivalent to the graph-theoretical concept of edge connectivity (as stated in Definition 3.5), which is analyzed in so-called *minimum cost k-edge-connected spanning subgraph* problems [Die06]. In these problems, starting from a given graph $G = (\mathcal{V}, \mathcal{E})$, algorithms determine a k-edge connected subgraph $G' = (\mathcal{V}, \mathcal{E}') \subseteq G$, $\mathcal{E}' \subseteq \mathcal{E}$, with minimum (edge) costs. This category of optimization problems is known to be NP-hard even for $k = 2$ [GJ03] and therefore generally requires heuristics for an efficient approximation of the optimal solution. QAPs and minimum cost k-edge-connected spanning subgraph problems are revisited in Section 6.2, where they form the basis to develop two algorithms for approximating optimal topologies in WBN.

In summary, Problem 1.1 as presented Section 1.3 can now be transformed into a formal (preliminary) problem statement:

Problem 3.3

$$\begin{aligned}
\text{minimize} \quad & -Q_{\mathrm{TC}}(\mathbf{A}, \Pi(\mathbf{D})) - E(\mathbf{A}, \Pi(\mathbf{D})) - F && \text{(3.3a)}\\
& = \alpha_Q \sum_{i_1, i_2} a_{i_1,i_2}\, \delta_{\Pi(i_1),\Pi(i_2)} - E(\mathbf{A}, \Pi(\mathbf{D})) - F \\
\text{s.t.} \quad & -\chi(G) \;\leq\; -k && \text{(3.3b)}\\
& \chi(G) \;\leq d_{\max} && \text{(3.3c)}\\
& \mathbf{A} - \mathbf{A}^{\mathrm{T}} \;= 0 && \text{(3.3d)}\\
& a_{i,i} \;= 0 \;\; \forall\, i \in \{1, \ldots, m\} && \text{(3.3e)}\\
& a_{i_1,i_2} \;\in \{0, 1\} \;\; \forall\, i_1, i_2 \in \{1, \ldots, m\} && \text{(3.3f)}
\end{aligned}$$

Here, constraints (3.3b) and (3.3c) control edge connectivity, constraint (3.3d) guarantees duplex (bidirectional) radio links, and constraint (3.3e) excludes loops. α_Q is a scaling factor to appropriately weigh the QoS-related part of the objective function. The following sections elaborate on the selected approach for modeling a WBN. This will allow the further

refinement of problem statement 3.3 and will pave the way to develop and apply various heuristic optimization techniques.

3.3 Evaluation Methodology

The complexity of deploying and operating a WBN stems from the numerous degrees of freedom developers and operators face. However, the availability of numerous design and control parameters also allows for customized networking solutions exhibiting high flexibility and superior capabilities to adapt to fluctuating throughput and latency requirements. It comprises a highly challenging task to combine available control levers in a way as to improve network performance according to a given set of interdependent (and possibly conflicting) objectives.

This challenge of designing a multi-domain network model and subsequently evaluating developed optimization strategies is tackled by a novel event-based simulation tool developed as part of this work. Its structure follows the methodological approach illustrated in Figure 3.1. A structured sequence of *measure-analyze-act* steps is the fundamental basis for all quantitative, simulation-based evaluations within the *integrated modeling and evaluation framework*. The first step comprises a thorough modeling of both the physical and the network domain of coordinated wireless point-to-point networks. Here, the author takes well accepted models as presented in Chapter 4, adapts and, if necessary, extends them. The simulation tool implements these models, which in turn allows to monitor temporal dependencies and interactions. An integrated context management system facilitates the acquisition of so-called *raw* context data as well as the inference of *semantically rich* context, as detailed in Section 4.3. Subsequently, these data can be fed into different kinds of dynamic algorithm classes, among them *Optimization and Search*, *First-Order Logic*, or *Probabilistic Reasoning* approaches. In conjunction with rather static rules from a policy repository (e.g., on minimum coverage requirements), the outputs of the algorithms determine the setting and adjustment of available control parameters in different control schemes, such as, *Bandwidth Adaptation*, *UE Balancing*, or *Backhaul Topology Optimization*. In turn, these adjustments directly affect model parameters of the simulation tool. Moreover, they determine key performance indicators (KPI) and metrics of the so-called *balanced scorecard* used for evaluating the effectiveness of the applied optimization techniques.

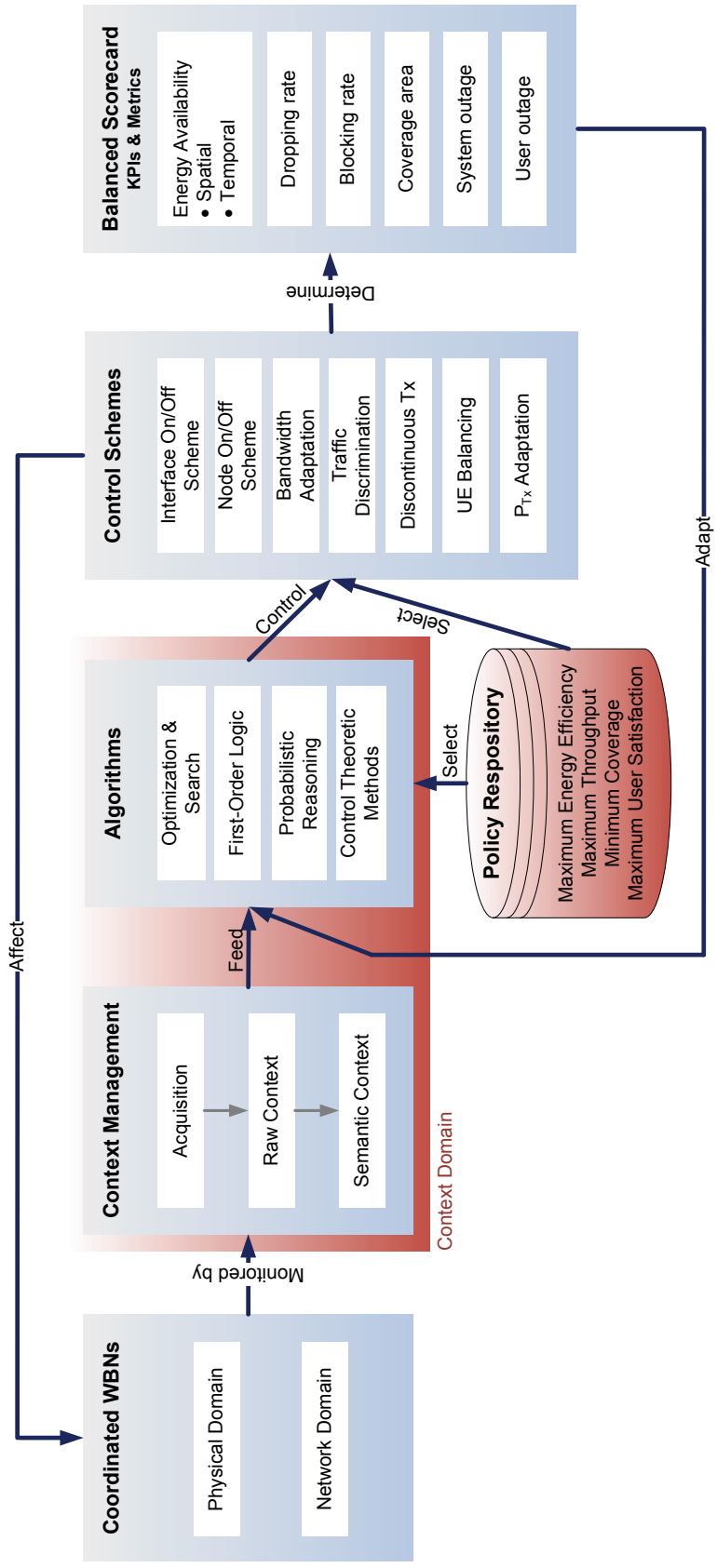


Figure 3.1: Methodology for an integrated modeling and evaluation framework

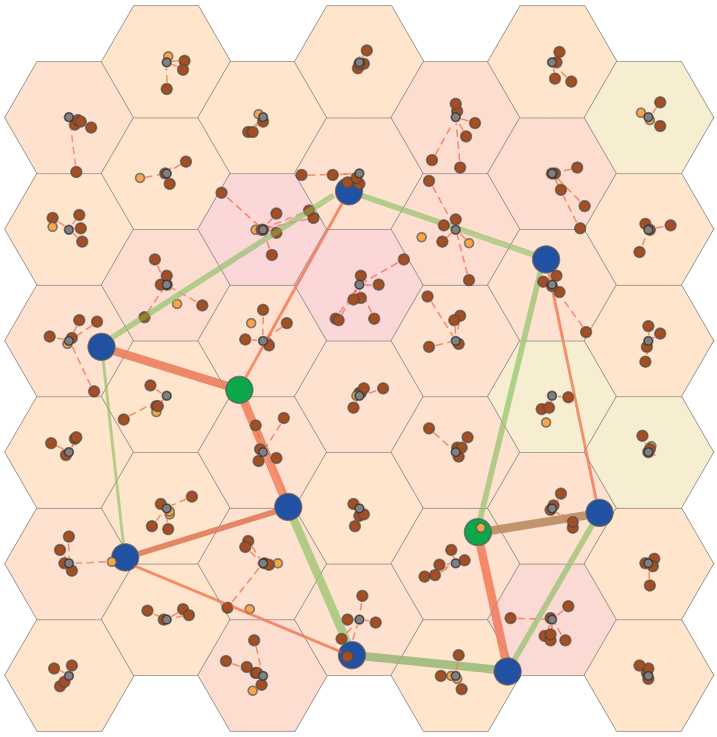
Typical KPIs include, but are not limited to, system and user outage, energy efficiency and availability, User Terminal (UT) dropping and blocking, as well as average throughput per user. The major objective of the balanced scorecard is to properly reflect the multi-criteria objective function by assembling an according set of metrics, which in turn are used to adapt algorithmic parameters. In this manner, the simulation framework incorporates two overlapping feedback loops:

- The **Upper Feedback Loop** assures a continuous consideration of optimization results within the simulation model, i.e., proposed changes in parameter settings are executed and their effectiveness can be evaluated subsequently.
- The **Lower Feedback Loop** implements a coupling between KPIs and algorithm parameters. This is introduced to influence algorithm behavior in the desired way.

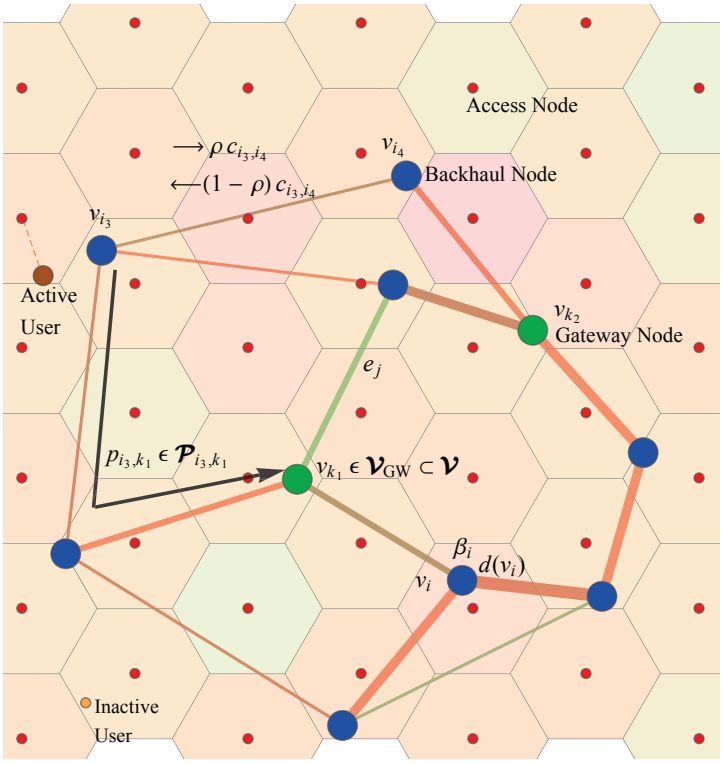
The presented integrated modeling and evaluation framework serves as a generic frame that can house many different approaches, algorithms, and evaluation metrics for different classes of networks. The upcoming sections will cover the implementation of selected methods and metrics within the framework, thereby depicting the benefits of the approach.

3.4 A Graph Model of the Backhaul Network

An efficient representation of the considered wireless point-to-point networks is essential for any optimization method that shall be applied to the network. As mentioned before and shown in Figure 3.2(a), a common way to model point-to-point radio networks is the representation as a graph since many of their characteristics can be transferred into according quantities of graph theory. This does not only include backhaul nodes themselves, but can be extended to access points and users, that can be introduced as a second and third type of vertex, respectively. Quantities, such as, edge capacity and weight, vertex weight, vertex degree, etc. have been chosen to model the according characteristics of the network (link capacity and utilization, battery level, number of links to other nodes, etc.). The utilized model is an extension of a similar, widely used model originating from the domain of wireless ad-hoc networks. More specifically, the author uses the subsequent notation that is also depicted in Figure 3.2(b).



(a) High degree of meshing



(b) Overview of notations and symbols

Figure 3.2: Topology examples of a WBN [MCS13d]

Let $\mathcal{V}$ denote the set of backhaul nodes or vertices v_i, $i \in \{1, \ldots, m\}$, of an undirected graph G and $\mathcal{E}$ the set of edges e_j, $j \in \{1, \ldots, n\}$, of G. Further, let $d(v_i)$ denote the vertex degree of vertex v_i. The subset $\mathcal{V}_{\text{GW}} \subseteq \mathcal{V}$ is called the set of gateway nodes in G, $|\mathcal{V}_{GW}| \in \mathbb{N}^+$. Gateway nodes connect a WBN to the backbone network, cf. Figure 2.4. The percentage of gateway nodes among all nodes is given as $g = \frac{|\mathcal{V}_{GW}|}{m}, g \in]0, 1]$. Vertex coordinates are determined according to a so-called placement function $\mathbf{p} : \mathcal{V} \times \mathcal{T} \rightarrow [0, l_1] \times [0, l_2] \times [0, l_3]$, limiting the considered space to a cuboid of side lengths $l_1, l_2, l_3 > 0$. For any time $t \in \mathcal{T}$, $\mathbf{p}$ assigns a three-dimensional position $\mathbf{p} = (p_1, p_2, p_3)$ to every node $v \in \mathcal{V}$. For coordinated WBNs, the author generally assumes stationary nodes that are located in a two-dimensional plane, i.e., the nodes' locations do not vary with time and elevation is not considered. Thus, $\mathbf{p}$ simplifies to $\mathbf{p} : \mathcal{V} \rightarrow [0, l_1] \times [0, l_2]$. $\mathcal{P}_{i,k}$ denotes the set of paths with non-zero capacity from a node v_i to a gateway node $v_k \in \mathcal{V}_{GW}$. Further, let $\mathbf{A}(G) = (a_{i_1,i_2})$ be the $m \times m$ adjacency matrix of G, $a_{i_1,i_2} \in \{0, 1\}$. Since G is undirected, $\mathbf{A}(G)$ is a symmetric matrix. Accordingly, $\mathbf{C} = (c_{i_1,i_2})$ denotes the $m \times m$ capacity matrix of G. The edges, modeling full-duplex wireless links, split the capacity between both directions according to a ratio $\rho_{i_1,i_2} = 1 - \rho_{i_2,i_1}$, $i_1 \neq i_2$, $\rho \in [0, 1]$. This work fixes $\rho_{i_1,i_2} = \rho_{i_2,i_1} = 0.5 \; \forall \; i_1, i_2 \in \{1, \ldots, m\}$. $\boldsymbol{\pi}_{\text{in}} = (\pi_{1,\text{in}}, \ldots, \pi_{m,\text{in}})$ denotes the m-dimensional vector of power consumption of the nodes v_i. $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m)$ denotes the m-dimensional vector of battery charging levels (in %) of the nodes v_i. Furthermore, each vertex v_i hosts a set $\mathcal{X}_i$ of wireless radio interfaces that are used for bidirectional communication to neighboring vertices ($|\mathcal{X}_i| = N_i$). Consequently, the degree $d(v_i)$ of a vertex is bound to an upper limited, $d(v_i) \leq N_i$. Using the path loss models as depicted in Section 4.2.1, the *power assignment* $\pi_{i,\text{out}} : \mathcal{X}_i \rightarrow [0, \pi_{\text{out,max}}]$ maps a transmission power between 0 and $\pi_{\text{out,max}}$ to each radio interface of a backhaul node v_i. Assuming a transmission range function $\delta(\pi_{\text{out}})$ that maps a coverage range to each power level π_{out}, the condition $||\mathbf{p}(v_1) - \mathbf{p}(v_2)|| \leq \delta(\pi_{\text{out,max}})$ is a prerequisite for a joint edge of an arbitrary pair of nodes v_1, v_2, i.e., v_1 and v_2 need to be placed in their mutual coverage area. Figure 3.2(b) depicts a graphical summary of the defined notations and symbols.

3.5 A Novel Balanced Scorecard Concept for Multi-Domain Evaluation

Performance evaluation of wireless communication systems serves the purpose of quantitatively verifying if a developed system (in terms of both hardware and software) meets the developer's expectations with respect to a pre-defined set of metrics, according threshold values, and use case scenarios. Going beyond conventional one- or multi-dimensional evaluation metrics, one of the contributions of this thesis is the design of a holistic *Balanced Scorecard (BSC)* for a coordinated WBN. The concept originates from strategic business management and has first been published by Kaplan and Norton [KN92]. The control tool links performance objectives and measures from different domains and perspectives (in the original sense from different organizational entities in a company, e.g., research and development, production, marketing and distribution, finance, etc.) and to translate strategies into operational actions. Therefore, a correspondingly designed BSC for WBNs not only serves the purpose of a comprehensive causal performance evaluation of wireless point-to-point networks exceptionally well, but can also be employed for facilitating the efficient operation of such a network. In the context of WBNs, the author will consider an adapted BSC featuring two evaluation dimensions. The first dimension distinguishes the physical domain and the network domain, as introduced in Sections 4.1 and 4.2, respectively. The second dimension differentiates between user, node, and system level evaluation metrics. In order to completely cover this two-dimensional space, several existing and novel metrics will be introduced in the following.

3.5.1 Network Domain Metrics

This section, besides introducing some established evaluation metrics for wireless systems, will cover some specific, custom-tailored KPIs for the analysis of network domain performance of WBNs.

Data Throughput and Capacity

Network capacity constitutes one of the most important parameters in the analysis of wireless networks since it has a significant impact on the approaches for network design, deployment, and operation. Generally, the notion of capacity has multiple dimensions. First, it differentiates between

access and backhaul capacity. Further, this work, which focuses on a WBN using a shared medium and uncoordinated multiple access schemes, separates between gross and net capacity. Moreover, network capacity can be measured using different base units: bandwidth in Hz, maximum data throughput in bit/s, or number of users. Finally, most of the quantities can be considered on different levels, e.g., overall system capacity or throughput per user. For the latter, posterior statistics on the allocation of physical radio resources yield very exact figures, not only on average capacity per user but also on distribution of available capacity and, hence, on fairness aspects.

Maximum throughput of a backhaul network is mainly determined by three factors: capacity of individual wireless links in the network, WBN topology configuration, and number and spatial distribution of gateway nodes. Assuming a graph representation of a WBN, as depicted in Section 3.4, a network's capacity with multiple source and sink vertices can be computed by introducing two virtual nodes, a single source and a single target node that connect to all source and target nodes, respectively. The new edges are assumed to provide unlimited capacity. Thus, the problem is reduced to the well-known maximum s-t-flow problem, which can be solved using the Fulkerson-Ford algorithm [FF55]. This algorithm works with residual graphs, where edge capacity is continuously reduced until no path from s to t with non-zero capacity remains. In case of real-valued edge capacities, the algorithm terminates after a finite number of iterations. An enhanced implementation has been developed by Edmonds and Karp [EK72], which, based on the technique of breadth-first search, guarantees the calculation of maximum s-t-flow in polynomial time. For computing the maximum flow for any possible combination of access and gateway node in a WBN, one can break down backhaul capacity to throughput per user, given the amount of connected users is known for every access node. While histograms, depending on bin width, contain more detailed figures, empirical mean and variance give aggregated information.

Fairness of Throughput Allocation

In order to evaluate the impact of WBN operation schemes on the fairness of throughput allocation to users, the analysis uses two metrics: Jain's fairness index as introduced in Section 6.3.2 and the Gini coefficient [CV12]. The latter constitutes a well-established metric to analyze inequalities in

income distributions and therefore also qualifies for evaluating fairness of bandwidth allocation in coordinated WBNs. For the given vector $\boldsymbol{\nu} \in \mathbb{R}_+^N$ of average data rates of N users, the Gini coefficient is computed according to

$$G(\boldsymbol{\nu}) = \frac{\sum_{i=1}^{N} (2i - N - 1)\, \boldsymbol{\nu}}{N \sum_{i=1}^{N} \boldsymbol{\nu}}. \tag{3.4}$$

Similar to Jain's index, $G(\boldsymbol{\nu}) \in [0, 1]$; however, $G = 0$ is equivalent to a perfectly uniform distribution, whereas a rising value of G means increasing inequality.

Delay and Outage Metrics

Latency belongs to the most crucial performance criteria in communication networks since the number of delay-sensitive services has steadily been increasing. Delay definitions vary depending on the context they are used in. Some definitions are restricted to propagation delay, others include end-to-end transmission delay, even more comprehensive ones include processing delay on application layer. Common quantities for delay include average delay per user, Cumulative Distribution Function (CDF) of delay, or x-percentile of delay.

In a second analytical step, outage metrics are directly derived from delay figures. For example, for a Voice over IP (VoIP) service, a user is generally presumed to be in outage if more than 2% of VoIP packets is not delivered successfully within a delay bound of 50 ms [SZJ+08]. Similar rules with adapted thresholds exist for other service classes, e.g., HyperText Transfer Protocol (HTTP)-based services or video streaming. Moreover, system outage is reached if more than 2% of users are in outage, independent of their current service. Finally, throughput outage, which is the percentage of users with an average throughput below a (service-specific) threshold value, is another common outage metric.

Coverage Metrics

Since the context-enabled optimization of point-to-point backhaul networks forms the focus of this work, this section only presents a selected overview on common coverage metrics. However, for reasons of comprehensiveness, the author also computes coverage metrics in the course of algorithm evaluation. In cellular networks, in particular cell radius (which in turn is a function of transmission power) and inter-site distances determine cover-

age. Common models for cellular network evaluation usually assume regular hexagonal layouts, so that the considered area is — technically speaking — completely covered. In these cases, received signal strength measurements and Signal-to-Interference-plus-Noise Ratio (SINR) statistics are a common means to evaluate radio systems. In the uplink, CDF of SINR observed by a Base Transceiver Station (BTS) for each UT serves this purpose. Accordingly, in the downlink, CDF of SINRs observed by each UT is analyzed. Now, coverage range is defined as the maximum radial distance that provides a fix percentage of the area with a SINR above a target SINR during a certain time [SZJ+08].

The situation becomes more complicated in cases where access nodes are not assumed to be evenly distributed, e.g., in Wireless Sensor Network (WSN) deployments, where nodes are distributed in a rather random fashion. Here, different concepts for coverage analysis come into consideration. Huang and Tseng [HT05] propose an algorithm of $\mathcal{O}(m\,\hat{m}\log\hat{m})$-complexity where, given a natural number k, the so-called *k-Non-unit-disk Coverage (k-NC)* problem is solved by determining whether all points in an area are k-covered or not. Here, k-coverage means that a position is within the coverage area of at least k nodes. Further, m is the number of nodes and $\hat{m}$ is the maximum number of sensors that may interfere with a given sensor's transmissions.

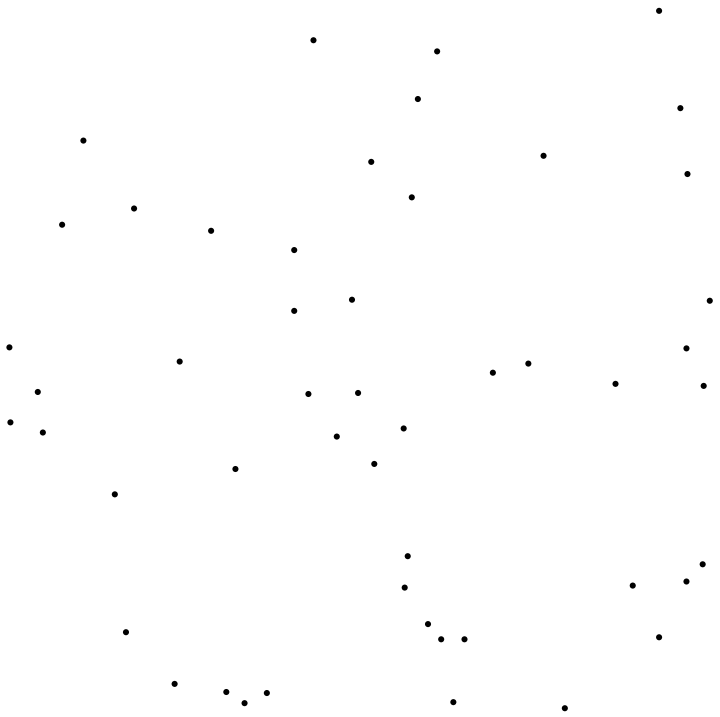
Another common coverage analysis is based on the so-called Voronoi diagram, e.g., [MKPS05], [Mul10]. The Delaunay triangulation, as depicted in more detail in Figure 3.3, is an equivalent dual representation of a Voronoi diagram (cf. Section 6.1.3). Given a set $\mathcal{V}$ of nodes v_i including their coordinates $\mathbf{p}(v_i)$, the Voronoi diagram constructs $|\mathcal{V}|$ cells. An arbitrary point in a plane P with coordinates $\mathbf{p}$ belongs to cell C_i if v_i is the closest node according to a distance function δ, i.e.,

$$C_i = \{\mathbf{p} \in P \,|\, \delta(\mathbf{p}, \mathbf{p}(v_i)) \leq \delta(\mathbf{p}, \mathbf{p}(v_k))\, \forall\, k \neq i\}\,. \tag{3.5}$$

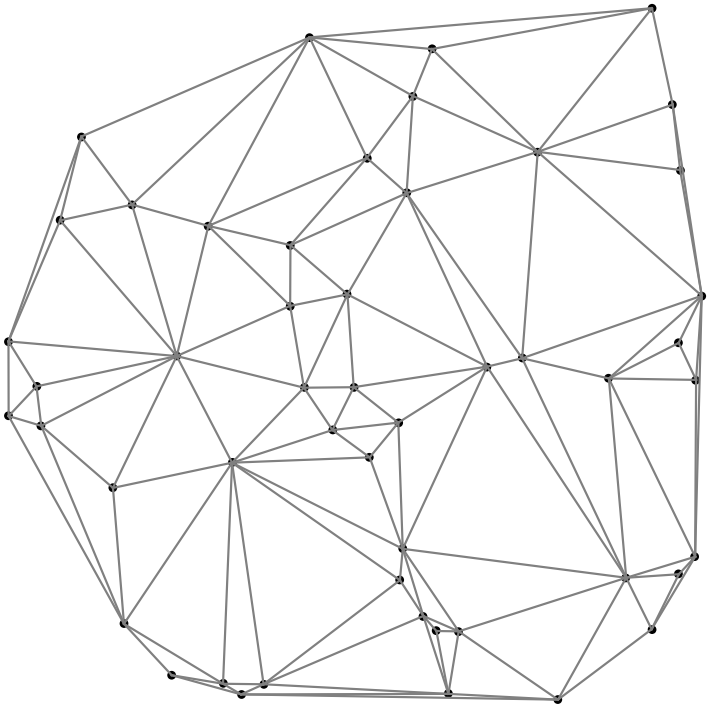
Based on resulting distribution, mean cell radius or area can be computed. Alternatively, CDFs of these quantities give a more detailed analysis with respect to the coverage characteristics of a given access node distribution.

3.5.2 Physical Domain Metrics

The most important aspect of physical domain evaluation of WBNs includes the analysis of energy consumption and efficiency, which requires



(a) Random set of vertices



(b) Resulting planar graph with maximized minimum angles

Figure 3.3: Construction of a Delaunay triangulation

technology-agnostic metrics covering various aspects of energy efficiency. While some systems outperform others in one metric, e.g., power per area, another metric, such as energy per throughput, might yield a completely different system ranking. Moreover, as mentioned above, energy consumption can be considered on different levels and therefore different indicators for energy efficiency have evolved as well. This section gives a brief overview on state-of-the-art energy efficiency metrics.

Lifetime Extension Measure

The Lifetime Extension Measure (LEM) is a novel metric introduced to cover two aspects of the balanced scorecard for WBNs: First and most importantly, it serves as a metric for evaluating the impact of optimization techniques, in particular with respect to lifetime extension of nodes and the entire network. Second, it also indicates the effectiveness of algorithms that aim at improving energy efficiency [MCS13a]. As stated in, e.g., [OE07], many of the available lifetime metrics suffer from (depending on the context) minor or major drawbacks, among them incompatibility with the specific application requirements (e.g., routing optimization [SGL11]) or the incorporation of additional quantities by using weighted sums or utility functions ([Ram08],[OE07]), thus diluting the weight of the lifetime quantity. Finally, many metrics fall short of including a scaling factor that serves as a baseline and as an indicator of relative improvement. In contrast, the LEM takes a network-centric (operator) perspective and uses a novel design approach, which yields several advantages. First, unlike many metrics, instead of determining lifetime based on the first failing node, it considers a percentage threshold of failed nodes that needs to be surpassed (e.g., 80%). Second, it measures the relative improvement in lifetime by comparing it to an upper bound (best case scenario). Finally, while suiting the specific scenario for WBNs, it maintains an acceptable degree of generality and is thus suitable for other application areas as well.

Figure 3.4 illustrates the basic construction of the measure. It shows the successive failure of nodes over time due to draining batteries. The dashed curve limiting area A_1 to the left presents the ratio of failed nodes in standard operation mode (no optimization applied), while the dashed curve limiting A_1 to the right shows the failure ratio when a lifetime extension algorithm is applied. Further, the continuous line on the very right indicates an upper bound reference for maximum operation time (lifetime) of

a backhaul node, which can, for example, be derived analytically, as shown in Section 7.1.

Upper bounds usually depend on, among others, network topology, node distribution and density, incurred user traffic, and required network capacity of the given network configuration. Clearly, area A_1 already indicates the improved performance if the algorithm is applied. In order to quantitatively capture the optimization gain, the author has introduced LEM as the ratio of the two areas A_1 and $A_1 + A_2$ shown in Figure 3.4:

$$LEM = \frac{G_{\text{opt}}}{G_{\text{max}}} = \frac{A_1}{A_1 + A_2}. \tag{3.6}$$

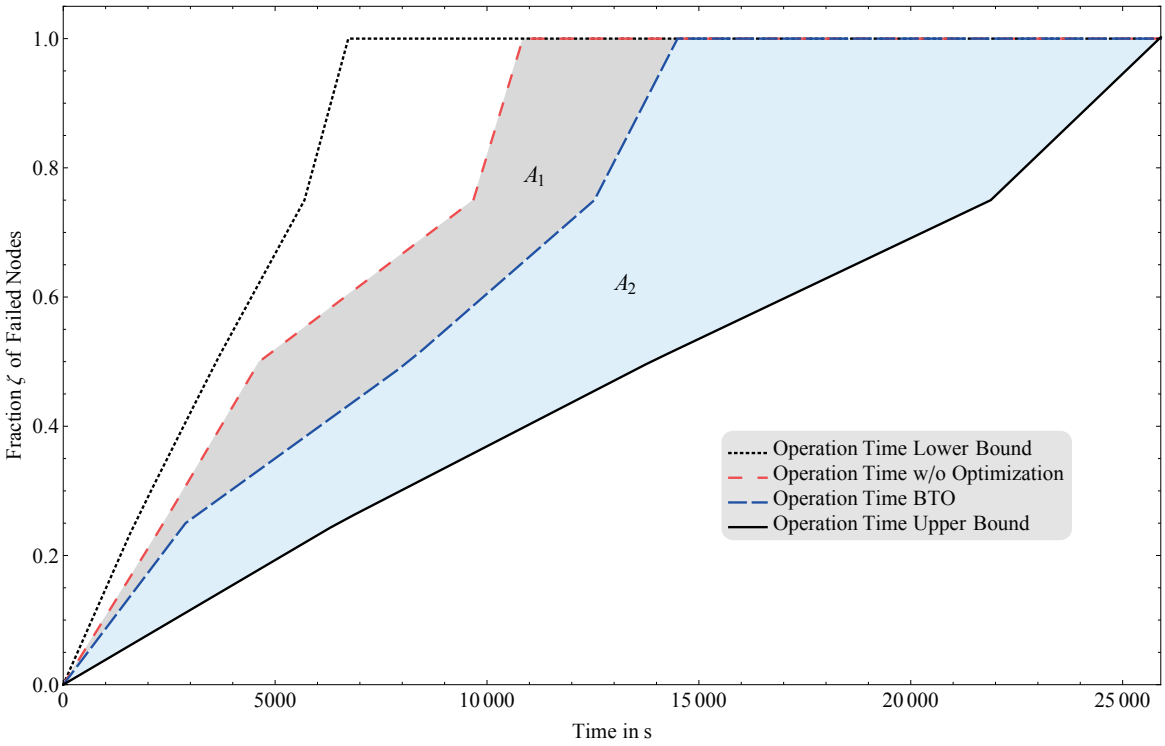


Figure 3.4: Illustration of Lifetime Extension Measure (LEM) [MCS13a]

A_1 is an absolute measure for the gain G_{opt} achieved by the optimization algorithm(s), whereas $A_1 + A_2$ serves as a measure for the gain G_{max} that can theoretically be achieved in the best case, thus scaling G_{opt} to a quantity that can be compared across different systems. The higher LEM, the more effective the impact of optimization with respect to conserving energy and extending network lifetime. For the exemplary network configuration of Figure 7.1, the ratio evaluates to $LEM = 0.2639$, i.e., the algorithm has achieved approx. 26% of the (theoretical) maximum improvement. In Section 7.1, a fundamental analytical assessment of upper and lower bounds

for energy consumption and lifetime of WBN nodes and how LEM relates to them is presented.

Overall Energy Consumption

One of the most obvious metrics is overall energy consumption [LKWG11] of a system. Being defined as the temporal integral over instantaneous power $\pi(t)$

$$E_{\text{tot}} = \int_{t_1}^{t_2} \pi(t)\, dt, \tag{3.7}$$

the metric also includes information about temporal variability of $\pi(t)$ between time t_1 and t_2. However, coming short of any performance indicator used for normalization, it can only give a vague idea of a system's energy efficiency since it does not give any information about the achieved system performance, e.g., with respect to served users or redundancy levels to prevent system failure. Peak power consumption, a similar quantity that is important in the dimensioning of power supply for a network, suffers from the same disadvantages.

Output-Input Power Efficiency Ratio

In analogy to traditional efficiency ratios from other areas, such as mechanical engineering, the ratio of output and input power, π_{out} and π_{in}, is used to measure energy efficiency:

$$\eta = \frac{\pi_{\text{out}}}{\pi_{\text{in}}}. \tag{3.8}$$

For wireless systems, input power comprises the total power supplied to the network, whereas output power usually refers to the transmission power, i.e., the output power at the amplifier [CZB+10]. However, due to its limited applicability (generally, the metric is only used to evaluate the efficiency of a transmitter, such as a cellular base station), other indicators are required for a more comprehensive system analysis.

Energy per Bit and Throughput per Watt

A commonly used metric for measuring the energy efficiency is the consumed power or energy per performance unit [CKY10]. Depending on the considered system, performance can be measured in various metrics, e.g., throughput in bit/s or the amount of transferred data in bits or bytes.

Thus, according metrics for energy efficiency are usually expressed by ratios, such as total energy E_{tot} over total data transferred ($\overline{\nu}\,(t_2 - t_1)$) or peak power consumption π_{pk} over average data rate $\overline{\nu}$

$$\frac{E_{\text{tot}}}{\overline{\nu}\,(t_2 - t_1)} = \frac{\pi_{\text{avg}}}{\overline{\nu}}, \tag{3.9}$$

and

$$\frac{\pi_{\text{pk}}}{\overline{\nu}}, \tag{3.10}$$

Reciprocal metrics, where performance is normalized using an energy or power measure, e.g., $\frac{\overline{\nu}}{E_{\text{tot}}}$ in bit/s/J constitute equivalent efficiency indicators [ABG+10]. Meanwhile, numerous variations of these basic metrics have been developed (among them Energy Consumption Rating (ECR) [ANK10], Telecommunications Energy Efficiency Ratio (TEER) [ATI09], or Telecommunications Equipment Energy Efficiency Rating (TEEER) [Tal09]). The common drawback of these metrics is their limited applicability to networks not operating at full load constantly, such as cellular networks. For them, a more suitable objective is to minimize the power for covering a certain area [CZB+10].

Power per Covered Area

For cellular systems, coverage is a major design and optimization issue, particularly in scenarios with low traffic loads. Consequently, along with power per throughput, power per covered area A_{cov} becomes an important metric [ABG+10]:

$$\pi_{\text{cov}} = \frac{\pi_{\text{avg}}}{A_{\text{cov}}}. \tag{3.11}$$

Generally, π_{cov} is a useful metric in case access networks are considered since they aim at covering the given area to the greatest possible extent. However, the foremost objective on the backhaul side of coordinated WBNs is to provide reliable broadband point-to-point links from one wireless router to the next. Clearly, π_{cov} is a considerably less useful energy efficiency indicator in this case.

User-Specific Energy Efficiency

Another class of metrics utilizes user-related figures to normalize statistics on energy consumption. This again is particularly useful for access net-

works, e.g., in order to compare their performance from a UT perspective. Suggested indicators include consumed energy per subscriber

$$E_{\mathrm{UT}} = \frac{E_{\mathrm{tot}}}{N}, \tag{3.12}$$

energy per subscriber per throughput

$$E_{\mathrm{UT},\nu} = \frac{E_{\mathrm{tot}}}{N \cdot \nu_{\mathrm{avg}}}, \tag{3.13}$$

or energy per subscriber per coverage area

$$E_{\mathrm{UT},A} = \frac{E_{\mathrm{tot}}}{N \cdot A_{\mathrm{cov}}}. \tag{3.14}$$

The disadvantage of user-specific metrics is their tendency to treat systems with higher user density more favorably, e.g., evaluating urban networks better than those in rural areas. Further, networks operating at full capacity receive better evaluation, despite the fact that user experience usually deteriorates in such situations. These issues reflect the general problem of exclusively using a single metric as a means to measure energy efficiency. Biases have to be accounted for by applying different perspectives (i.e. heterogeneous metrics) to the analysis.

3.6 Summary

Chapter 3 has motivated the formal structure of the problem statement. Objective function and a first set of constraints have been identified. Incorporating the concepts and specific constraints presented in this chapter, the optimization problem 1.1 introduced in Section 1.3 and extended in Section 3.2 can now be refined as follows:

Problem 3.15

$$\text{minimize} \quad \alpha_Q \sum_{i_1,i_2} a_{i_1,i_2}\, \delta_{\Pi(i_1),\Pi(i_2)} - E(\mathbf{A}, \Pi(\mathbf{D})) - F \tag{3.15a}$$

$$\text{s.t.} \quad -\chi(G) \leq -k \tag{3.15b}$$

$$\chi(G) \leq d_{\max} \tag{3.15c}$$

$$\mathbf{A} - \mathbf{A}^{\mathrm{T}} = 0 \tag{3.15d}$$

$$a_{i,i} = 0 \;\; \forall\, i \in \{1, \ldots, m\} \tag{3.15e}$$

$$a_{i_1,i_2} \in \{0, 1\} \;\; \forall\, i_1, i_2 \in \{1, \ldots, m\} \tag{3.15f}$$

$$-\sum_{i_2} a_{i_1,i_2} \leq -d_{\min} \;\; \forall\, i_1 \in \{1, \ldots, m\} \tag{3.15g}$$

$$\sum_{i_2} a_{i_1,i_2} \leq N_{i_1} \;\; \forall\, i_1 \in \{1, \ldots, m\} \tag{3.15h}$$

$$\pi_{i,\text{out}}(x) \leq \pi_{\text{out,max}} \;\; \forall\, i \in \{1, \ldots, m\}, \; x \in \{1, \ldots, N_i\} \tag{3.15i}$$

$$a_{i_1,i_2}\, \delta_{\Pi(i_1),\Pi(i_2)} \leq \delta(\pi_{\text{out,max}}) \tag{3.15j}$$

$$\mathcal{P}_{i_1,i_2} \neq \emptyset \;\; \forall\, i_1, i_2 \in \{1, \ldots, m\}, \; i_1 \neq i_2 \tag{3.15k}$$

$$-\beta_i \leq 0 \;\; \forall\, i \in \{1, \ldots, m\} \tag{3.15l}$$

Constraints (3.15g) and (3.15h) constitute bounds for the degree of a backhaul node, while constraint (3.15i) limits transmission power. Coverage limitations are considered in Equation (3.15j). Finally, constraints (3.15k) and (3.15l) state requirements with respect to graph connectivity and battery levels of backhaul nodes, respectively. The set of equations (3.15) reveals the complex non-linear structure of the multi-domain QAP to be solved. Therefore, a novel modeling and evaluation framework has been implemented in a multi-domain simulation tool and been used as a guidance for this work. It outlines the approach for modeling a WBN, facilitates the refinement of the preliminary problem statement, and guides the application as well as evaluation of optimization techniques. Moreover, a representation of WBNs using concepts from graph theory has been identified as the preferred option since it allows for the flexible integration of specific models from wireless networking, as outlined in Chapter 4, and the application of well established optimization techniques. In subsequent chapters, the problem statement will be continuously detailed to finally account for, among others, energy-autarkic operation of WBNs and fairness of resource allocation.

Chapter 4

Multi-Domain Modeling of Wireless Backhaul Networks

The notion of multi-domain modeling has become popular with the advent of CPS theory. CPSs deal with the interdependence and convergence of physical, networking, and computing resources in modern engineering systems. The first systematic definition of CPSs has been contributed by [Lee08]:

> "Cyber-physical systems (CPS) are integrations of computation with physical processes. Embedded computers and networks monitor and control the physical processes, usually with feedback loops where physical processes affect computations and vice versa."

Differentiating between the *cyber* (i.e., computing and networking) and the *physical* domain, Lee argues that existing abstractions (i.e., models) of both worlds require significant improvement and, more importantly, mutual alignment as well as reconciliation in order to offer a substantial value in the development and evaluation of today's engineering systems. Further, a multi-domain perspective on system engineering is promoted in order to fully recognize and unleash the capabilities of these systems. In order to transfer this multi-domain perspective into the context of this work, the author decides to introduce three domains for modeling coordinated Wireless Backhaul Networks (WBNs) in an integrative manner: **physical**, **network**, and **context** domain [MLK+12]. The following section present state-of-the-art approaches for radio network modeling and categorize them according to the defined domains. Further, their applicability in the context of point-to-point radio networking is evaluated and necessary modifications

and extensions are realized. According to the evaluation methodology outlined in Section 3.3 , the author integrates the (partially modified) models to build a formal multi-domain representation of a WBN.

4.1 The Physical Domain of Wireless Backhaul Networks

In the context of this work, three aspects determine the impact of the physical domain on WBN performance. User mobility describes the movement patterns of subscribers and therefore has a strong impact on traffic load distributions. Further, as mentioned before, the analysis and optimization of power budgets in energy-autarkic radio networks play an extraordinary role in the scope of this work. The author presents state-of-the-art energy consumption and supply models and depicts how they are extended to properly reflect the characteristics of WBNs.

4.1.1 User Mobility

User mobility is an inherent feature in any wireless communication system. However, there is no uniform mobility behavior that can be applied for any user, system, or use case. For example, IEEE 802.11 Wireless Local Area Network (WLAN) clients usually show considerably less dynamic mobility than cellular clients. Moreover, user movements follow different patterns when considered at different spatial and temporal scales, hence requiring an according classification [LCW97]. Mobility affects important network aspects, such as handover, registration and signaling overhead, Quality of Service (QoS), resource allocation, etc. [MD97]. Mobility models usually try to represent a two-dimensional movement (velocity $\mathbf{v}$) with either Cartesian ($\mathbf{v} = (v_x, v_y)$) or polar ($(|\mathbf{v}|, \arctan\frac{v_y}{v_x})$) coordinates. Typical approaches for modeling velocity of mobile users include:

- Random walk (also referred to as Brownian motion) [CBD02],
- Gauss-Markov mobility (random walk with rather continuous 2D user movements, typically used for simulation of cellular systems) [LH03],
- interrupted random walks ("random way-point") [JM96],
- metropolitan mobility (routes are limited by street network) [MN06],
- column and pursue mobility (group mobility, many users share the same route) [SM01],

- group mobility with reference point (superposition of group movement and individual User Terminal (UT) mobility within the group) [HGPC99].

For a uniform symbolic depiction, the following notation is used throughout the next sections:

$s(t)$ – user speed (absolute value) at time t,
$\theta(t)$ – direction of user movement at time $t, 0 \leq \theta(t) < 2\pi$,
$\mathbf{v}(t)$ – (two-dimensional) velocity of UT, $\mathbf{v}(t) = \begin{pmatrix} |\mathbf{v}(t)| \cos(\theta(t)) \\ |\mathbf{v}(t)| \sin(\theta(t)) \end{pmatrix}$.

Random Walk Mobility Models

Random walk mobility describes a class of mobility models based on *Brownian Motion*, reflecting the underlying movement pattern that has its origins in the random movement of gas or liquid particles attributed to the thermal molecular movements. For each such particle, the movement is assumed to be independent from the movement of all other particles. Moreover, the movement of any particle only shows negligible auto-correlative behavior after a short amount of time. Then, according to [Ein05], particle density $\rho(p,t)$ at time t and (one-dimensional) location p is described by the following differential equation:

$$\frac{\partial \rho(p,t)}{\partial t} = D \frac{\partial^2 \rho(p,t)}{\partial p^2}, \tag{4.1}$$

where D is the so-called diffusion coefficient ($[D] = 1\ \frac{\mathrm{m}^2}{\mathrm{s}}$), indicating how fast particles of a given substance diffuse in a fluid. A solution for particle density $\rho(p,t)$ in Equation (4.1) is given by [Ein05]

$$\rho(p,t) = \frac{n}{\sqrt{4\pi D}} \frac{e^{-\frac{x^2}{4Dt}}}{\sqrt{t}}, \tag{4.2}$$

where n is the number of particles solved in the fluid. The analogy to mobile user density in wireless communication systems can easily be drawn by manipulating the diffusion coefficient D in a way that it constitutes a measure for real-world user mobility. Therefore, it is necessary to approximate D by estimating the square root of the arithmetic mean of the squares of

displacement in direction of p (or simply a user's mean displacement Δd_p (in m) in direction of p):

$$\Delta d_p = \sqrt{\overline{p^2}} = \sqrt{2Dt}. \tag{4.3}$$

Consequently, given an initial user density $\rho_0 = \rho(p, t_0)$, the user density for any location p at any time t and, derived from that, user location can be computed. With according extensions for two-dimensional particle density problems, user speed $s(t)$ and direction $\theta(t)$ at time t are computed by

$$s(t) = \overline{s} + s_R, \tag{4.4}$$
$$\theta(t) = \overline{\theta} + \theta_R, \tag{4.5}$$

where

$\overline{s}, \overline{\theta}$ – mean values of speed and angle of direction, respectively,
s_R, θ_R – speed and angle realizations of a specified probability distribution.

Gauss-Markov Mobility Model

The Gauss-Markov mobility model is an evolution of the random walk model that addresses the (frequently unrealistic) assumptions of complete randomness, lack of memory, and consequently, lack of auto-correlation of the random walk model [LH03]. In order to overcome these drawbacks, user mobility is modeled as a Gauss-Markov process $\Theta(t)$. A stationary Gaussian process is of Gauss-Markovian nature if its auto-correlation function $\Phi_\Theta(\tau)$ obeys Equation (4.6) [Gel74]:

$$\Phi_\Theta(\tau) = \mathrm{E}[\Theta(t)\,\Theta(t+\tau)] = \lambda\sigma^2 + \mu^2, \tag{4.6}$$

where

σ – standard deviation of $\Theta(t)$,
μ – mean of $\Theta(t)$,
$\lambda = e^{-\alpha|\tau|}$ – degree of memory ($\alpha \geq 0, 0 \leq \lambda \leq 1$).

In practical terms, the mobility model facilitates the application of customized degrees of randomness by introducing the tunable memory pa-

rameter λ. In case of complete memory ($\lambda = 1$), the process $\{\mathbf{v}(t)\}$ is equivalent to (deterministic) linear movement. In contrast, setting $\lambda = 0$, $\{\mathbf{v}(t)\}$ becomes a purely random walk as defined in Section 4.1.1, i.e., a completely memoryless process. Moreover, the usage of a simple Finite State Machine (FSM) with memory span of 1 for both speed s and direction θ helps implementing the correlation of current velocity with that from the previous time step. The following recursive equations realize this behavior:

$$s(t) = \lambda s(t-1) + (1-\lambda)\mu_s + \sqrt{1-\lambda^2}\,\sigma_s\,\varsigma_{s,t}, \tag{4.7}$$

$$\theta(t) = \lambda\theta(t-1) + (1-\lambda)\mu_\theta + \sqrt{1-\lambda^2}\sigma_\theta\varsigma_{\theta,t}, \tag{4.8}$$

where

σ_s, σ_θ – standard deviation of s and θ, respectively,
μ_s, μ_θ – mean of s and θ, respectively,
ς_t – realization of a random variable with specified probability distribution.

Initial speed s_0 and direction θ_0 are drawn from Gaussian distributions $\mathcal{N}(\mu_s, \sigma_s)$ and $\mathcal{N}(\mu_\theta, \sigma_\theta)$, respectively. Designed in this manner, $s(t)$ and $\theta(t)$ obviously satisfy Equation (4.6), making them a Gauss-Markov processes. In the depicted approach, the degree of memory λ is chosen to be the same for both speed and direction. Operating with two different memory degrees for speed and direction (λ_s and λ_θ) results in a completely independent modeling of the two dimensions, at the price of additional computational complexity. The general advantage of the Gauss-Markov mobility model originates from its flexibility of continuously switching between (completely) random and (completely) deterministic movement behaviors, e.g., by making $\lambda = \lambda(t)$ a function of t. The memory characteristics of the model result in a further advantage: User movement becomes smoother by reducing the likelihood of sudden directional changes or accelerations. Obviously, this can be achieved by increasing the memory level beyond one (higher order Markov models).

Example 4.1

A common variation of Gauss-Markov mobility is utilized in the scope of this thesis. Two-dimensional user velocity $\mathbf{v}$ *is expressed in terms of speed* $s = \|\mathbf{v}\|_2$ *and angle* $\theta = \arctan\frac{v_y}{v_x}$. *Based on previous values* $s(t-1)$

and $\theta(t-1)$, velocity updates are derived as follows:

$$s(t) = s(t-1) + 0.8 s_m \varsigma_{s,t} \tag{4.9}$$

and

$$\theta(t) = \theta(t-1) + \frac{\pi}{2} \mathrm{e}^{-\frac{s(t-1)}{s_m}} \varsigma_{\theta,t}, \tag{4.10}$$

where
s_m – mean speed across all users,
$\varsigma_{\cdot,t}$ – realizations of a random variable of uniform distribution $U[-1,1]$.

In other words, Δs is uniformly distributed in the interval $[-0.8 s_m, 0.8 s_m]$, whereas $\Delta\theta$ follows a uniform distribution over an interval I whose width is controlled by the term $\mathrm{e}^{-\frac{s(t-1)}{s_m}}$, i.e., the faster a user moves, the smaller I and thus the smaller the maximum possible change of θ. I exhibits maximum width π if a user does not move: $I(s=0) = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Figure 4.1 depicts the movements of three exemplary users with medium ($\overline{s}_1 = s_m$), low ($\overline{s}_2 = \frac{1}{2} s_m$), and high ($\overline{s}_3 = 2 s_m$) average speed $\overline{s}$ in a rectangular area (that can be bent into a torus) during a time span of 15 minutes. While directional changes of the slow user occur frequently and occasionally even abruptly, significantly fewer and smoother turns can be observed for the fast moving user. Moreover, the difference in speed is also visible with respect to the traveled distance. The torus-like wrapping of the considered area guarantees its closure, i.e., users do not leave the area, rather, they reappear at a defined location.

Random Waypoint Mobility

Being an evolution of the random walk, random waypoint mobility shares many characteristics with models of the same origins. The basic idea is a repeated alternation between the *moving* and the *resting* state. At initialization, destination coordinates and speed are drawn randomly. The terminal then moves towards the destination at the according (constant) speed, either on a straight line or on allowed paths (e.g., a road network). Upon arrival, a random variable determines the amount of time the terminal pauses until a new random destination and speed are drawn. With these new figures, the procedure starts all over again. Figure 4.2 schematically depicts the sequential steps of the model.

The choice of probability distributions for the random variables is a crit-

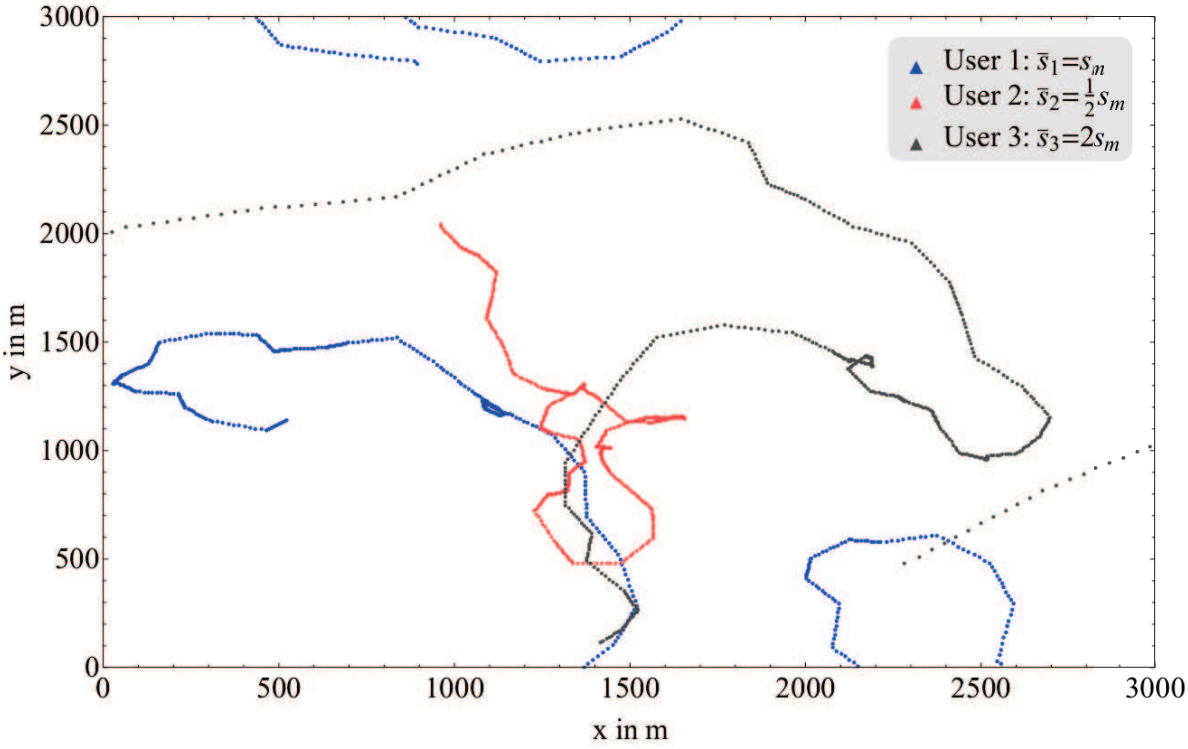


Figure 4.1: Realizations of Gauss-Markov mobility

ical design issue in this model, i.e., results have to be evaluated carefully with regard to the parameters of the probability distributions. Particular focus should be given to the ratio of the mean time a user spends in the two states. Finally, the model is also comparatively sensitive to initialization. The ability to realistically model user mobility and flexibility are among the advantages of the widely used model, e.g., for routing protocol performance evaluation [BMJ+98, GLAS99, JLH+99].

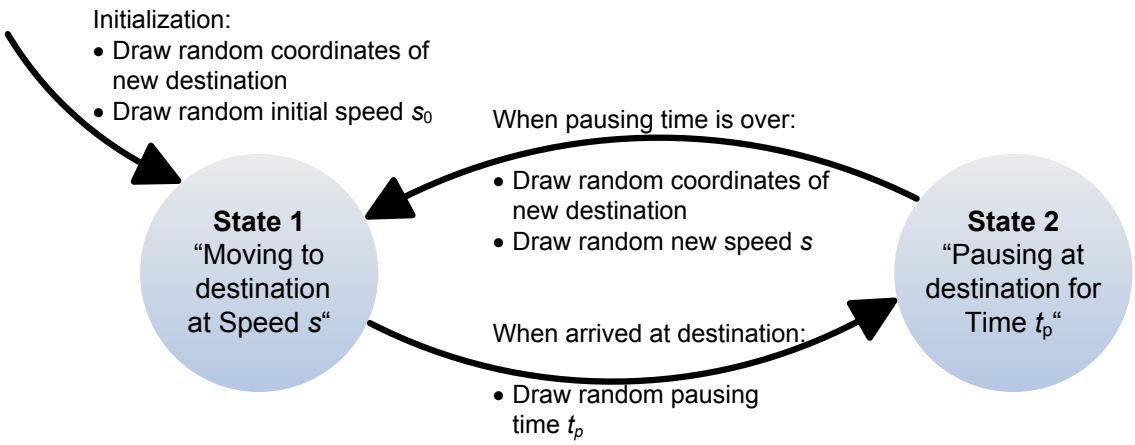


Figure 4.2: State machine for random waypoint mobility

Metropolitan Mobility Models

Metropolitan mobility depicts a class of models that take any kind of maps, typically a road network, as input. The users' mobility is limited to selected areas, such as roads or pedestrian zones (possibly including additional traffic regulation information, e.g., speed limits, traffic lights, etc.). Other areas on the map usually cannot be reached by the user [MN06]. Drawbacks of this significant reduction of possible user locations are very specific user distributions and, as a consequence, less generic and comparable results. However, map-based models reflect real-world user trajectories exceptionally well, thus generating valid results for the specific city layout under consideration, particularly for densely populated urban areas. Operational details of the model are similar to the random waypoint model as described in the previous paragraph. For initialization, all users are put in a random location of the allowed map area. After randomly selecting a respective destination, the shortest path (w.r.t. to a selected parameter, usually time or distance) is calculated and the user starts traveling along the computed route. Upon arrival (and after pausing time has passed), a new destination is selected and the process starts all over again.

Group Mobility Models

The models depicted so far assume mutual independence between individual user movements. In many scenarios, this has proven to be an unrealistic assumption since user movements are correlated more often than not, e.g., in public transportation or vehicular scenarios. This inconsistency is resolved by so-called group mobility models that contain mechanisms for a correlated user mobility behavior:

- **Reference Point Group Mobility** [HGPC99] - In this model, individual user movement mainly depends on the trajectory of the logical group center moving according to a selected model. Users belonging to the group are assigned a fix reference point (position relative to the group center). In each time step, the spatial deviation of the user from this reference point is calculated using a random variable. Hence, every user stays within a maximum distance from the group center.
- **Column Mobility** [San01] - This model defines equidistant reference points on a straight line, one point for each user. The line as

a whole encounters a random two-dimensional shift over time, hence keeping the reference points in a constant relative position. In parallel, users are allowed to oscillate around their respective reference point within a certain radius in a random manner, thus breaking the strictly straight line pattern.

- **Pursue Mobility** [SM01] - The pursue mobility model takes the movement of a chosen user as reference for all other users. For each time step, a user's displacement from his previous position is calculated using the movement vector of the reference user and adding an individual, time-variant white noise vector of limited magnitude to preserve the pursue nature.

Further models reflect other typical characteristics of moving groups (such as the *Nomadic Community* model where users share a common reference point around which they move independently) or simply try to mathematically create a correlation (such as the *Exponential Correlated Random Mobility* model [HGPC99]).

4.1.2 Energy Supply

Energy supply plays a crucial role in the operation of autarkic WBNs since they are required to be self-sustaining with respect to their fluctuating balance of power supply and consumption. This particular requirement is paramount in deployment areas where no or only insufficiently reliable energy grids exist. Typically, this is the case in remote areas (*Alps* scenario) or developing countries (*Sub-Saharan Africa* scenario). Consequently, alternative sources of energy have to be made available for reliable network operation. Two reasonable options are power generation from wind using wind turbines or from sun light using PV modules. This work will focus on the latter option: it has nowadays become the preferred solution in practice, usually requires less deployment and maintenance effort, and provides the best fit with the considered scenarios (in particular the *Sub-Saharan Africa* scenario).

PhotoVoltaic (PV) modules can be a viable alternative when planning for autarkic power supply, e.g., for wireless mesh routers. However, before installation, the economic and technical soundness of a deployment has to be analyzed. This usually requires the computation of a set of criteria, among others, installation and operation cost as well as yield of PV modules. This section describes the major parameters the latter depends upon,

how it can be modeled (e.g., for simulation purpose) and what metrics are used in practice. A PV module consists of a series connection of multiple so-called solar cells. A solar cell is a semiconductor element (usually silicon) exploiting the internal photoelectric effect [Qua06]. Here, light of sufficient energy (i.e., frequency) can lift electrons from the valence band to the conduction band, thus activating the intrinsic conduction of semiconductors. Crystalline silicon possesses the advantage of high spectral sensitivity across a broad frequency band. In other words, a wide range of wavelengths can be exploited for transforming irradiation into electrical power. Modern high volume PV concentrators can convert up to 27% of irradiation into electrical power [Qua06]. In laboratory setups, researchers have demonstrated efficiency ratios of almost 45% [Dim13]. Meteorological and geo-information services, e.g., *German Weather Service (Deutscher Wetterdienst) (DWD)* [Deu12b], PVGIS [Joi12], or SolarGIS [Geo14], offer maps and tables indicating mean global irradiation for the last decades. A world map of global horizontal irradiation is illustrated in Figure 4.3. For Germany, global irradiation values as measured on a horizontal plane are available with a 1 km spatial resolution. Tilt and orientation of a PV panel, shading, cloudy skies, etc. can increase or decrease this value. Using the so-called "performance ratio" PR and peak power P_p, the monthly or yearly yield of a PV module can be computed, usually assuming non-shaded modules with free horizon [Deu12a]. Mean yearly yield Y (in kWh) is calculated

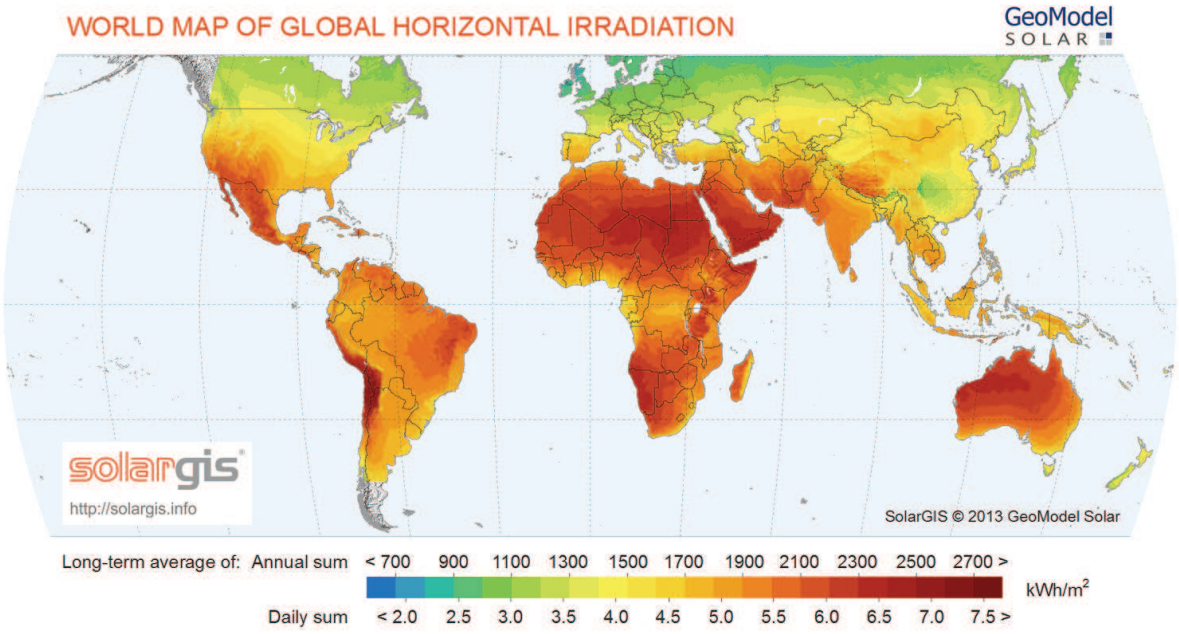


Figure 4.3: World map of global horizontal irradiation [Geo14]

according to [HB95]

$$Y = \eta\, A\, H = \eta A \int_{t_1}^{t_2} I(t)dt, \tag{4.11}$$

where

η – efficiency of PV module,
A – surface area of the solar panel(s),
H – sum of global irradiation at PV module location, $[H] = 1\ \frac{\text{kWh}}{\text{m}^2}$,
$I(t)$ – instantaneous global irradiation at module location, $[I] = 1\ \frac{\text{W}}{\text{m}^2}$.

H, besides including *direct* solar irradiation, also accounts for *diffuse* irradiation coming from atmospheric dispersion and reflections from objects or the ground [Rit06]. On overcast days, diffuse irradiation accounts for almost 100% of total irradiation. The efficiency η of a PV module nowadays typically lies in the range of 0.09 to 0.11 [Laq03].

Since solar panel producers usually do not publish η but rather peak (also called effective) power output P_{stc}, i.e., *power output under standard test conditions (stc)*, assuming a standardized irradiation power of $G_{\text{stc}} = 1\ \frac{\text{kW}}{\text{m}^2}$ [Qua06], Equation (4.11) can be rewritten as

$$Y = P_{\text{stc}}\, \frac{H}{G_{\text{stc}}} \tag{4.12}$$

by exploiting $\eta \times A = \frac{P_{\text{stc}}}{G_{\text{stc}}}$. Obviously, panel surface area is now included in P_{stc}. However, Equation 4.12 does not account for tilt and orientation (Azimuth) of the solar panel, as global irradiation H refers to a horizontal plane (whereas P_{stc} already assumes a vertical solar irradiation on the panel, i.e., a tracked panel). In order to incorporate panel tilt θ and orientation φ, a correction factor $c(\theta, \varphi)$ is introduced [HB95]:

$$Y_{\text{ref}} = P_{\text{stc}}\, \frac{H}{G_{\text{stc}}}\, c(\theta, \varphi), \tag{4.13}$$

where Y_{ref} is the so-called reference yield in kWh. The actual value of $c(\theta, \varphi)$ depends on the season and is usually derived experimentally. Table 4.1 depicts typical correction factors for orientation and tilt during different seasons of the year.

Table 4.1: Correction factor $c(\theta, \varphi)$ for global irradiation [Joi12]

Module Direction		**Season**			
Orientation φ	**Tilt θ**	**Spring**	**Summer**	**Fall**	**Winter**
South	30°	1.08	0.99	1.31	1.42
	45°	1.06	0.94	1.39	1.54
	60°	0.99	0.85	1.4	1.59
South-East South-West	30°	1.07	0.99	1.27	1.35
	45°	1.05	0.94	1.33	1.45
	60°	1.0	0.86	1.32	1.48
East West	30°	0.98	0.96	1.0	1.01
	45°	0.95	0.93	0.98	0.99
	60°	0.89	0.87	0.94	0.95

The actual yield, also termed final yield Y_f, of a PV module further depends on additional parameters, such as shading, snow coverage, surface dust layer, generator losses, efficiency losses in case of weak irradiation, reflection, cable losses, inverter losses, etc. [Qua06], which can be categorized into capture losses L_c and system losses L_s. The according decrease in yield is indicated by [HB95]

$$Y_\mathrm{f} = Y_\mathrm{ref} - P_\mathrm{stc}\,(L_c + L_s), \tag{4.14}$$

where

$$\begin{aligned} Y_\mathrm{f} &- \text{final yearly yield, } [Y_f] = 1 \text{ kWh},\\ L_c &- \text{yearly capture losses, } [L_c] = 1 \text{ h},\\ L_s &- \text{yearly system losses, } [L_s] = 1 \text{ h}. \end{aligned}$$

From numerous research projects, it has been shown that the major percentage of losses can be attributed to shaded solar panels (i.e., to L_c). The important non-dimensional quantity of performance ratio PR can now be computed with the following quotient:

$$PR = \frac{Y_\mathrm{f}}{Y_\mathrm{ref}} = 1 - \frac{G_\mathrm{stc}\,(L_c + L_s)}{H\,c(\theta, \phi)}. \tag{4.15}$$

With Equation (4.15), Equation (4.14) can be rewritten to the common form of

$$Y_\mathrm{f} = PR\,Y_\mathrm{ref} = PR\,P_\mathrm{stc}\,\frac{H}{G_\mathrm{stc}}\,c(\theta, \phi). \tag{4.16}$$

For simulation purpose, yearly yield, as considered so far, has to be broken down to daily or hourly distributions of irradiation power, thus account-

ing for weather conditions (e.g., clouds), air dust, and other sources of disturbance of optimal yield on a smaller time scale. Yield fluctuations as expected during cloudy days can be incorporated in a random fashion. The approach is illustrated in Example 4.2.

Example 4.2

In this example, the author illustrates the model-based computation of power yield of a PV module (e.g., for simulation purpose). The following input parameters are required to calculate solar irradiation [HB14]:

- *Sun declination angle* Γ, $\Gamma(t_y) = \Gamma_{max} \sin\left(\frac{360}{365}(t_y - 81)\right)$, $t_y = 1, \ldots, 365$ *(day of the year),* $\Gamma_{max} = 23.45°$,
- *hour angle* ι, $\iota = 15° \, (t_d - 12)$, $t_d \in [0, 24[$ *(hour of the day), the earth rotates* $15°$ *per hour,*
- *panel size* A,
- *efficiency* η *of PV module,*
- *geographical latitude* Λ, *e.g.,* $\Lambda = 47.8°$ *(approximate geographical latitude of the Alps),*
- *solar constant* $H_0 = 1,353 \, \frac{W}{m^2}$,
- *percentage* h *of irradiation incident on the outer atmosphere perpendicularly arriving on the ground.*

From these parameters, daily power yield Y_d *of a PV module as well as sum of solar irradiation* H *in* $\frac{kWh}{m^2}$ *during a sunny day can be computed according to*

$$\begin{aligned} Y_d &= \eta \, A \, H \\ &= \eta \, A \, H_0 \, h^{\cos^{-0.678}\left(\frac{\pi}{2} - \arcsin(\cos(\iota) \, \cos\Gamma \, \cos\Lambda + \sin\Gamma \, \sin\Lambda)\right)}. \end{aligned} \tag{4.17}$$

The exponent to h *is used to approximate field measurements [HB14] and scales down* Y_d *by accounting for two effects. First, irradiation power is reduced by non-perpendicular propagation of light through the atmosphere due to the geographical latitude of the considered location, current sun declination, and time of day. Second, nonuniformities in atmospheric layers further attenuate light intensity. In the described approach, temporal variations of solar irradiation are captured on a seasonal scale by* Γ *and on a daily scale by* ι. *Spatial characteristics are reflected by* Λ. *Finally, weather conditions are considered by the clearness index* k. *The clearness index describes the ratio of mean hourly sum of terrestrial and extraterrestrial irradiation on a horizontal plane. Figure 4.4 illustrates a sample*

probability distribution $p(k)$. The so-called three-state model [HS13] captures this empirical shape by parametrizing three (truncated) normal distributions $\mathcal{N}_i(\mu_i, \sigma_i)$, $i \in \{$"sunny", "partially cloudy", "overcast"$\}$, using (location-specific) mean and standard deviation (μ_i, σ_i), each reflecting one of the three basic weather conditions i. A weighted linear combination superimposes the three distributions, yielding $p(k)$:

$$p(k) = \sum_i w_i \, p_i(k), \quad p_i(k) \sim \mathcal{N}_i, \, \sum_i w_i = 1. \tag{4.18}$$

Figure 4.5 depicts example irradiation patterns (in $\frac{W}{m^2}$) during different seasons under changing cloudiness conditions as derived from the presented method [VH98, KDE07]. Integration of these curves gives a day's irradiation sum in $\frac{kWh}{m^2}$. Besides varying widths of the Gaussian bell curves due to seasonal changes (time of sun rise and set, maximum sun declination), the figures show the (very irregular) ups and downs of instantaneous irradiation due to cloudy skies. For computing PV yield in the course of a particular day, empirical data for the clearness index serves as the basis for fitting the location-specific three-state model.

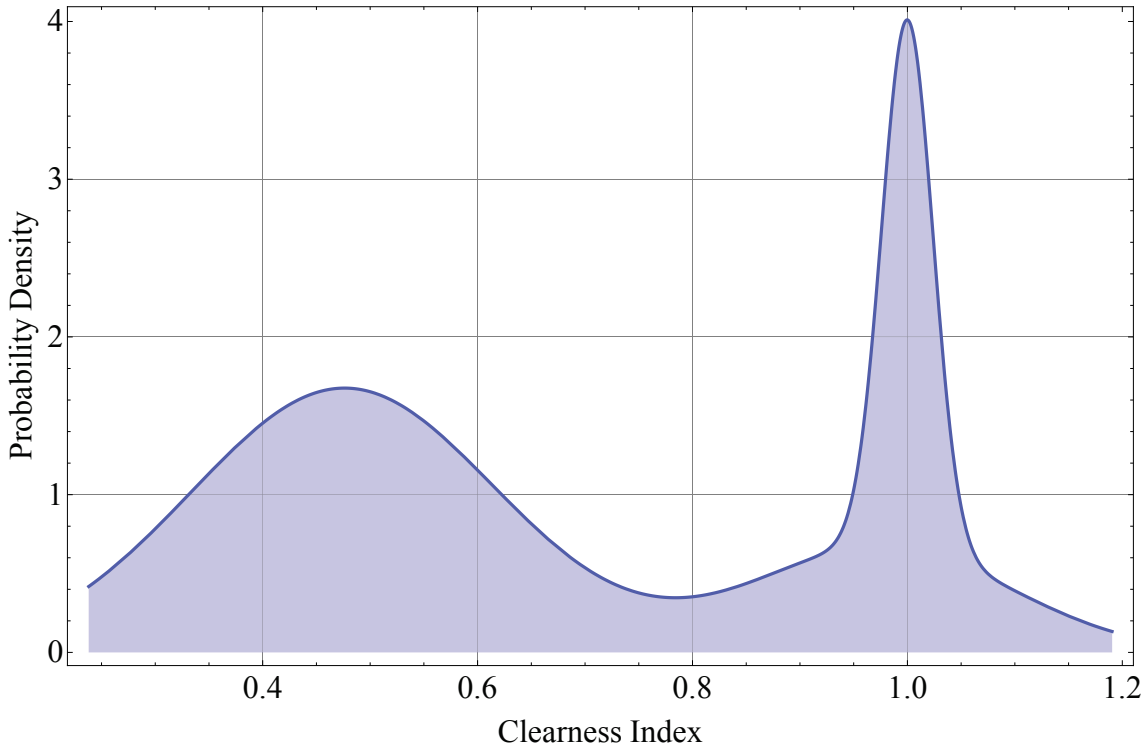


Figure 4.4: Probability distribution $p(k)$ of clearness index k [MCS14]

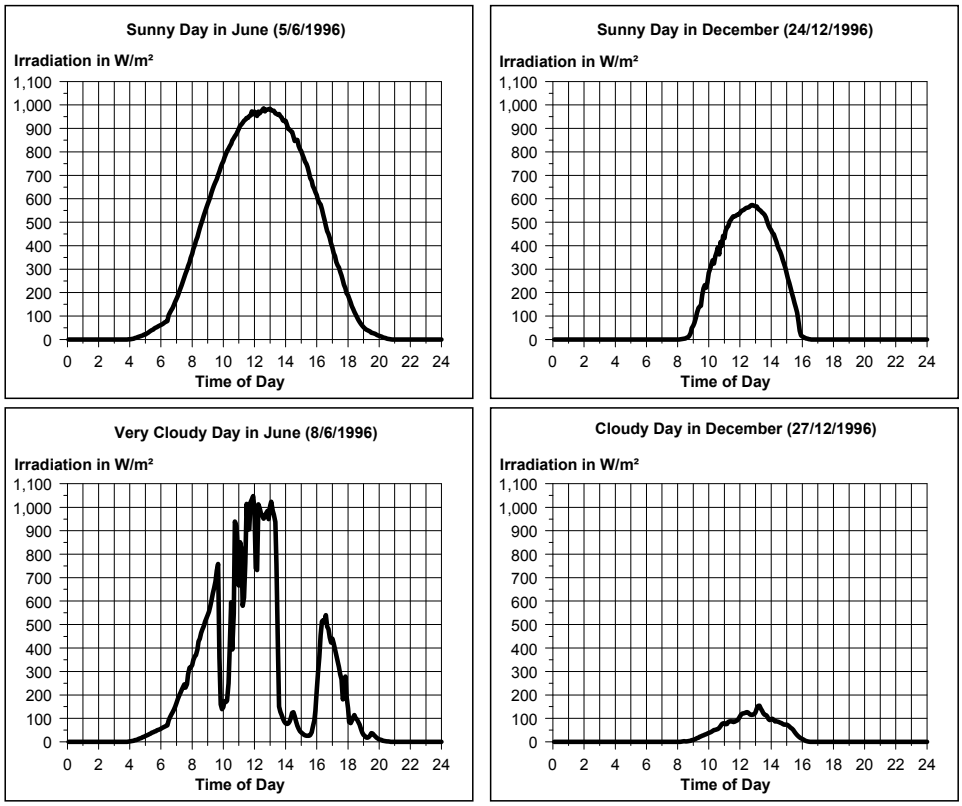


Figure 4.5: Intraday variation of global irradiation [VH98]

4.1.3 Energy Consumption

Energy consumption of wireless network infrastructures has become a significant contributor to operational costs for network operators. The significant improvements in data rates provided by mobile networks as well as the increased variety of services available have come at the cost of increased power consumption of access, backhaul, and backbone network components, such as power amplifiers and high-performance processors. Consequently, reducing power consumption meanwhile enjoys a high priority among network operators and network optimization is carried out with two equally important objectives: QoS maximization (in particular, maximization of throughput and minimization of end-to-end latency) and minimization of energy consumption [JSK+13]. However, due to the multi-domain nature of wireless backhaul networking, controlling energy consumption is a very complex process depending on heterogeneous variables, including deployment aspects, traffic patterns, and user mobility. Moreover, variations in these parameters occur on both small- and large-scale spatial and temporal dimensions. Respective models have to reflect this heterogeneity whenever required. Generally, energy consumption models can be classified into three major categories:

1. Continuous models that derive energy consumption as a function of a correlated quantity, such as data throughput, output power of amplifier units, system utilization,
2. discrete models that compute energy consumption depending on the state of a network component (e.g., radio interface states: off, sleep, idle, transmitting, receiving),
3. static models that approximate energy consumption by a single correlated, though static quantity (e.g., number of users, number of deployed base stations).

The following sections present a selection of state-of-the-art models for power consumption models wireless networks.

Component-Level and System-Level Energy Consumption

Component-level energy consumption models focus on describing energy consumption of isolated network components, e.g., user terminals (UTs) or base transceiver stations (BTSs). The according results are aggregated in a later stage to derive the power demand of the overall network. In contrast, system-level energy consumption is slightly more challenging to approximate since contributing components are numerous, very heterogeneous, and dependent on system configuration. Therefore, approaches for computing power consumption of an entire network are usually based on one or several reference quantities, such as total traffic volume, number of cells, system utilization, etc. Power consumption per base unit of such reference quantities, according multiplications, and totaling yield overall power demand of a network.

State-Dependent Energy Consumption of Components

Designing very accurate, continuous-domain, closed-form models for energy consumption in wireless networks (e.g., based on data traffic or output power at antennas) is a challenging task. Moreover, in many cases, the additional efforts for model development are not fully compensated for by an according increase in accuracy. Moreover, many such models, due to their inherent overhead, are ill-suited for implementation and evaluation purposes. Therefore, a common approach for modeling energy consumption in wireless communication systems employs discrete states or modes a node (i.e., a UT or a BTS) can be in, assigning a fix power consumption to each of

them. This offers the flexibility of introducing as many states as required for a sufficiently accurate representation of reality. Moreover, engineers can benefit from these rather simple models by easily integrating them into real systems or evaluation frameworks. In the context of this work, it is worthwhile to exploit these advantages and create a state-dependent power consumption model. At the level of the backhaul network nodes, the following states have been identified:

- Node on/off,
- in case of "node on", for each radio interface of the node,
  - interface down,
  - interface up and in idle state,
  - interface up and receiving (Rx state),
  - interface up and transmitting (Tx state).

Example 4.3

For illustration purpose, the author considers a wireless backhaul node with two IEEE WLAN radio interfaces. Such a configuration yields eleven possible states. The following list depicts a sequence of operational modes and the resulting power consumption [Qua04]:

– *Node off*	0 *W*,
– *node on, radios down*	12 *W*,
– *first radio up (idle)*	21.13 *W*,
– *first radio up and receiving*	23.2 *W*,
– *first radio up and transmitting*	23.4 *W*,
– *additionally, second radio up (idle)*	32.53 *W*,
– *second radio up and receiving*	34.6 *W*,
– *second radio up and transmitting*	34.8 *W*.

The maximum total power consumption of a two radio node adds up to 34.8 W and scales accordingly with an increasing number of radio interfaces. In the context of this thesis, the number of radio interfaces of a node is limited to four, which amounts to a maximum power consumption of 57.6 W per backhaul node. Figure 4.6 illustrates the eight consecutive consumption levels as listed above, from state node off *at 0 W to state* two radios transmitting *at approximately 35 W, as well as the additional power required for each state transition.*

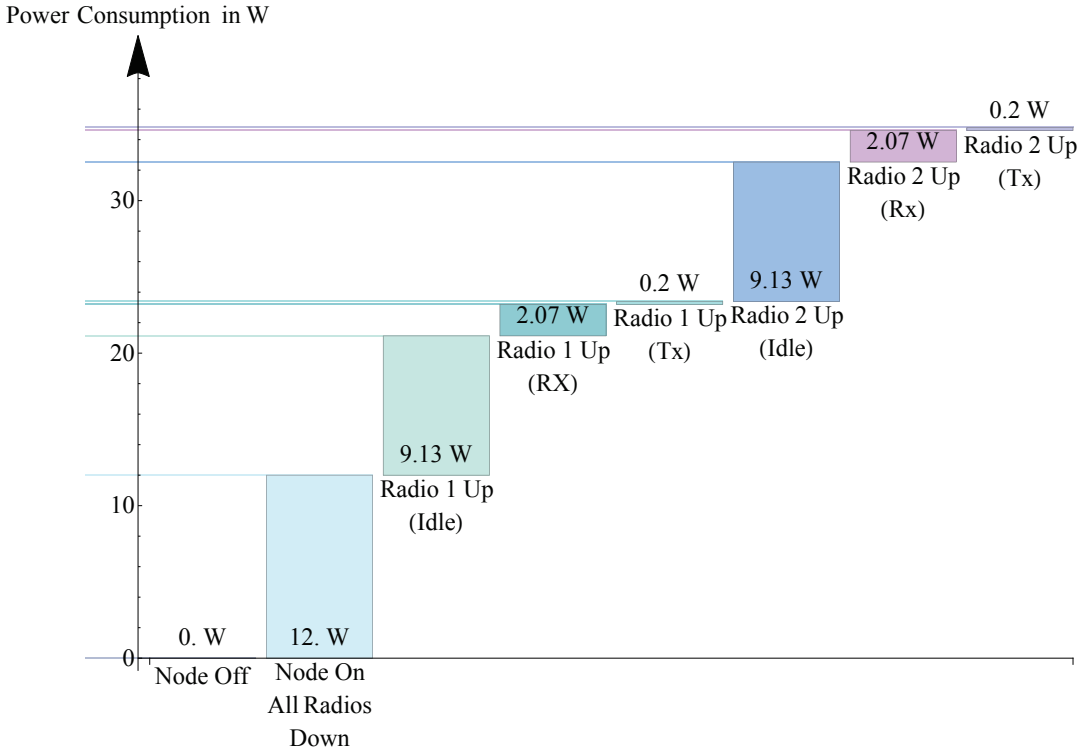


Figure 4.6: State-dependent energy consumption of a network node [Qua04]

Power Consumption of a Base Transceiver Station

BTS power consumption models concentrate on capturing the power consumption of BTSs, thus being less generic since models are tailored according to the characteristics of a single network component. The rationale is that, from an operator's perspective, BTSs are the major consumers of energy, any other components can be considered negligible.

Typically, BTS power consumption models assume a linear relationship between Radio Frequency (RF) power output π_{out} and total power consumption π_{in}, which can be attributed to the fact that the Power Amplifier (PA), besides being the major power consumer, is the only subcomponent that scales according to varying traffic load [AGG+11]. Other subcomponents, such as base band processors and coolers, have a rather constant power consumption independent of instantaneous traffic load. Therefore, π_{in} is modeled with a base consumption π_0 that grows commensurate to output power π_{out} [ABG+10].

$$\pi_{\text{in}} = N_i \left(\pi_0 + m_\pi \, \pi_{\text{out}}\right), \quad 0 \leq \pi_{\text{out}} \leq \pi_{\text{out,max}}, \tag{4.19}$$

where

N_i – number of radio interfaces (transceiver chains) of BTS i,
π_0 – power consumption without any traffic load,
$\pi_{\text{out,max}}$ – output power at full traffic load,
m_π – incremental power consumption per additional watt of π_{out}.

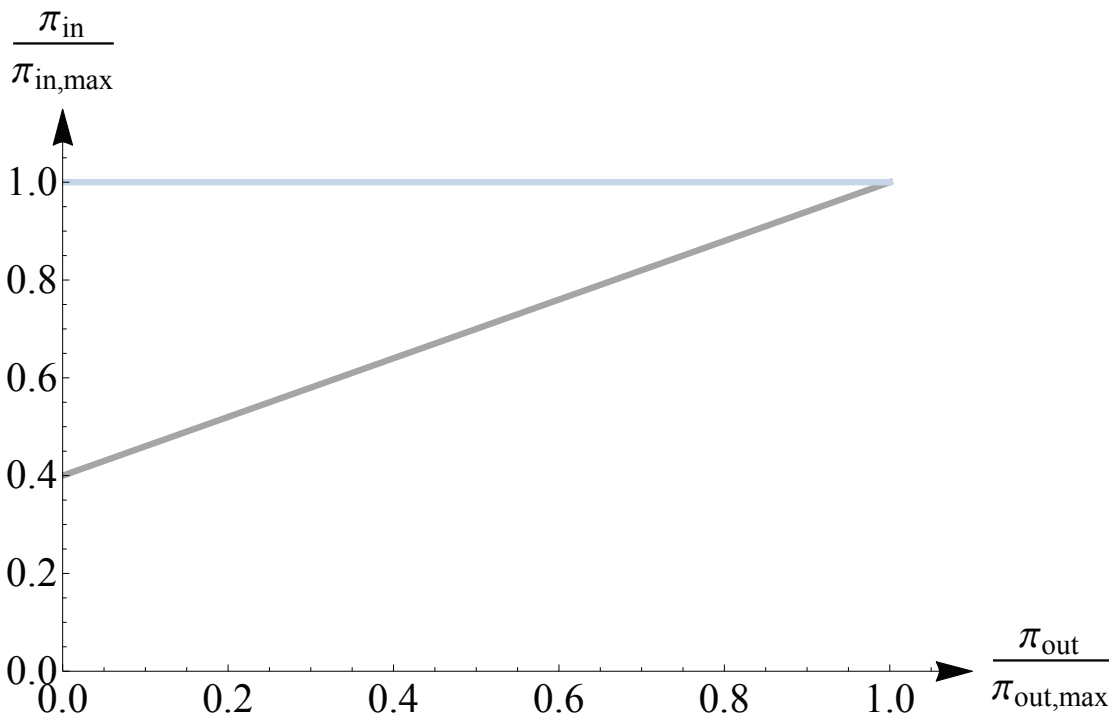
$\pi_{\text{out,max}}$ and π_0 are derived depending on the power consumption of individual subcomponents in the respective load situations, e.g., in [AGG+11], this includes the antenna interface, power amplifier, small signal RF transceiver, baseband engine, power supply, and cooling. Figure 4.7(a) shows an example of this model with $\pi_0 = 0.4\,\pi_{\text{in,max}}$.

Energy Consumption Based on System Throughput

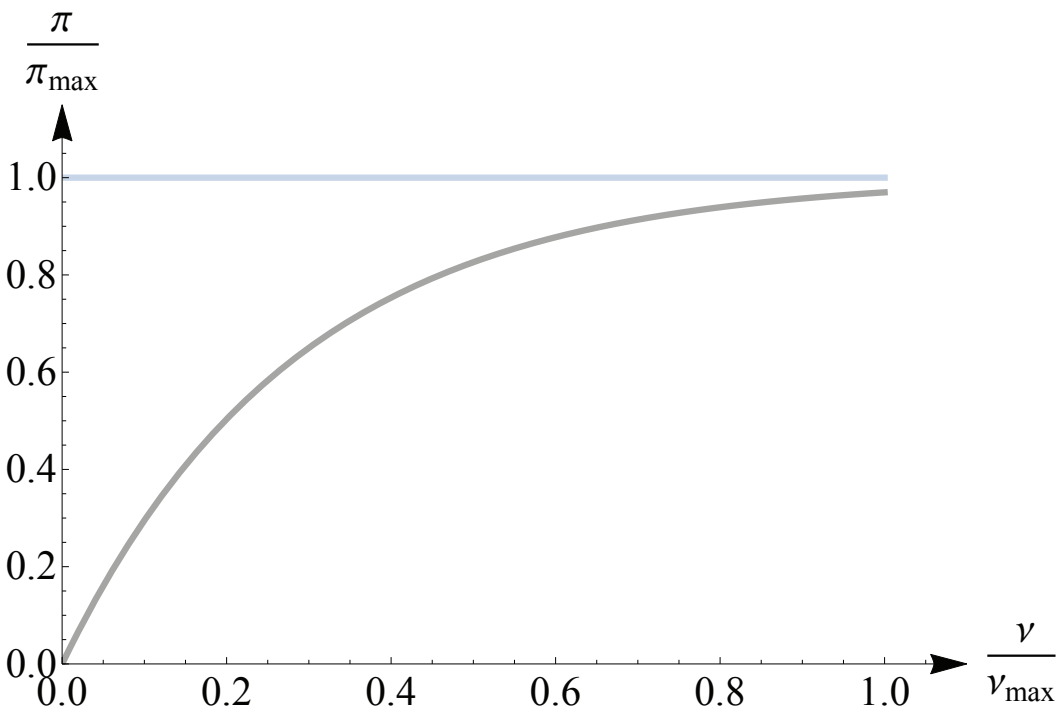
A common approach for estimating power consumption on system level is based on total throughput. Thus, the amount of data transferred in a certain time window as well as a quantity that states power consumption per base unit (e.g., $\frac{\text{W}}{\text{bit/s}}$ or $\frac{\text{W}}{\text{MB}}$) need to be computed. For deriving total data volume, a further breakdown is required. A common approach is to approximate data throughput of an average base station (or backhaul node), which is then multiplied with the total number of base stations. However, further assumptions need to be taken into account, among them average load of a base station (in terms of connected users or radio resource utilization) and data traffic incurred by an average user or per radio resource unit. The degree of granularity in calculating these statistics can vary significantly and has a strong impact on both accuracy of the figures and effort of deriving them. Auer et al. [ABG+10] depict a degressive function for estimating power consumption π based on overall throughput ν:

$$\pi(\nu) = \pi_{\max}(1 - \mathrm{e}^{-\alpha\nu}). \tag{4.20}$$

Figure 4.7(b) illustrates Equation 4.20. Obviously, the function can be interpreted as total power consumption of multiple individual base stations (as shown in Figure 4.7(a)) that are switched on successively with increasing overall traffic load, e.g., as part of a BTS densification strategy.



(a) BTS level



(b) System level

Figure 4.7: Power consumption models [AGG+11]

4.2 The Network Domain of Wireless Backhaul Networks

In this section, the author illustrates the major concepts for modeling the network domain of a Wireless Backhaul Network (WBN), including necessary adaptations to these concepts. This comprises spatial models for slow (large-scale) fading, Signal-to-Interference-plus-Noise Ratio (SINR), and link capacity for both access and backhaul networks. Further, the author presents models for voice and data traffic as well as routing protocols that are suited for realistically reflecting WBN operation.

4.2.1 Path Loss and Capacity of Radio Links

Backhaul Network

Path loss for communication over the backhaul links is modeled according to the commonly used spatial channel model, e.g., from [EG04, KMH+07]. The IEEE 802.11n-based links of the backhaul network operate in the 5 GHz band with sufficient channel separation among neighboring nodes and no interference resulting from IEEE 802.11n 2.4 GHz access links. Point-to-point backhaul links are assumed to obey the following path loss L_p law [TV05]:

$$L_p = -20 \log_{10} \frac{\lambda}{4\pi d_0} + 10\alpha \log_{10} \frac{d}{d_0}. \tag{4.21}$$

Signal attenuation is proportional to d^{-2} for all distances $d \leq d_0$ (free space loss). For $d > d_0$, the path loss coefficient becomes α. The resulting signal-to-noise ratio $\mathrm{SNR}_{\mathrm{Rx,backhaul}}$ at the receiving node is derived by a link budget calculation as follows [Mol11]:

$$\mathrm{SNR}_{\mathrm{Rx,backhaul}} = P_{\mathrm{Tx}} - L_{\mathrm{Tx}} + G_{\mathrm{Tx}} - L_p - M + G_{\mathrm{Rx}} - N_{\mathrm{Rx}} - N_{\mathrm{th}}, \tag{4.22}$$

where

P_{Tx} – transmission power (amplifier output),
L_{Tx} – losses at transmitter (cables, combiners, etc.),
G_{Tx} – antenna gain at transmitter,
M – fading margin,

G_{Rx} – antenna gain at receiver,
N_{Rx} – receiver noise,
N_{th} – thermal noise at receiver.

For simplification, interference is not considered on the backhaul link level for the following reasons: channels in the 5 GHz band are well separated from each other, frequency reuse is low, and interference suppression mechanisms of the Medium Access Control (MAC) layer are available [LYGL05].

Depending on the experienced SNR, different Modulation and Coding Schemes (MCSs) are selected on the Transmitter (Tx) side, with modulation schemes ranging from Binary Phase-Shift Keying (BPSK) to 64-Quadrature Amplitude Modulation (QAM) and code rates ranging from $\frac{1}{2}$ to $\frac{5}{6}$. The mapping of SNR to gross channel capacity C_{backhaul} (in Mbit/s) is performed according to [IEE09], assuming a channel bandwidth of 20 MHz and Multiple Input and Multiple Output (MIMO) operation with $N_{\mathrm{ss}} = 2$ spatial streams (as depicted in Table 4.2):

$$C_{\mathrm{backhaul}} = \sum_{i=1}^{N_{\mathrm{MCS}}} C_i \, u(\mathrm{SNR}_{\mathrm{Rx,backhaul}} - \mathrm{SNR}_{i,\mathrm{MCS}}), \tag{4.23}$$

where

$u(\cdot)$ – unit step function,
N_{MCS} – number of MCSs,
C_i – additional capacity of MCS i in comparison to MCS $i-1$,
$\mathrm{SNR}_{i,\mathrm{MCS}}$ – required SNR for MCS i.

Access Network

In the access domain, modeling SINR as well as capacity becomes more complex due to medium contention in case several UTs are connected to the same Access Point (AP). Moreover, due to the reduced amount of orthogonal channels in the utilized 2.4 GHz band, interference from other stations (i.e., UTs and APs) has to be accounted for. Therefore, the following quantities are considered for modeling the capacity of access links:

- Number of users served by AP i,
- medium contention due to multiple UTs and random access nature of MAC layer mechanism (Carrier Sense Multiple Access/Collision

Table 4.2: MCS parameters for IEEE 802.11n WLAN 20 MHz channel bandwidth with $N_{\text{ss}} = 2$ spatial streams [IEE09]

Modulation	**Coding rate**	$\mathbf{N_{CBPSC}}$	$\mathbf{N_{CBPS}}$	$\mathbf{N_{DBPS}}$	**Channel capacity**	**Minimum SNR**
BPSK	$\frac{1}{2}$	1	104	52	13.0 Mbit/s	14 dB
QPSK	$\frac{1}{2}$	2	208	104	26.0 Mbit/s	17 dB
QPSK	$\frac{3}{4}$	2	208	156	39.0 Mbit/s	19 dB
16-QAM	$\frac{1}{2}$	4	416	208	52.0 Mbit/s	22 dB
16-QAM	$\frac{3}{4}$	4	416	312	78.0 Mbit/s	26 dB
64-QAM	$\frac{2}{3}$	6	624	416	104.0 Mbit/s	30 dB
64-QAM	$\frac{3}{4}$	6	624	468	117.0 Mbit/s	31 dB
64-QAM	$\frac{5}{6}$	6	624	520	130.0 Mbit/s	32 dB

$\mathbf{N_{CBPSC}}$ – number of coded bits per single carrier
$\mathbf{N_{CBPS}}$ – number of coded bits per OFDM symbol
$\mathbf{N_{DBPS}}$ – number of data bits per FDM symbol
Guard interval – 800 ns

Avoidance (CSMA/CA), Carrier Sense Multiple Access/Collision Detection (CSMA/CD)),

- collision probability as a function of number of UTs, MAC Service Data Unit (MSDU), and Maximum Transmission Unit (MTU) of Internet Protocol (IP) (usually 1,500 bytes),
- path loss between transmitter and receiver,
- Receiver (Rx) noise and interference from neighboring stations.

The approach for analytically estimating IEEE 802.11n MAC protocol capacity is to evaluate, in asymptotic conditions (i.e., all stations have a non-empty scheduling queue at any time), the ratio of average transmission time in case of an ideal, delay-free channel (also referred to as packet length), and the average time the channel is actually occupied in transmitting a message successfully (also referred to as the average virtual transmission time). Hence, the latter additionally includes channel delay, ACKnowledgment (ACK) transmission time, inter-frame spacings, and back-off times in case of repeated transmission attempts. Eventually, this approach resolves into the calculation of a collision probability p_c depending on the number of users n sharing the medium. p_c is computed numerically since it cannot be separated in the resulting equation [CCG98, CCG00]:

$$p_c + \left(1 - \frac{2(1-2p_c)}{31(1-p_c-32p_c^6)}\right)^{n-1} = 1. \tag{4.24}$$

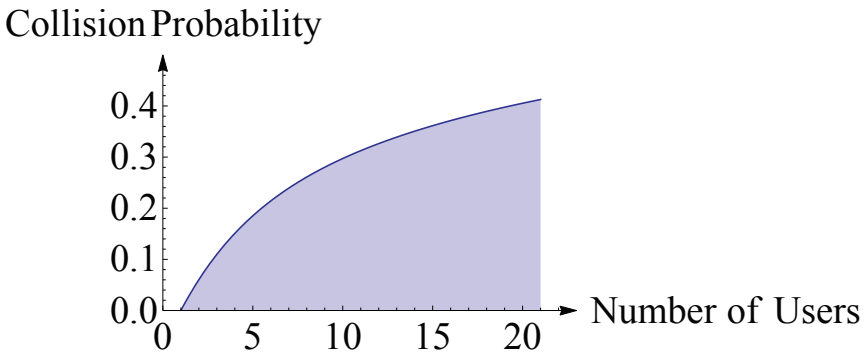


Figure 4.8: Collision probability depending on number of users [CCG00]

Figure 4.8 shows an interpolation of the (numerical) solutions $p_c(n)$ of Equation (4.24) for $n = 1, 2, \ldots, 21$. Using collision probability and propagation statistics [LYGL05], a factor can be computed by which access channel capacity has to be reduced due to medium contention. Figure 4.9 depicts the ratio of mean channel capacity $c(n)$ and gross channel capacity $C_{\max}$ (available if only a single user is connected, i.e., $n = 1$) as a function of n (number of associated users).

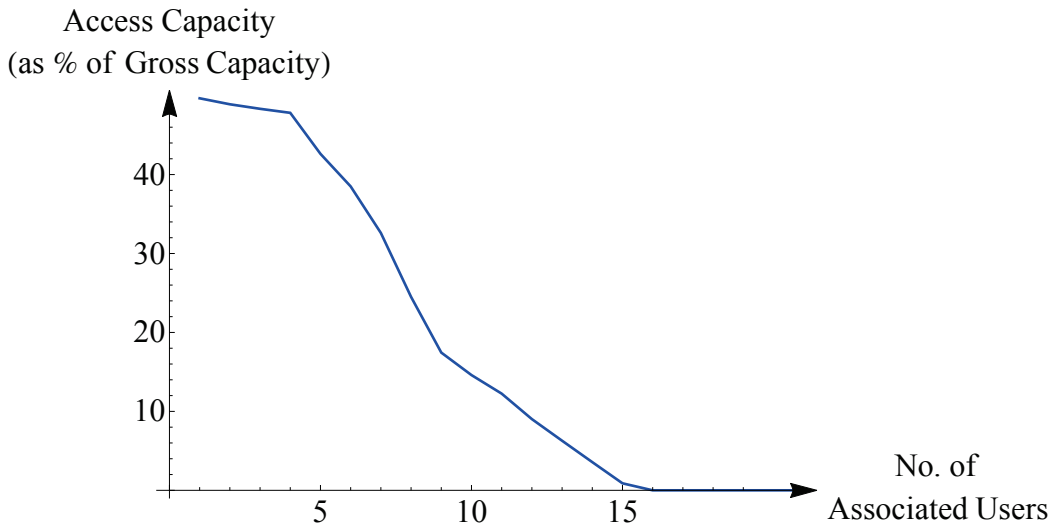


Figure 4.9: Capacity of an AP depending on number of associated users [Hay05]

Capacity C_{access} of an access channel can then be computed as a function of Receiver (Rx) Signal-to-Interference-plus-Noise Ratio $\text{SINR}_{\text{Rx,access}}$, similar to Equation (4.23). Additionally, a factor $\frac{c(n)}{C_{\max}}$ that accounts for medium contention is incorporated:

$$C_{\text{access}} = \frac{c(n)}{C_{\max}} \sum_{i=1}^{N_{\text{MCS}}} C_i \, u(\text{SINR}_{\text{Rx,access}} - \text{SINR}_{i,\text{MCS}}). \tag{4.25}$$

Available modulation schemes and code rates are the same as for the backhaul network (cf. Table 4.2). Due to only using $N_{\text{ss}} = 1$ spatial stream for access links, maximum gross data rates are halved compared to those of the backhaul links [IEE09].

4.2.2 Traffic Models

File Transfer Protocol Traffic

File Transfer Protocol (FTP) [PR85] traffic is frequently used for representing the class of best effort services with no strict bandwidth requirements. Therefore, the characteristics of the protocol are modeled by initially grouping user behavior into two temporal phases [CCLS01, CM99]: transmission time (both downlink or uplink) and reading time. The former obviously depends on the file size as well as the Quality of Service (QoS) parameters (particularly bandwidth and Packet Error Rate (PER)), the latter mainly is a function of two variables: time for processing the downloaded data (e.g., reading a text file, watching a video) and time until the next FTP request is sent. The usual approach is to derive the distribution of the required quantities from empirical data. Hence, file size S_{file} (in MB) is modeled according to a truncated log-normal function $p_{\text{FS}}(x)$ [SZJ+08]:

$$p_{\text{FS}}(x) = \frac{1}{\sqrt{2\pi}\sigma x} \mathrm{e}^{\frac{-(\ln x - \mu)^2}{2\sigma^2}}, \tag{4.26}$$

where $x \geq 0$, $\mu = 14.45$, and $\sigma = 0.35$. This translates into a mean file size of 2 MB and a standard deviation of 0.722 MB. The maximum file size is chosen to be $S_{\text{file,max}} = 5\,\text{MB}$.

Reading time is modeled in analogy to inter-arrival times of a queuing process. Consequently, an exponential distribution with probability density function $p_{\text{RT}}(x)$ is employed for generating random reading times [SZJ+08]:

$$p_{\text{RT}}(x) = \lambda \mathrm{e}^{-\lambda x}, \tag{4.27}$$

with $\lambda = 0.006\ \frac{1}{\text{s}}$, i.e., mean reading time $\frac{1}{\lambda} = 180$ s. For MTU, a size of 1,500 bytes is assumed. Further, a user is considered to be in outage if the data rate for FTP download falls below 128 kbit/s. Moreover, PER must not exceed 1%.

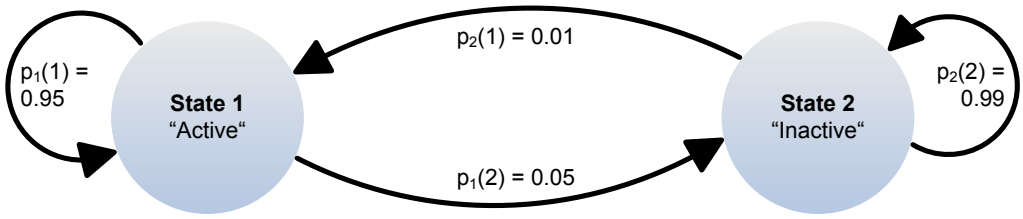


Figure 4.10: State machine for modeling user VoIP activity [SZJ+08]

Voice over IP Traffic

In system-level simulations and evaluations, Voice over IP (VoIP) traffic generally represents a second important class of network traffic, namely delay-sensitive traffic with limited bandwidth requirements. Similar to FTP traffic, modeling is done by analyzing the characteristics of empirical data.

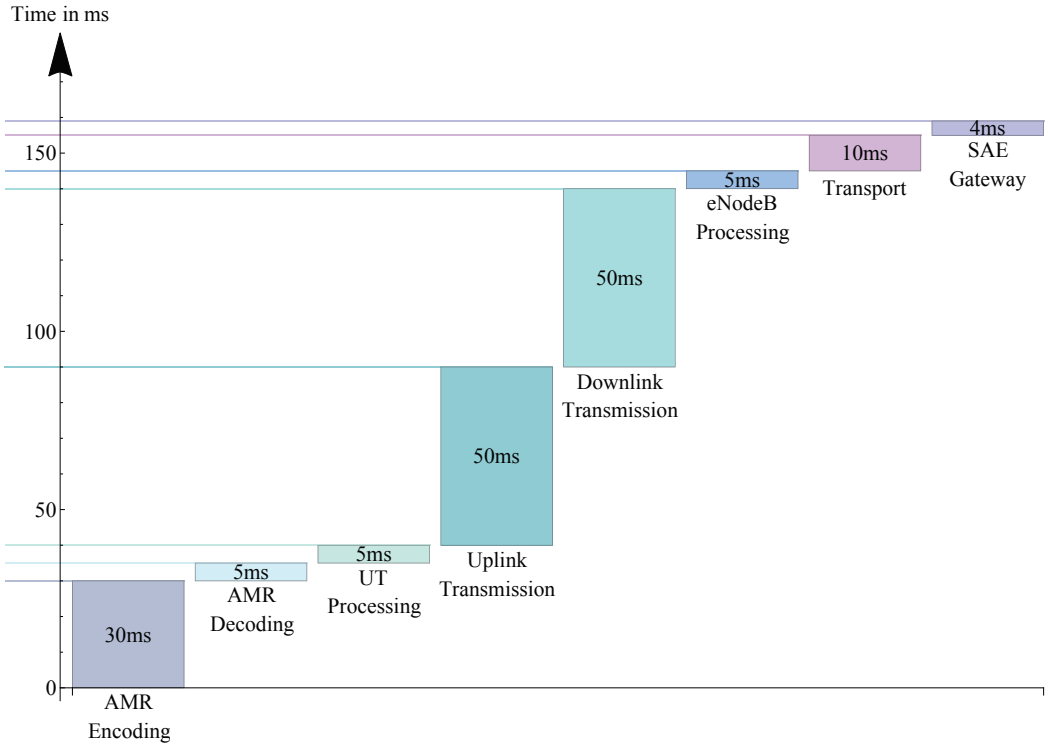


Figure 4.11: Delay budget of LTE VoIP service [HT10]

In the case of VoIP traffic, a Finite State Machine (FSM) is introduced for each user in the system. This FSM exhibits two states indicating a user's VoIP activity, "active" or "inactive". For each state, transition probabilities $p_i(j)$ are given, where i is the current state and j the possible user state for the next update. The model is evaluated for each user at rate $1/T$, T being the speech encoder frame duration (e.g., 10 ms G.729). Figure

4.10 visualizes the according finite state machine and depicts examples of numerical values for the transition probabilities. Prescinding from these concrete probabilities, active users are generally modeled to be more likely to switch to inactive state ($p_1(2)$) than inactive users starting a phone call ($p_2(1)$). Consequently, it is more probable to remain in the inactive state ($p_2(2)$) than to continue an ongoing phone call ($p_1(1)$). For QoS analysis, a user is considered to be in outage if downlink (uplink) transmission time of more than 2% of VoIP packets exceeds the bound of 50 ms [SZJ+08]. End-to-end delay, including Adaptive Multi-Rate (AMR) encoding and decoding, should preferably remain below 200 ms, which is the typical delay in circuit-switched networks. Figure 4.11 depicts a typical end-to-end delay budget calculation for packet-switched voice services, e.g., LTE VoIP.

4.2.3 Routing in Coordinated Wireless Backhaul Networks

Efficient multi-objective multi-hop routing poses a significant challenge in WBNs and therefore has been subject to numerous research efforts. In IP-based networks, routing is usually executed on layer 3 (network layer) of the Open Systems Interconnection (OSI) reference model. Extracting source and destination IP addresses of each IP packet and comparing them with the available routing table, a router can decide on the next target node of a packet. Routing tables are compiled by interior and exterior gateway protocols for delivering packets within and across Autonomous Systems (ASs), respectively [SS04, RLH06]. Using specified metrics for route computation, these protocols follow various optimization criteria. Reactive protocols produce routing information on demand, while proactive protocols keep an updated table of routes in each node. When faced with larger networks, reactive protocols suffer from QoS problems (increasing and varying latency) when determining a route, while proactive protocols are accompanied by rising overhead [BMJ+98]. In order to overcome these problems, professional operators usually introduce hierarchy levels to their networks, thereby improving the performance of proactive routing approaches. For overcoming latency limitations, several protocols, such as Multi Protocol Label Switching (MPLS) and Better Approach To Mobile Ad-hoc Networking (B.A.T.M.A.N.) advanced, introduce layer 2 switching as an efficient alternative to layer 3 routing. Appendix B presents typical routing approaches and protocols that can be applied in WBNs. Further, a set of common routing metrics is presented.

4.3 The Context Domain of Wireless Backhaul Networks

In today's world of pervasive mobile communications, context has become a key enabler for any kind of service execution and adaptation. For this purpose, state-of-the-art system implementations follow a multi-layer model for context management: context provisioning, context processing, context distribution and consumption, each of them implementing defined (disjunct) functionalities that, as a whole, represent the context domain of a wireless communications system. In the context of this thesis, the author introduces two classes of services that exploit context information. Firstly, conventional user services, including applications for mobile devices, and secondly, network and transport services, referring to any kind of functionality that optimizes network management. For both service classes, context is gathered from heterogeneous sources, such as networks, physical and logical sensors (e.g., temperature, ambient noise, Global Positioning System (GPS), databases, social networks, etc.) and utilized for optimized service adaptation and performance improvement purposes. Figure 4.12 summarizes the concept of continuous sequential sensing, processing, and action as well as adaptation for facilitating context-aware services. Following this logic, this section presents state-of-the-art and new developments in the field of context, context awareness, and context management systems.

4.3.1 Definitions of Context and Context Awareness

The notion of context and context awareness can be found in manifold variations in different research areas of computer science. Hence, similar to other concepts, there is no unique definition of the two terms, rather, depending on the application area, different definitions are considered and used.

One of the most commonly utilized definitions for context is given by Dey and Abowd [DA00]. They refer to context as "any information that can be used to characterize the situation of entities (i.e., a person, place or object) that are considered relevant to the interaction between a user and an application, including the user and the application themselves." More precisely, Dey [DAPW98] explicitly enumerates emotional state, focus of attention, location, and orientation for describing a user's context. Other authors, such as Ryan et al. [RPM98], list the user's location, environment,

identity, and time. More abstract definitions by Hull et al. [HNBR97] or Brown [Bro96] consider context as the aspects of the current situation or the elements of the user's environment the computer knows about, respectively. More systematically, Zimmermann [Zim07] categorizes context into five classes, namely time, location, activity, relations, and individuality.

Context-awareness has been used as a system or object characteristic in both narrow and broad meanings. Additionally, in earlier research work, it was defined in a less comprehensive manner, serving as synonym for single characteristics. Examples for the latter category can be found in [Bro96] ("adaptive"), [CTB+95] ("reactive"), [EHC+93] ("responsive"), [HNBR97] ("situated"), [RAH98] ("context-sensitive") and [FKS97] ("environment directed"). Examples for the former category include one of the first definitions of the term, where context-awareness refers to a software capable of adapting to its location of use, nearby people and objects, and to changes to those objects over time [ST94]. Similarly, the notion of adaptability can be found in [DAPW98] describing a software system or in [SDA98] characterizing a computational service. Further definitions refer to applications being able to dynamically react to changes of their own or the user's context, such as [WJH97], [KSB98], [DMCB98], or [BBX97]. While some authors assume a completely autonomous adaptation of a system or software, others explicitly require a user input or user guidelines before executing any change, e.g. [FKS97]. This thesis will basically use the popular definition of Dey and Abowd [DA00]:

> *A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task.*

In the scope of this thesis, *user* will be replaced by the more comprehensive term *entity*, also including any hardware or software component of a system.

The distributed nature of context-aware services and applications requires information to be exchanged between large numbers of entities, as for example shown in Figure 4.12. For instance, particular classes of services and applications require near real-time execution in order to provide useful results and outputs. The correct representation [DN09] of context and the efficiency associated to its processing thus become key in the development of next generation context-aware services. Solutions from industry, such as those based on eXtensible Mark-up Language (XML) (e.g., Web Ontology Language (OWL) [WZGP04]) constitute an important contribu-

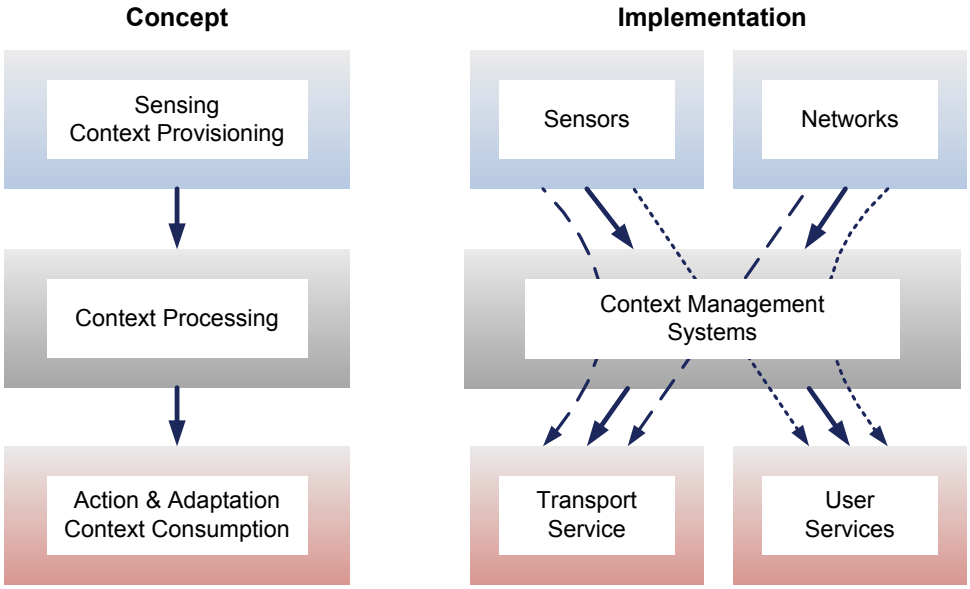


Figure 4.12: Concept and implementation of layered context data handling

tion, but need to be revised under the scope of integrating new and smaller devices that do not have the processing power to produce or consume context represented in XML. Furthermore, the representation must lend itself to reasoning and inference techniques to be used, such as classification, taxonomies, data mining, clustering, prediction, and pattern recognition [APH07]. In order to generate reliable results, the inclusion of a model for data quality, such as accuracy of location information or level of anonymization is required [EIKP08]. The joint consideration of these aspects has led to the development of heterogeneous architectures for context management systems, as depicted in the following section.

4.3.2 Architectural Concepts for Context Management

Context management systems have been designed to acquire, process, manage, and distribute context information according to the specific needs of applications and services. Starting from the very beginnings in the early 1990s, a broad variety of such systems have been developed, ranging from purely centralized to distributed architectures [RE12]. As a direct consequence, nowadays, location-based services are easily available for mobile phones, either running on a remote server or downloaded for local execution, using context coming from built-in GPS modules. Manifold applications exploiting other sensors deployed on high-end phones, such as gyro-

scopes or microphones, have become commonplace. However, context data can also be utilized for tasks less visible to end users, going beyond conventional end consumer services. Optimizing radio access network performance [MKSS09a] and policy-based networking approaches, such as adaptive resource admission control [SV02], context-aware network policies [PB02], and dynamic traffic engineering and service level specifications [TFP+02], are groundbreaking examples. However, independent of the final exploitation, a context management system is required in order to properly supply the collected context, thus making its consumer "context-aware".

Evolution of Context Management Architectures

Legacy context management systems typically exhibit application-specific and frequently proprietary structures, functionalities, and data representation formats. The *Active Badge Location System* by Want et al. [WHFG92] was one of the first context-aware applications. The current location of a user was determined using infrared-based techniques. In the 1990s, location-aware tourist guide systems were developed that replaced or complemented traditional tourist guides [AAH+97, SEF+98, CDM+00]. In the following years, there have been improvements with regard to both architectural frameworks for context-aware systems and the diversity and quality of available context information (e.g., advanced positioning technology [JLL+07]). European projects such as *SPICE* [SPI08, ZZM+06] and *MobiLife* [Mob06, FPN+05] have set milestones in integrating context awareness functionalities into their respective architectures.

The work of this thesis has significantly contributed to the development of the context management framework developed in the European Union (EU) research project *C-Cast* [CC10]. Figure 4.13 depicts the building blocks of the C-CAST context management platform. Moreover, the underlying producer-consumer model has been further refined for a versatile embedding of the platform in various settings, among them

- intelligent radio network access [MKSS09a, MKSS09b, SJMS13],
- heterogeneous access management [MKSS10, KMSS10a, CMK+10],
- mobile cloud computing [KMSS10b],
- user classification for efficient multicasting [MKSS11a, MSKS11],
- Radio Resource Management (RRM) [SWP+10, KMSS11, KLM+11],
- indoor and outdoor localization [MRS+10, SWM+11],

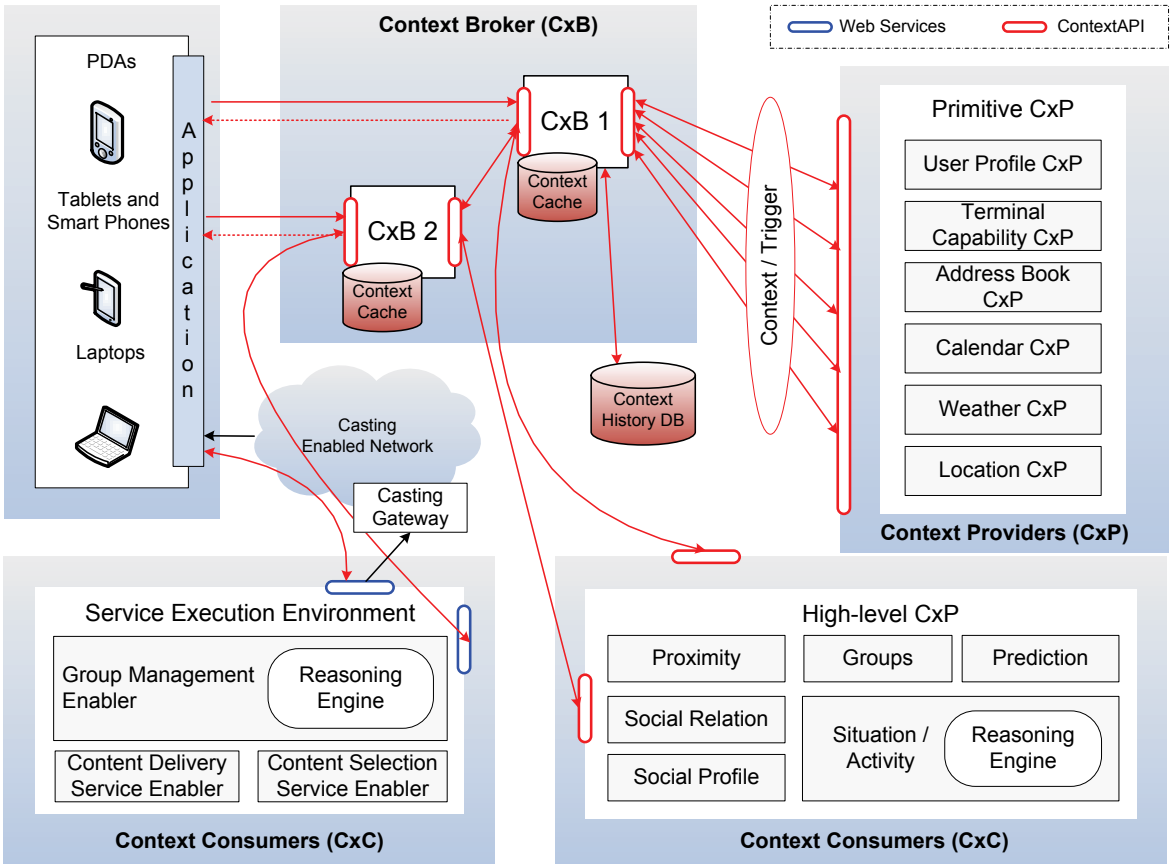


Figure 4.13: C-CAST context management architecture [MKSS09a]

- smart public environments and smart home [MSMS10, SSWM+10, SKMS12, MMS10],
- the Internet of Things [SMKS09, SKMS11],
- distributed systems [MAS+10, SMKa12, MSK+13],
- security and privacy [MKSS11b].

In the scope of this thesis, the context management platform has been used for supporting the planning and operation of energy-autarkic WBNs. As an alternative to the producer-consumer model, other authors have proposed agent-based approaches [Che04]. An extended overview of various context management systems and their characteristics is given in Table 4.3.

The multitude of today's engineering systems labeled "context-aware" requires a basic system classification. Winograd [Win01] differentiates between three architectural styles:

- **Widgets** are composed of Graphical User Interface (GUI) elements and provide a public interface for a physical or logical sensor. Hiding

low-level sensing details and allowing for a reusable and easy application development. A disadvantage is their low robustness with regards to component failures.

- The more general **networked service** model exploits specific discovery techniques to find networked context services. Besides increasing robustness, the approach makes use of available protocols and interface specifications, allowing for a more extensive applicability to heterogeneous context sources. However, usability efficiency decreases.
- The **blackboard** model represents a data-centric view. By means of asynchronous communication (e.g., publish-subscribe protocols), context data are placed on the blackboard and forwarded based on event subscription. The registration of new context sources is simplified, while communication efficiency decreases due to limited scalability of single blackboards. Distributed approaches can overcome these limitations.

For a long period of time, many of the context-aware systems in science and practice have suffered from at least one of the following drawbacks: limited data variety (e.g. exclusive focus on geographic data for location-awareness), proprietary architecture (including data format and interfaces), limited scalability and extensibility, limited flexibility (application-specific design), costly accessibility, no or limited functionalities for Authentication, Authorization, and Accounting (AAA), insufficient mechanisms for data security and trust [BZMK09]. Despite these drawbacks and the heterogeneity of available architectures, the majority of the systems consist of common elements. Since building context management systems for every application domain has not proven practical, one of the primary objective of any modern, state-of-the-art context management architecture is to decouple context management from application in order to enable system reuse and support in many areas [vA01]. Consequently, most modern context management systems are designed according to the producer-consumer model with one or several intermediaries illustrated in Figure 4.14. Generally, such systems deal with:

- Context discovery and acquisition as a mechanism to locate and retrieve context from diverse sources (described in Section 4.3.2),
- context brokering as a means to decouple context provisioning and context consuming as well as to perform context data life cycle management tasks (described in Section 4.3.2),

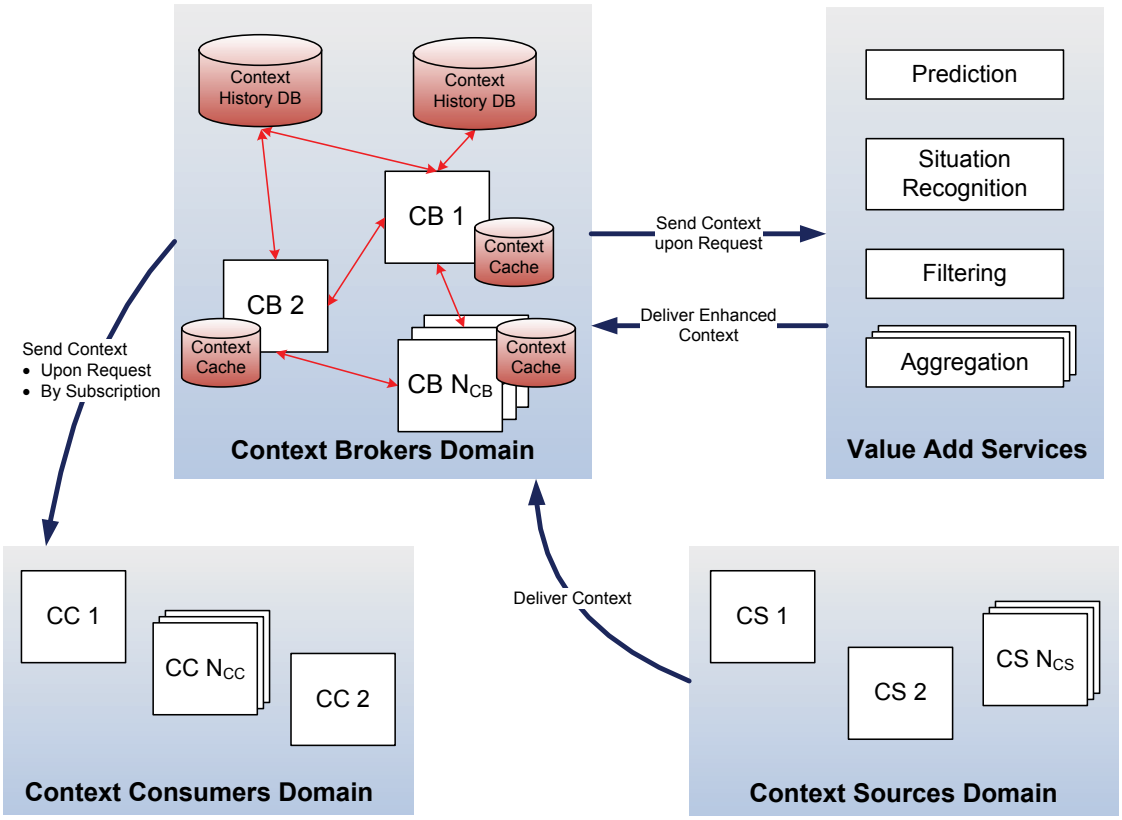


Figure 4.14: Building blocks of a state-of-the-art context management architecture

- context reasoning methods for context analysis and abstraction (described in Chapter 5),
- context diffusion/dissemination for efficiently propagating context while ensuring availability and reliability (described in Section 4.3.2).

Context Acquisition

Context acquisition refers to the process of collecting context data from heterogeneous sources. For categorization of context sources, Chen [Che04] proposes the following classes:

1. **Direct access to sensors** - The availability of wireless sensor networks, deployed in arbitrary locations, as well as sensors integrated into devices (e.g., user terminals) has facilitated an easy access to environmental context, such as location or temperature, as well as network context information. However, due to individual hardware and software realizations, direct access requires many different implementations since no common interfaces exist.

2. **Middleware infrastructure** - Overcoming shortcomings of the direct access approach, middleware infrastructure aims at strictly separating the processes of context acquisition and further processing. Introducing such an abstraction layer improves system extensibility and reusability.
3. **Context server** - In this approach, a resource-rich context server takes over the task of managing context data received from various context sources. It relieves sensors and terminals from handling context requests from other, context consuming entities. Context servers can be organized in a centralized or distributed manner.

In the domain of communications as well as in the scope of this work, it has proven sensible to differentiate between the following context source categories: physical sensors (e.g., for temperature, light, humidity, power meters), logical sources (e.g., data from databases, social networks, or advanced context reasoning functions), and network sources providing various data, such as QoS parameters (typically bandwidth, delay, jitter), access point coverage, network capacity, or network utilization.

Context Brokering and Storage

Centralized or distributed entities in charge of managing collected context data using an aligned representation format (such as key-value tuples or XML derivatives) form the cornerstone components of any context management system. Their design determines the extent of scalability and extensibility of a context management system as a whole as well as of individual components, i.e., the capability of accommodating both an increasing number of the same entities and data as well as new kinds of entities and data, respectively. Independent of their actual capabilities, according to Chen [CFJ03a], context broker entities

> *manage and maintain a shared model of context. [...] A context broker acquires contextual information from heterogeneous sources and fuses it into a coherent model that is then shared with computing entities in the space.*

To fulfill these tasks, a set of fundamental system functionalities is required, a selection of which is presented in Appendix C.

Context Distribution

Context distribution refers to the process of context sharing between multiple context producing and consuming entities in a context aware system. For this purpose, context brokers as well as many context sources implement standardized interfaces facilitating the provisioning of context to consuming entities of a system. Two major communication schemes for context dissemination can be differentiated: asynchronous publish-subscribe mechanisms and synchronous, query-based mechanisms. The two approaches provide important means of decoupling for facilitating scalable, distributed communication with low coordination overhead [CC08]:

- **Space Decoupling** - The interacting entities do not need to know each other. The publishers (providers) publish information through an event/information service and the subscribers (consumers) receive information indirectly through that service. Publishers and subscribers usually do not hold references to each other.
- **Time Decoupling** - The interacting entities do not need to synchronize their communication, i.e., the publisher might publish some information at time t_0, while the subscriber can request this information at any time $t_1 \geq t_0$.
- **Scale Decoupling** - Context sources (that frequently have very limited computing and energy resources) publish new data only once. The resource-intensive task of fulfilling numerous, potentially concurrent requests for context retrieval is performed by a resource-rich entity. Scaling problems at the context source can hence be eliminated.
- **Operator Decoupling** - Using open and standardized interfaces (e.g., RESTful designs [Rod08]), different components in a context management systems can be operated by different entities without sacrificing the advantages of proprietary component development. This allows for distributed, plug and play system layout; entrance barriers are lowered to a minimum and diversity as well competition are fostered.

In Table 4.3, the author has aggregated an overview of state-of-the-art context management systems and their major functions.

Table 4.3: Classification of context management systems

Reference System	Year	Context Data	Data Format	Available Functions	Interface Design	Application Domain
Active Badge Location System [WHFG92]	1992	location	ID-location-time triple	context cache, visualization	serial (RS-232), TCP/IP	localization, proximity, alarm
Context Toolkit [DA99]	1999	heterogeneous, multi-level	attribute-value tuples	context cache, reasoning engine, context widgets for abstraction	HTTP over TCP/IP, publish/subscribe, query-based	context-aware applications
GUIDE [CDM+00]	2000	geographical GPS, personal, environmental	object-based	context cache (centralized), entity referencing, visualization	TCP/IP, query-based	advanced tourist guide (route composition, hotel bookings)
CoBrA [CFJ03b]	2003	heterogeneous, multi-level	OWL-based	context cache, persistent storage, reasoning engine, security & privacy	SOAP, TCP/IP	smart space
SOCAM [GPZ04]	2004	heterogeneous (internal and external)	OWL, RDF, multi-level	context cache, persistent storage, provisioning, reasoning engine, security & privacy, service locating mechanism	OSGi service platform, push and pull mode	context-aware mobile and vehicular services
MobiLife Context Management Function (CMF) [FPN+05]	2005	heterogeneous (network, terminal, environment, web)	OWL-based, XML, description logic	context cache, persistent storage, provisioning, reasoning engine, privacy & trust, visualization	HTTP, publish/subscribe, query-based	context-aware applications and services in mobile communications
SPICE CMF [ZZM+06]	2006	heterogeneous (e.g., physical sensors, profiles, system data)	XML	context cache, persistent storage, provisioning, user profile management	RESTful (HTTP over TCP/IP)	mobile service platforms
C-CAST Context Management System	2010	network, environment, logical (e.g., user profile, weather)	XML	context source registry, context cache, persistent storage, entity relationships, visualization	RESTful (HTTP over TCP/IP), publish/subscribe, query-based	multicasting, content selection, group forming

4.3.3 Ontologies for Contextual Reasoning

Contextual reasoning refers to any (both elementary and advanced) method for context data analysis and abstraction. As illustrated in Figure 4.15, advanced reasoning algorithms are designed to infer complex context knowledge (e.g., user intentions, roles, and duties) that cannot be directly acquired from sources (such as physical sensors) or to detect and resolve inconsistent knowledge frequently occurring as a result of imperfect sensing. Therefore, while context brokering usually contents itself with rather simple models for data representation, reasoning algorithms have to operate on versatile context data models (ontologies) capable of representing contextual objects on different semantic levels, describing the relationships among those objects, reflecting dynamic changes in their attributes and relationships by introducing the notion of time, as well as formally describing events and their consequences on an object's parameters.

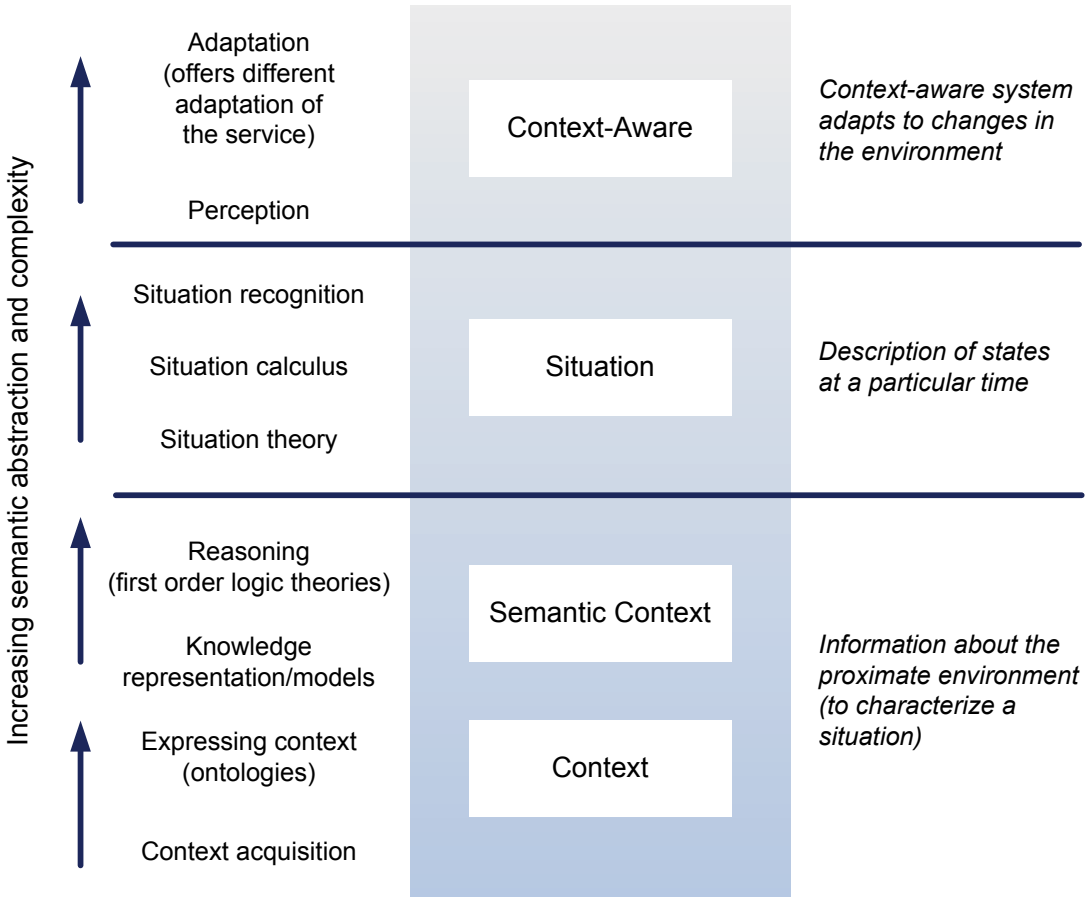


Figure 4.15: Semantic layers of context (adapted from [BZMK09])

Summarizing, Korpipää et al. [KM03] define the following principles for designing ontologies: (1) domain compatibility, (2) simplicity, (3) practical access, (4) flexibility and extensibility, (5) support for reasoning, (6) generality, (7) efficiency, and (8) expressiveness. Context ontologies, such as developed in [CFJ03b, KMK+03, GPZ04], are implementations striving for meeting these requirements. Particularly referring to principle (5) and (7), such ontologies have to facilitate reasoning with flexible specification of input and output (as defined by the respective requesting application) and to incorporate according reliability and availability metrics.

4.4 Summary

In Chapter 4, the author has motivated a multi-domain approach for modeling coordinated WBNs. The benefits of such a modeling perspective include a more realistic representation of actual network deployments that also reflect the interdependence between physical, networking, and computing resources. Moreover, it allows the application of selected innovative multi-domain analytical methods, e.g., known from CPS theory, for WBN performance evaluation. For this purpose, three domains, namely physical, network, and context domain, have been introduced. With respect to each domain, the author has presented and, if necessary, modified and extended relevant state-of-the-art concepts and finally integrated them according to the approach outlined in Section 3.3. Besides referring to proven models from evaluation methodology frameworks (e.g., [SZJ+08]), this chapter has outlined how system performance benefits from the new context domain and how a context management framework is integrated from a system architecture perspective. By means of the developed ideas, Problem 3.15 can now be refined as follows:

Problem 4.28

$$\begin{aligned}
\text{minimize} \quad & \alpha_Q \sum_{i_1,i_2} a_{i_1,i_2}\, \delta_{\Pi(i_1),\Pi(i_2)} - E(\mathbf{A}, \Pi(\mathbf{D})) - F && \text{(4.28a)}\\
= \; & \alpha_Q \sum_{i_1,i_2} a_{i_1,i_2}\, \delta_{\Pi(i_1),\Pi(i_2)} - \alpha_E \sum_i \Big(Y_i - \int \pi_{\text{in},i}(t)\, dt \Big) - F \\
= \; & \alpha_Q \sum_{i_1,i_2} a_{i_1,i_2}\, \delta_{\Pi(i_1),\Pi(i_2)} \\
& - \alpha_E \sum_i \Big(\int \big(\eta_i\, A_i\, I(t) - (\pi_{0,i} + d_i(t)\, \pi_{\text{out},i}) \big)\, dt \Big) - F
\end{aligned}$$

s.t.

$$-\chi(\mathbf{A}) \leq -k \tag{4.28b}$$

$$\chi(\mathbf{A}) \leq d_{\max} \tag{4.28c}$$

$$\mathbf{A} - \mathbf{A}^{\mathrm{T}} = 0 \tag{4.28d}$$

$$a_{i,i} = 0 \;\; \forall\, i \in \{1, \ldots, m\} \tag{4.28e}$$

$$a_{i_1,i_2} \in \{0, 1\} \;\; \forall\, i_1, i_2 \in \{1, \ldots, m\} \tag{4.28f}$$

$$-\sum_{i_2} a_{i_1,i_2} \leq -d_{i_1,\min} \;\; \forall\, i_1 \in \{1, \ldots, m\} \tag{4.28g}$$

$$\sum_{i_2} a_{i_1,i_2} \leq N_{i_1} \;\; \forall\, i_1 \in \{1, \ldots, m\} \tag{4.28h}$$

$$\pi_{i,\mathrm{out}}(x) \leq \pi_{\mathrm{out},\max} \;\; \forall\, i \in \{1, \ldots, m\}, \; x \in \{1, \ldots, N_i\} \tag{4.28i}$$

$$a_{i_1,i_2}\, \delta_{\Pi(i_1),\Pi(i_2)} \leq \delta(\pi_{\mathrm{out},\max}) \tag{4.28j}$$

$$\mathcal{P}_{i_1,i_2} \neq \emptyset \;\; \forall\, i_1, i_2 \in \{1, \ldots, m\}, \; i_1 \neq i_2 \tag{4.28k}$$

$$-\beta_i \leq 0 \;\; \forall\, i \in \{1, \ldots, m\} \tag{4.28l}$$

In a subsequent step, Chapter 5 outlines basic and advanced optimization methods to be applied to the specified problem of context-aware, energy-autarkic network operation.

Chapter 5

Techniques for Context-Enabled Optimization

Techniques for basic, statistical context analysis, such as feature extraction (mean, variance, etc.), aggregation, and fusion, belong to the standard functions of any context management system for network management. However, real differentiation among value-add services is achieved by providing advanced methods for context reasoning originating from the area of artificial intelligence. Neural and Bayesian networks, (hidden) Markov models, Kalman filters, and clustering algorithms constitute common methods for facilitating machine learning, spatio-temporal data mining, rule-based processing and situation recognition. Optimization and search methods complement this set of techniques. They are applied for purposes such as user classification (e.g., users, content), content selection and matching, and situation prediction. The following sections describe state-of-the-art techniques for optimization and search problems as well as for contextual reasoning and motivate their application in context-enabled optimization of an energy-autarkic Wireless Backhaul Network (WBN). The analysis serves as the basis for developing novel TC algorithms as presented in Chapter 6.

5.1 Optimization and Search

5.1.1 Linear Programming

Linear programming problems (or briefly Linear Programs (LPs)) are the best investigated and the least complex class of optimization problems. LPs are termed *linear* because the constraints they include are expressed as linear equations or linear inequalities. Moreover, the objective function linearly depends on the set of variables to be determined. LPs have first

been expressed and investigated formally during World War II, e.g., for solving goods transportation problems [NT93]. It was not until 1947 that Dantzig introduced the *Simplex Algorithm*, which until to date is the most popular universal algorithm for solving LPs. It first converts an LP into the *augmented* or *slack form* and afterwards exploits the fact that the optimum objective function value is to be found on one of the extreme points of the polytope that represents the feasible region of a multi-dimensional LP [DT97].

Formally, a linear optimization program can be expressed in the following canonical form [DT97]:

$$\underset{\mathbf{x}}{\text{minimize}} \quad \mathbf{c}^{\mathrm{T}}\mathbf{x} \tag{5.1}$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \geq \mathbf{b} \tag{5.2}$$

$$\mathbf{x} \geq \mathbf{0} \tag{5.3}$$

Here, $\mathbf{x}$ is the m-dimensional vector of variables that need to be determined as to minimize the objective function 5.1, while respecting the constraints 5.2 and 5.3. From a more illustrative perspective, LPs are nothing else but optimizing a linear function on a convex polytope. An example is illustrated in Appendix D.1. LPs and according solving algorithms are a very versatile instrument. Therefore, they enjoy widespread application in many domains (particularly in industry) to find or approximate optimum solutions. For example, LPs are used by logistics companies to optimize delivery routes, by network providers to support network planning, by energy companies to plan power grids, or in production environments to optimally source and allocate raw materials and other input factors. Generally, network flow, matching and assignment, transportation, matrix games, or transshipment comprise some of the problem categories solved by LPs and variants of the Simplex Algorithm [NT93], [DT97].

5.1.2 Non-Linear Programming

Frequently, real-life problems from different domains are inherently non-linear. Since LPs cannot be applied in these cases, Non-Linear Programs (NLPs) require a transformation to an LP. Alternatively, non-linear optimization techniques can be considered. In the following, the author briefly presents *gradient-based* methods and *direct search* methods for solving NLPs.

Gradient-Based Methods

Sequential Quadratic Programming (SQP) comprises a set of numerical methods for solving a NLP. The basic idea is to iteratively find a solution $\mathbf{x}^k$ to an updated local approximation to the NLP, where the local approximation consists of a quadratic program QP_k. This yields a sequence of quadratic programs and according solutions that shall converge to the optimal solution $\mathbf{x}^*$ [BT95]. The motivation for choosing quadratic subproblems (more specifically, programs with quadratic objective functions and linearized constraints) stems from the observation that, on the one hand, they are able to reflect the non-linearities of the original problem quite well and, on the other hand, reliable and efficient methods for solving such problems exist. The basic idea of SQP is further specified in Appendix D.4. Generally, the *steepest descent* and *line search* methods [Sch83], as well as *interior point* techniques [BGN00] belong to the most popular SQP techniques. In case of unconstrained problems or problems having only equality constraints, the well-established Newton's method [BSS06] can be applied. Generally, SQP is applied in various practical areas, among them science (e.g., control theory, structural design), engineering (e.g., mechanical design, electrical networks), and management (e.g., vehicle routing, stochastic resource allocation) [BSS06].

Clustering [Fri98] is a further important optimization technique. While there are also several direct search clustering techniques, self-organizing maps (SOM) are a prominent example for a direct search clustering technique. Further details on clustering are depicted in Appendix D.2.

Direct Search Methods

Genetic Algorithms (GAs) are meta-heuristic optimization methods belonging to the class of *evolutionary algorithms*. They are inspired by the evolutionary development of species. Accordingly, they try to mimic the mechanisms underlying the evolution process, e.g., selection, crossover, and mutation.

As depicted in [Mic94], a genetic algorithm has to consist of five fundamental components:

1. A genetic (coded) representation of potential solutions to the problem.
2. A method to generate an initial population of potential solutions to the problem.

3. An evaluation or fitness function determining the quality of a solution.
4. Genetic operations controlling the composition of the next generation solution.
5. Parameters for algorithm execution (such as population size, probabilities of applying genetic operations, etc.).

For implementation, a GA realizes three fundamental operations: crossover, mutation, and selection [Bod01]. *Crossover* refers to the rules for merging the genetic information of selected parent solutions to form offspring solutions, *mutation* to how to deform the genetic information of an individual solution, and *selection* to the process of choosing solutions from both the parents and their offsprings to form the next generation. For further details, the reader is referred to Appendix D.11.

GAs are utilized in many areas where optimization is the key objective, among them engineering (communications networks, navigation, object classification, e.g., in [GCL08]), operations research (logistics and fleet optimization, job scheduling, financial mathematics, e.g., in [LLT07, KP99]), and bio-informatics [vGP95]. Particularly, in network modeling, considerable results have been achieved in topology optimization, delay and cost minimization, and capacitated multi-point network design problems (e.g., [ES96, LC00]). Due to their analogy to evolutionary processes in nature, GAs can very well handle problems showing evolutionary behavior. Moreover, they are frequently capable of finding good solutions relatively fast. However, computation of (close-to) optimal solutions is generally accompanied by time- and resource-intensive algorithm executions. Moreover, the lack of general design rules for the appropriate choice of parameters (e.g., coding function, mutation and selection probabilities, mutation and crossover routines, termination condition) makes the efficient use of GAs a challenging task. In order to mitigate some of these drawbacks, extensions to the conventional GA have been introduced, among them adaptive GAs featuring time-variant algorithm parameters (e.g., adaptive mutation probability, varying crossover routines), hybrid GAs that integrate conventional optimization methods for a refined local search, thus exploiting additional information such as the derivative of the fitness function, and self-organizing GAs where crossover, mutation, and selection routines themselves are optimized applying GA logic [Bod01].

Simulated Annealing (SA) has first been introduced in 1983 as an algorithm belonging to the class of *hill climbing* algorithms [KGV83]. Hill

climbing refers to the capability of escaping from local minimums (i.e., climbing up the slope surrounding the minimum) by allowing the objective function to decrease to a certain extent. Consequently, these methods show an improved performance with respect to finding the global optimum solution instead of only a good local one. Simulated annealing is usually used for discrete optimization problems and mimics the according thermodynamic process. Certain solids, when heated up and, subsequently, annealed very slowly, settle into their most regular possible crystal lattice configuration, i.e., the lattice configuration with the lowest energy state and hence greatest stability. Appendix D.12 summarizes the formal concept of simulated annealing for optimization.

SA and its variants have successfully been used for solving optimization problems such as different scheduling problems (factory jobs, maintenance crew), routing, multiphase system analysis, and real-time multi-agent coordination. In signal processing, it has been applied to Infinite Impulse Response (IIR) filter design, equalizer evaluation, and channel estimation [Tan08]. The most important advantage of simulated annealing when compared to other methods is the seamlessly built-in ability to escape from local optima. Nevertheless, the algorithm also shows very good convergence behavior and is very well suited for parallel execution. However, these advantages come at the cost of considerably difficult design choices for the most influential parameters, such as neighborhood function, cooling schedule, and termination criterion. Moreover, simulated annealing is computationally expensive and comparatively slow when executed in a purely sequential manner [HJJ03]. Important extensions to the algorithm include improved neighborhood selection strategies, optimal termination criteria design, as well as the adaptive simulated annealing method for faster convergence in case of non–smooth cost functions (e.g., [ARP08, CL99]).

5.1.3 Dynamic Programming

The modern notion of Dynamic Programming (DP) has been introduced by Bellman in 1954, who coined the term and formalized the underlying theory [Bel54]. DP breaks down a complex problem into a set of smaller, potentially nested, subproblems to be solved recursively. The optimum solution to the overall problem can be derived by generating the optimal solutions for the smaller subproblems. These subproblems are modeled as a sequence of state transitions $\mathbf{x}_k = t(\mathbf{x}_{k-1}, \mathbf{u}_k)$, where $t : (\mathbb{X}, \mathbb{U}) \rightarrow \mathbb{X}$ is the

transition function, $\mathbf{x}$ is a state from the state space $\mathbb{X}$, and $\mathbf{u}$ the decision vector from space $\mathbb{U}$. N $(k = 1, \ldots, N)$ denotes the length of the sequence [BK65]. The derivation of the Bellman equation is depicted in Appendix D.3.

Besides its natural application area of operations research, e.g., military and industrial planning problems, DP is frequently utilized in domains, such as sequence alignment in speech recognition and genome analysis, optimal ordering for performing chain matrix multiplications, hidden Markov model problems, or machine learning (*temporal difference reinforcement learning*) [Dre10]. In these areas, DP can fully unleash its strengths with respect to multistage decision making. However, DP techniques continuously fail in applications requiring a large number of state variables. This problem of high dimensionality (also referred to as the *curse of dimensionality*) is reinforced by the fact that DP needs to store state variables across multiple temporal stages [Dre10].

5.2 Contextual Reasoning

5.2.1 Regression Analysis

(Multivariate) linear models that aim at explaining and forecasting a depending variable by utilizing one or several explanatory variables are called regression models. Explanatory and depending variables are assumed to have a causal relationship, such as inflation and interest rate. Auto-regression refers to the case where previous values of the depending variable itself are used to generate forecasts. Given the vector of explanatory variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, the unknown depending variable y is approximated by $\hat{y}$ according to

$$y(k) = \hat{y}(k) + e(k) = \beta_0 + \langle \boldsymbol{\beta}, \mathbf{x}(k - \delta) \rangle + e(k), \tag{5.4}$$

where $e(k) \sim \mathcal{N}(0, \sigma_e^2)$ is the independent and identically distributed (iid) residual $(y(k) - \hat{y}(k))$ and δ denotes the time lag between depending and explanatory variables (usually $\delta \in \mathbb{R}_0^+$). $\boldsymbol{\beta}$ and β_0 are the model parameters to be chosen so as to minimize the mean squared error $\frac{1}{N-1} \sum_k e^2(k)$, $k = 1, \ldots, N$ ("ordinary least squares" estimation strategy) [Die01]. Variations of the standard linear regression model include trend-based seasonal models, Moving Average (MA) models, or Auto-Regressive Moving Aver-

age (ARMA) models. For evaluation of the explanatory power of regression models, a set of statistics is employed, among them the (adjusted) R^2 metric ("goodness of fit") giving the percentage of variance of y explained by the regression model, the *Durbin-Watson* statistic checking for serial correlation of regression disturbances, or simply sample variance $\hat{\sigma}_e^2$ working as an estimator for σ_e^2 and thus indicating the magnitude of the residual [Die01]. Moreover, Support Vector Machines (SVMs) are a supervised learning technique used for regression analysis and, in addition, for solving classification problems. Appendix D.5 depicts further details on SVMs. Regression analysis can be flexibly applied in time series analysis, such as stock market analysis or macro-economic models for explaining quantities, such as interest rates, Gross Domestic Product (GDP), or unemployment rates. However, despite the assumption of a causal relationship between explanatory and depending variables, regression analysis can only confirm (or reject) a *correlation* between y and $\mathbf{x}$. Applications in context processing include analysis of general correlation characteristics, e.g., entity correlation. Entities (e.g., persons, objects, rooms, etc.) and their context (e.g., activity, location, temperature, etc.) have exploitable interdependencies with each other. For example, correlated location data among persons might serve as an indicator for some form of joint activity, which might be further confirmed by similar calendar entries (given the according data is available). Similar logic applies to other context and entity categories and can be exploited in regression analysis.

5.2.2 Fuzzy Logic

The incorporation of feedback loops has become a common feature in system optimization, thus combining methods from optimization and theory of closed-loop control systems. In this context, Fuzzy Logic-based control is a practical alternative for a variety of challenging control applications since it provides a convenient method for representing and implementing a human's heuristic knowledge by designing nonlinear controllers using a static mapping between time-variant inputs and outputs. The term *fuzzy logic* has first been introduced in the context of fuzzy set theory [Zad65]. Belonging to the class of probabilistic reasoning methods, fuzzy inference employs fuzzy set theory for assigning to an element s its degree of membership μ to a fuzzy set $\mathbb{A}$, i.e., a set where membership can only be defined vaguely (e.g., set of fast moving terminals). In fuzzy logic, instead of requiring an

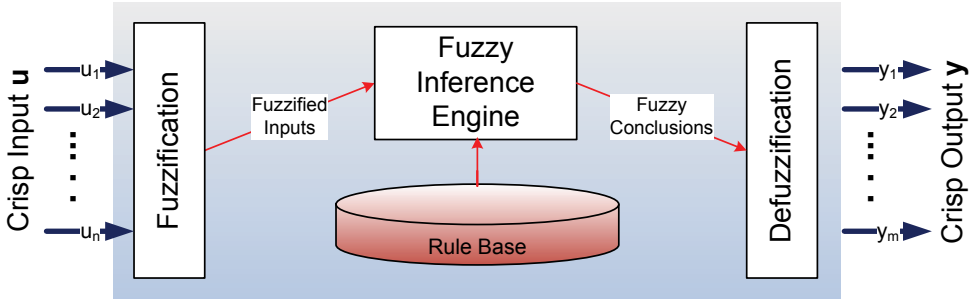


Figure 5.1: Schematic of a fuzzy controller [PY98]

exact binary true/false (i.e., 1 or 0) evaluation of a statement, a certainty $0 \leq \mu \leq 1$ is assigned to each statement. This is particularly helpful when heuristic knowledge and according rules are given as a vague description ("the user moves fast") rather than as a precise statement ("user velocity is $v = 10\ \frac{\mathrm{m}}{\mathrm{s}}$"). This can frequently occur in fuzzy control when rules for more abstract situations or system states and consequent actions are defined, e.g., **if** *traffic load is high* **then** *switch on additional radio interface.*

Fuzzy inference usually comprises four steps as illustrated in Figure 5.1 and depicted in Appendix D.6: fuzzification, rule matching, inference, and defuzzification. As input, the process takes n exact (sometimes also referred to as *crisp*) variables $(u_1, u_2, \ldots, u_n)$, the output comprises m exact (crisp) variables $(y_1, \ldots, y_m)$, representing the system variables to be controlled or optimized. Fuzzy logic has been applied in a variety of fields, among them control engineering, qualitative modeling, pattern recognition, signal processing, information processing, machine intelligence, decision making, management, finance, medicine, motor industry, and robotics (e.g., [TC01, ANNM04, BMVSSO04, BGF04, Fen06, CHG+01]. The most important advantage of fuzzy logic, besides well handling non-linear systems, is the capability of integrating heuristic human knowledge by allowing the mapping of both numerically exact quantities and non-measurable input (e.g., human observations) to semantically abstract situations. The required rules for system control and optimization can then be expressed on a semantically augmented level, i.e., there is no explicit need for complete unambiguity of defined situations and derived rules. Further, fuzzy control can easily incorporate multiple objectives by combining and weighing according rules. Obviously, this comes at the cost of introducing fuzziness with regard to applicable rules. Moreover, the development of an appropriate fuzzy controller requires significant experience since there are

neither any general guidelines nor any fuzzy systems theory for an efficient design of membership functions, matching and inference rules, as well as defuzzification methods [PY98]. Finally, for the same reasons, guaranteeing system stability in fuzzy control is a major challenge in any implementation [Fen06, Sug85]. Important extensions to conventional fuzzy logic control as presented here include fuzzy Proportional-Integral-Derivative (PID) control, neuro-fuzzy control, and adaptive fuzzy control [Fen06].

5.2.3 Kalman Filters

In 1960, R.E. Kalman published the concept of the *Kalman Filter* [Kal60] as a recursive algorithm for the linear filtering and prediction of discrete data. Ever since, the filter has been utilized for modeling stochastic processes by means of optimizing the minimum mean squared error, e.g., in mobile communications for tracking and predicting user position and velocity. Essentially, the filter, which bases on a Bayesian model, is comprised of two fundamental, alternating operations: quantity estimation (prediction) based on current model parameters and, subsequently, estimate correction and update of model parameters based on the according process measurement/observation. Mathematically, the Kalman filter shall estimate the state $\mathbf{x}_k \in \mathbb{R}^n$ of the underlying discrete-time, linear process at time k. Implementing the first of the two basic Kalman filter steps (prediction), an a-priori state estimate $\hat{\mathbf{x}}_k^-$ is defined as the expected state at time k. Similarly, in the second step (model update), $\hat{\mathbf{x}}_k$ is defined as the a-posteriori estimate of $\mathbf{x}_k$ that incorporates the additional knowledge gained at time step k. $\mathbf{x}_k$ is approximated using a linear combination of the a-priori and the a-posteriori estimates. The algorithmic details of the Kalman Filter are depicted in Appendix D.7.

The Kalman filter has several advantages: It allows for estimating past, present, and future behavior of quantities without requiring a full-fledged system model. Further, it operates reliably in the presence of noisy measurements and can be applied to many fields, among them navigation, signal processing, and economics [KB61]. While the conventional filter can be applied to processes governed by a time-discrete linear model according to Equations (D.25) and (D.26), several extensions of the filter have been developed in order to address other classes of processes. The Extended Kalman Filter (EKF) [Sor60] and the Unscented Kalman Filter (UKF) [JU97] have been designed to support processes following a non-

linear stochastic difference equation model for state transition and observation. For this purpose, the EKF continuously linearizes about the current mean and covariance, in other respects following the same methodology as the conventional Kalman filter, whereas the UKF implements the so-called unscented transform, a deterministic sampling technique for approximating mean and covariance of highly non-linear processes. With an underlying continuous-time state space model, the Kalman-Bucy filter [BJ87] abstains from the requirement of an underlying discrete-time process.

5.2.4 Particle Filters

Particle Filters are a generalization of the Kalman filter and thus also belong to the category of sequential Monte Carlo methods. Mathematically, particle filters estimate Bayesian models, with first applications coming from the area of growing polymers [HM54]. Their applicability for signal processing has first been shown in [GSS93]. In contrast to the Kalman filter, particle filters do not make the assumption of linear, Gaussian-distributed quantities, they rather operate on non-linear state-space and observation models without any pre-defined probability distribution [DKJ+03]. For further details, the reader is referred to Appendix D.8.

Extensions to the conventional filter are the so-called *Gaussian Particle Filter* [KD03a] that allows an improved approximation of the posterior distribution as well as the *Gaussian Sum Particle Filter* [KD03b] featuring approximation based on a linear combination of Gaussian distributions. The particle filter, like the Kalman filter, has the advantage of very flexible application areas, among them positioning, navigation, and tracking in wireless networks [GGB+02], wireless channel equalization, detection over flat fading channels [DKJ+03], and multi-user detection for Code Division Multiple Access (CDMA) systems [ADT00]. However, the filter can easily impose significant requirements on data processing and storage, as the number M of necessary particles has to increase with the dimension of the state vector $\mathbf{x}$ in order to keep estimation variance at an acceptable level [PMP+10].

5.2.5 Bayesian Networks

Bayesian networks fall into the category of symbolic but non-logicist approaches in Artificial Intelligence (AI). The term has its origins in a subdomain of statistics called *Bayesian statistics* that deals with conditional

probabilities as introduced in Bayes' theorem. The notion of Bayesian networks was introduced by Pearl in 1985 [Pea85]:

> "Bayesian networks are directed acyclic graphs in which the nodes represent variables, the edges signify the existence of direct causal dependencies between the linked nodes, and the strengths of these dependencies are quantified by conditional probabilities."

More specifically, a Bayesian network is an annotated directed acyclic graph $G = (\mathcal{V}, \mathcal{E})$ representing a joint probability distribution over a set of random variables $\mathcal{V}$ forming the vertices of G. The directed edges in $\mathcal{E}$ describe a hierarchy of nodes, indicating the dependency of children nodes on their parent nodes [FGG97]. Accordingly, parent variables are independent of their descendents. The formal concept as well as an example are depicted Appendix D.9.

Due to their theoretically solid design principles, Bayesian networks comprise a robust technique for processing uncertain information. They provide a consistent calculus for uncertain inference, i.e., one always derives unambiguous conclusions. Moreover, in contrast to neural network models, which usually appear to the user as a "black box", all the parameters in Bayesian networks have an understandable semantic interpretation and variables can be of arbitrary type (binary, discrete, continuous, etc.). As another benefit, Bayesian networks also exhibit a robust and smooth behavior with respect to model changes. In other words, marginal alterations of the network structure or conditional probabilities do usually not result in disruptive output changes. This also implies that the designer does need to provide exact input in order to get results within acceptable tolerance intervals [Myl05]. However, Bayesian networks also reveal several limitations, among them the possibly high number of variables that need to be considered, the difficulty to express subjective human beliefs and estimates in adequate (conditional) probabilities, and the high computational effort (Non-deterministic Polynomial time (NP)-hard) to perform full inference over the entire network [Pea88].

Similar to other probabilistic methods, Bayesian networks are applied in the areas of classification, signal processing, autonomous decision systems, and semantic search within different research fields such as biology, computer science, engineering, chemistry, and information technology. More specifically, they have been successfully used, among others, for building

intelligent agents and adaptive user interfaces, fault diagnosis, medical diagnosis, or process control [Myl05].

5.2.6 Neural Networks

The concept of Artificial Neural Networks (ANNs) is a direct imitation of neural networks in central nervous systems of biological organisms. In biological neural networks, neurons possess so-called axon terminals to interface to other neurons using synapses and dendrites as links, exchanging electrical signals of small amplitude with each other. If the amplitude of incoming signals surpasses a certain threshold, the neuron becomes activated and starts sending an action potential to connected neurons itself. In this manner, neurons form a complex network to propagate information through an organism. The fundamental ideas for ANNs as an analogy of their biological counterparts date back to 1943, when McCulloch and Pitts modeled neural activity using propositional logic [MP43].

As depicted in Figure 5.2, state-of-the-art artificial neural networks are set up in three layers: *input*, *hidden*, and *output layers*. Each layer contains a set of nodes (so-called *neurons*) that are connected with each other by weighted links. As an extension to Figure 5.2, the hidden layer can also consist of a more complex network of neurons and the output layer can be made up of more than one neuron. Generally, neurons of the hidden and output layers possess at least one incoming link, one outgoing link (that can be duplicated), and an *activation function* that decides whether the incoming stimuli are sufficient to activate the outgoing link. A more detailed description of how artificial neurons are modeled is presented in Appendix D.10. More complex, so-called *recurrent* ANNs do not only contain feed-forward paths, but the hidden layer also contains cycles. For example, recurrent networks allow for feedback loops, similar to closed loop control systems, and hence allow for decision making in challenging environments. Moreover, ANNs of limited structural complexity are also well suited for implementing learning algorithms, i.e., they can be trained to independently adapt their behavior. This is done by using training data to manipulate (optimize) one or more parameters of the activation function, according to a given objective function. However, when becoming sufficiently expressive to represent non-linear functions and thus more suited for making more abstract decision, ANNs become much harder to train. Due to these characteristics, ANNs nowadays are mainly applied, among

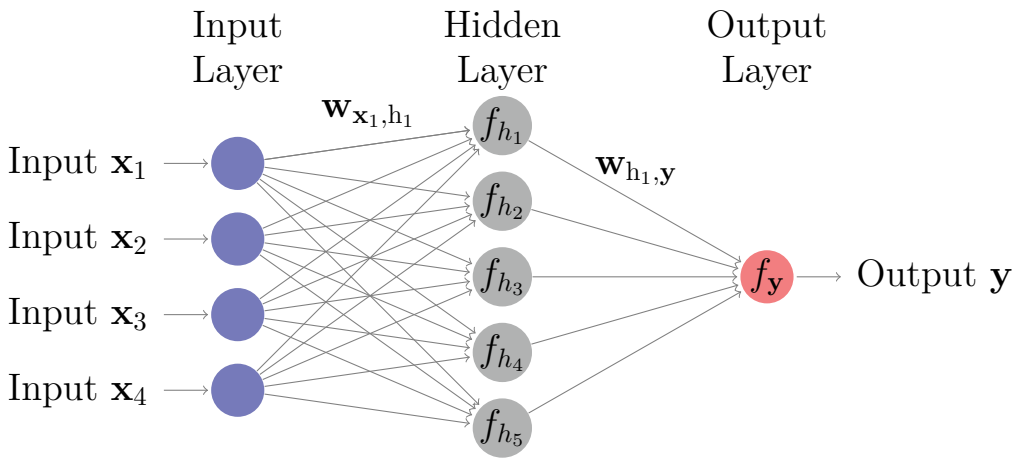


Figure 5.2: Basic Setup of a neural network

others, in sequential decision making, classification, pattern recognition, signal processing, as well as system and process control. Moreover, in the financial industry, they have found widespread use in automated trading systems [AM90].

5.3 Evaluation of Techniques

Sections 5.1 and 5.2 have delivered insights into a selection of optimization and learning techniques to generate context-awareness in the operation of energy-autarkic, coordinated WBNs. In Tables 5.1 and 5.2, the author gives an overview on strengths and limitations of the presented algorithms with respect to the application area and according problem categories of this work. Further, possibilities for extensions are listed. For solving different variations of optimization problem (4.28) and QAPs in general, the following optimization and search techniques are preferred:

- Linear Program (LP) – for LPs, very fast and efficient solution techniques are available. Moreover, many methods exist that conveniently transform non-linear problems to linear ones (e.g., reformulation linearization technique).
- Quadratic Program (QP) – convex QP relaxations [AB01] are used to obtain node-specific bounds. Using a branch and bound algorithm, this approach produces near-optimum results for QAP-class problems, such as topology optimization.
- Genetic Algorithm (GA) – GAs exhibit very good robustness, i.e., their ability of finding a good solution is independent from inaccurate

Table 5.1: Overview of optimization techniques

Method	Strengths	Limitations	Applications	Extensions
Linear Programming	well-researched, multitude of efficient solution techniques available	restriction to linear problems	network flow, matching and assignment, transportation	algorithms for converting non-linear programs into LPs
Sequential Quadratic Programming	proven and tested solving algorithms, QPs capture non-linearities well	convergence failure, implementation difficulties (approximation of Hessian matrix)	mechanical engineering, control theory, routing problems, resource allocation	advanced Hessian approximation techniques, control of convergence rates
Clustering Techniques	allow for hard and fuzzy clustering, simple implementation	limited stability, results may depend on initialization	user grouping, object classification, pattern recognition	self-learning techniques (e.g., adaptive cluster size)
Genetic Algorithms (GA)	finds near-optimum solutions fast, high robustness	challenging design, computationally intensive in finding optimum	topology optimization, logistics and fleet optimization, scheduling	time-variant parameters, hybrid GAs
Simulated Annealing (SA)	*hill climbing* capability, good convergence characteristics	computationally expensive, slow execution, complex design	routing optimization, filter design, real-time multi-agent coordination	adaptive SA, advanced neighborhood selection
Dynamic Programming (DP)	dynamic optimization over multiple time steps, multi-stage decision making	cannot handle high-dimensional state variables well, computation- and memory-intensive	sequence alignment, bioinformatics, machine learning, repeated matrix multiplication	stochastic DP, non-separable DP, discounted DP

input variables, which frequently is the case in real-world systems (e.g., SINR measurements).

- Simulated Annealing (SA) – SA is well capable of escaping from local optimums, which is an important characteristic since the given QAP (Problem 4.28) can exhibit enormous fluctuations of objective function values, even for "neighboring" solution vectors.

Implementations of these techniques have been available in *Wolfram Mathematica* [Wol14], the technical computing software utilized for the implementation of the simulation framework of this thesis. They have been parametrized according to the specific optimization objectives and restrictions of coordinated WBNs. Further, the context management system as described in Section 4.3 includes components implementing several of the algorithms to facilitate contextual reasoning. The following methods have been used primarily:

- Regression analysis – regression analysis is a fast and efficient way to describe a depending variable as a function of selected explanatory variables. Consequently, it can identify correlation between variables and further indicate causal relationships.

- Kalman Filter (KF) – the KF performs well in explaining time series by balancing observed values with statistical models; in case of suf-

Table 5.2: Overview of reasoning methods

Method	Strengths	Limitations	Applications	Extensions
Support Vector Machines	flexible w.r.t. classification criteria, stability	input data must be linearly separable, computationally expensive	localization, pattern recognition, page ranking	probabilistic kernel functions, support vector regression
Regression Analysis	can detect autoregressive behavior, broad applicability	cannot determine causal relationship between quantities	analysis of quantity correlations, time series analysis, and prediction	seasonal models, moving average methods
Artificial Neural Networks	autonomous decision making, broad applicability, learning possible	training of complex networks usually challenging	sequential decision making, pattern recognition, signal processing	recurrent ANNs, multi-layer ANNs
Fuzzy Logic	integrates heuristic human knowledge	complex logic design, ambiguity of fuzzy rules	control engineering, signal processing	fuzzy PID control, neuro-fuzzy control
Kalman Filter	estimation of past, present, and future values, robustness	limited to time-discrete linear Gaussian processes	navigation, signal processing, economics	Extended Kalman Filter, Unscented Kalman Filter
Particle Filter	generalized Kalman filter, high flexibility in design and applicability	usually significant requirements w.r.t. data processing and storage	positioning, navigation, channel estimation in wireless systems	Gaussian Particle Filter, Gaussian Sum Particle Filter
Bayesian Networks	exact inference possible, robust and smooth behavior	estimation of conditional probabilities, full inference is NP-hard	classification, agent systems, semantic search	advanced inference techniques (e.g., message passing)

ficient explanatory power of the filter, it can be used for predicting time series (e.g., user mobility).

- Bayesian Network (BN) – BNs excel in robustly deriving conclusions and abstractions from basic input data. Hence, they are ideal for contextual inference and are used for recognizing specific situations, e.g., network overload, high User Terminal (UT) densities, spatial distribution of energy buffers. Therefore, they are used for triggering topological reconfigurations and other network management decisions (e.g., resource allocation in the backhaul network).

Chapter 6 describes the specific algorithms that have been developed for addressing the problem of context-aware Topology Control (TC) in more detail.

Chapter 6

Topology Control for Wireless Backhaul Network Optimization

This chapter introduces several new techniques to improve the performance of a point-to-point Wireless Backhaul Network (WBN) in order to approach carrier-grade quality for this increasingly popular networking solution. More specifically, the author particularly targets two areas of WBNs:

- **WBN deployment** - In the network deployment phase, it is crucial to properly select node sites and to design an overall network topology that balances performance requirements, such as sufficient Quality of Service (QoS) and reliability, on the one hand, and cost efficiency on the other. This phase is characterized by more fundamental, long-term decisions and a more flexible choice between different design options. However, once taken, decisions can usually not be reverted without incurring significant additional costs.
- **WBN operation** - During the network operation phase, the optimization focus changes. Objectives have a rather limited time horizon (hours, minutes, seconds), cope with the restrictions imposed by the initial setup, and can only be optimized within defined boundaries. Since according actions are taken in higher frequency, they can also be reversed more easily and usually without major incremental expenses. However, according algorithms have to operate faster and can only process a limited amount of data.

In WBNs, the issue of network topology is a key lever in both phases. The following sections therefore present new initial network deployment techniques, particularly focusing on initial topological layout on the one hand and operational methods on the other. These operational methods, besides covering Topology Control (TC), focus on optimizing selected target

metrics, among them, user outage, fairness of resource allocation, battery usage of nodes with autarkic energy supply, and network coverage and thus significantly contribute to network performance optimization.

6.1 Fundamentals of Topology Control

TC constitutes an important instrument in the control of any kind of wireless network. In access networks, it determines Inter-Site Distance (ISD) and site coverage, in backhaul networks, it influences node density, network capacity, energy consumption, and robustness to node failures. In a Mobile Ad-hoc NETwork (MANET) or a Wireless Mesh Network (WMN), TC assures network connectivity and path redundancy, while at the same time controlling end-to-end packet delay. Typically, TC algorithms have a direct impact on latency and jitter statistics and are closely related to fundamental issues, such as power control, interference management, link calibration, Frame Error Rate (FER), routing, and many others. The following subsections first categorize TC approaches and present a definition of TC. Subsequently, state-of-the-art reference algorithms for TC are described.

6.1.1 Definition and Categorization of Topology Control

In previous research, TC subsumes a multitude of methods that aim at optimizing particular Key Performance Indicators (KPIs) of different classes of wireless networks. Over time and depending on the focus of the algorithms, a set of groups of TC methods have been developed.

In a substantial number of publications, particularly on large-scale wireless sensor and ad-hoc networks, authors apply clustering algorithms to introduce clusters and according head nodes for inter-cluster communications. This imposition of master/slave relations among nodes is similar to Bluetooth networks. However, it is technically more correct to label it hierarchy control instead of TC. Other researchers consider power control concepts as a form of TC. They typically define TC as

> "the process of dynamically changing the transmission range of nodes in order to maintain some property of the communication graph (e.g., connectivity), while reducing the energy consumed by node transceivers (which is strictly related to transmission range)" [San05].

Although both concepts have a strong mutual impact, they are usually employed in different contexts. More specifically, while power control usually considers the minimization of interference, Wideband Code Division Multiple Access (WCDMA) systems apply it to adjust the transmission power of user terminals (UTs), whereas systems operating in the Industrial, Scientific, and Medical (ISM) frequency bands, such as IEEE 802.11 or IEEE 802.15.4, employ TC to allow for spatial frequency reuse. In the former case, optimization is only performed locally (within cell boundaries). In the latter case, a global upper limit on transmission power restricts coverage to a bounded area, i.e., a static *one-size-fits-all* approach (resulting in possibly inappropriate power settings) is selected. Further, the vast majority of existing TC algorithms assume omni-directional single-radio nodes operating in a single-channel environment. As a consequence, the degrees of freedom as well as constraints differ substantially from those in point-to-point, multi-radio, multi-channel network setups.

The heterogeneity of TC methods is categorized and summarized in Figure 6.1. On the first level, the categorization differentiates between *homogenous* methods that apply the same parameter settings (e.g., transmission power, number of active radios) across all nodes in the network and *non-homogenous* methods that allow for individual settings for each node. For the second group, a further differentiation splits between methods for single-radio nodes in single-channel networks and multi-radio nodes in multi-channel environments. Due to the limited degrees of freedom resulting in less complexity in single-channel single-radio networks, it is possible to perform both centralized and distributed optimization with reasonable results. Location-based TC exploits the globally and permanently available knowledge of exact node positions. Distributed algorithms are implemented on each node and executed independently from each other. In direction-based approaches, nodes estimate the (relative) distance and direction of neighboring nodes. Finally, neighbor-based methods exploit logical neighborhood information, e.g., from routing table entries produced by proactive routing protocols. In multi-radio, multi-channel networks, methods for omni-directional and directional radio transmissions can be applied. In multi-channel multi-radio networks, optimization usually deals with an extended number of variables to be controlled. Therefore, distributed optimization quickly becomes very complex to control, thus favoring coordinated game theoretic approaches that maintain a minimum of reconciliation among nodes. Centralized algorithms are easier to implement

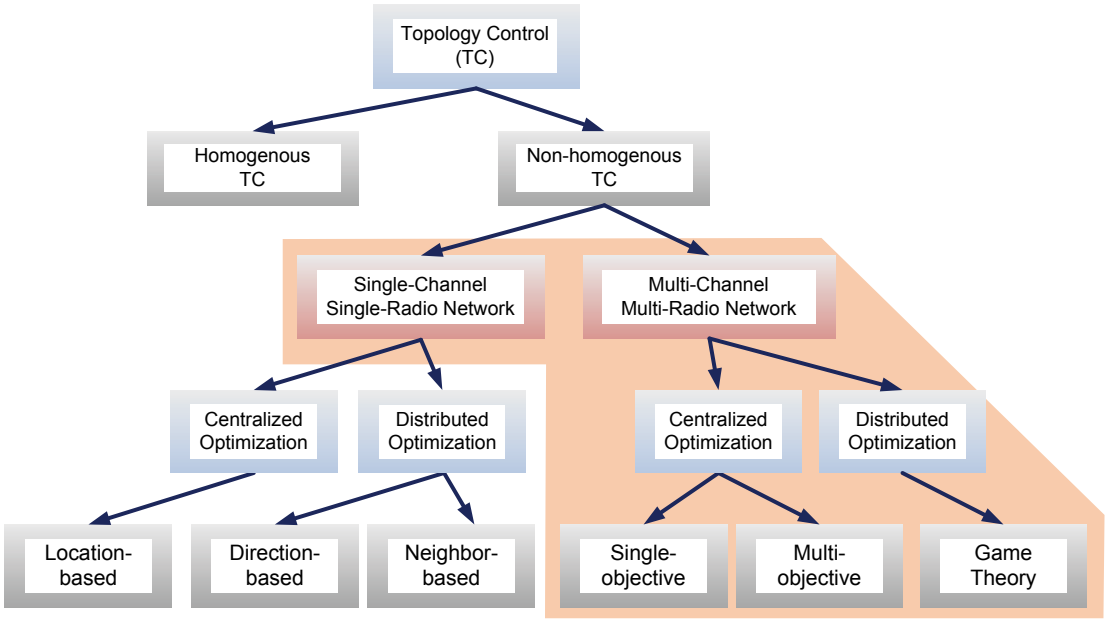


Figure 6.1: A taxonomy of topology control (based on [San05], extensions by the author are highlighted)

and available in two variants: single-objective and multi-objective.

Within the scope of this thesis, i.e., in the context of point-to-point WBNs with multi-radio nodes, the author employs an extended TC definition based on [San05]. TC comprises the set of algorithms that dynamically change

1. operational state (on / off),
2. transmission range / power,
3. transmission angle,
4. utilized radio channel, and
5. associated interface of a neighboring node

of each wireless interface (also referred to as a Network Interface Card (NIC)) in a point-to-point wireless network in order to maintain some property of the communication graph (e.g., connectivity, capacity), while reducing overall energy consumption of nodes.

Based on that definition, the process of finding an optimal topology inherently comprises two fundamental steps: First, relevant context data enabling a profound decision making needs to be available. Second, it requires the application of suitable search and optimization techniques that

allow for a formal mapping of desired network graph properties to multi-objective programs, multi-criteria objective functions, as well as feasible constraints in order to find the optimal solution. The author has elaborated on according context collection and distribution frameworks as well as on optimization and search algorithms in previous chapters.

From the definition of TC, the author explicitly excludes control and execution of routing decisions. In the context of this work, routing follows TC since it utilizes the result of TC to optimize the routing of packets through the backhaul network according to some metric. The following subsections give an overview on state-of-the-art TC algorithms. They first cover the category of centralized algorithms before switching to distributed methods. By presenting the most prominent and representative examples in terms of fundamental assumptions, objectives, and effectiveness, they aim at covering the abundance of TC approaches that have been developed (particularly in the area of wireless ad-hoc networks) in recent years.

6.1.2 Centralized Topology Control Algorithms

In case node positions are exactly known, global algorithms usually yield better results than distributed ones since they can search for an optimal topology, whereas distributed approaches need to apply some kind of heuristic. Therefore, location-based network topology forming generally employs centrally controlled, hierarchical optimization.

Centralized Topology Control Using On-Off Schemes

On-off schemes for nodes and their radio transceivers, respectively, belong to the simplest approaches to conserve energy in wireless multi-hop networks that, at the same time, maintain some minimum level of connectivity and capacity. However, since they usually require precise data on node coordinates, they are not well suited for distributed TC implementations. Xu et al. [XHE01] present an energy-conserving on-off scheme labeled Geographic Adaptive Fidelity (GAF) that exploits application and system information for switching off node radios for extended periods of time. Furthermore, location data of nodes is used to adaptively control node density and thus routing reliability. The overall goal is to preserve and, by means of energy conservation, temporally extend the so-called routing fidelity, in other words, to ensure connectivity in mobile ad-hoc networks. For this purpose, Xu et al. cluster nodes according to a rectangular overlay grid

shaped according to transmission ranges of nodes. In each cell, only one node needs to be active to reach a node in any neighboring rectangle, others can remain in a sleep state. In order to properly control the transition between *active* and *sleep* states, a third state labeled *discovery* as well as tunable time limits $T_\mathrm{a}, T_\mathrm{s}$, and T_d for the respective states are introduced. Nodes periodically return into discovery state and, based on a parametrized priority ranking for each cell, determine which of them will become active next, thus balancing the load and energy consumption among member nodes. Since the optimal scenario of one active node per cell is jeopardized by node mobility, Xu et al. suggest that sleeping time T_s is adapted according to the average expected time mobile nodes remain in the given cell, so that more frequent updates of the active node can be realized. A simulation-based quantitative evaluation featuring different energy, traffic, and mobility patterns compares the algorithm with a basic scenario where all nodes are switched on permanently. Analyzing a configuration of 50 nodes, Xu et al. demonstrate that node lifetime can be more than doubled for more than 30% of nodes and the average reduction of power consumption reaches a level of up to 60%, when compared to the standard scenario without any TC. Although GAF is compared to other routing protocols, it falls in the category of TC algorithms since Xu et al. evaluate it in conjunction with Ad-hoc On-Demand Distance Vector (AODV) routing and Dynamic Source Routing (DSR), i.e., routing protocols utilize the network topology generated by GAF.

Centralized Topology Control Using Specific Graphs

The properties of specific graphs can be exploited for forming a desired network topology. As an example, methods computing Minimum Spanning Trees (MSTs) emphasize a low number of edges for connecting the graph. This usually yields a low energy consumption since fewer links need to be maintained. Some other commonly desirable properties in wireless multi-hop networks include rather uniform and limited vertex degree (degree boundedness) as well as a rather regular and uniform topological structure, i.e., for each node, links to neighboring nodes should occur in rather uniform angular offsets to each other. *Delaunay triangulations* [For92] form a class of graphs that incorporate these characteristics to a high extent. A *Delaunay triangulation* G_D for a set $\mathcal{V}$ of vertices in a plane is a triangulation, such that no vertex $v \in \mathcal{V}$ is inside the circumcircle of any triangle.

Delaunay triangulations form a planar graph and maximize the minimum angle of all triangles in the planar triangulation.

Hu [Hu93] presents a centralized algorithm that, starting from a Delaunay triangulation, further optimizes graph topology with respect to two objectives: network connectivity and capacity. More specifically, utilizing maximum node degree $d_{\max}$, a node placement function $\mathbf{p} : \mathcal{V} \rightarrow \mathbb{R}^2$, and maximum transmission range $\delta(\pi_{\max})$ (where $\pi_{\max}$ is the maximum transmission power) as input parameters, the algorithm once passes through the following sequence of operations to obtain the final graph G:

1. Construct set $\mathcal{E}$ of edges for Delaunay triangulation $G_{\mathrm{D}} = (\mathcal{V}, \mathcal{E})$ of $\mathcal{V}$,
2. for all edges $(v_1, v_2) \in \mathcal{E}$: remove (v_1, v_2) from G_{D} if $\|\mathbf{p}(v_1) - \mathbf{p}(v_2)\| > \delta(\pi_{\max})$,
3. for all remaining edges (v_1, v_2) (sorted by length $\|\mathbf{p}(v_1) - \mathbf{p}(v_2)\|$, in descending order): remove (v_1, v_2), if it causes a degree $d(v) > d_{\max}$ at either v_1 or v_2,
4. for all edges $(v_1, v_2) \notin \mathcal{E}$ (sorted by length $\|\mathbf{p}(v_1) - \mathbf{p}(v_2)\|$, in ascending order): add (v_1, v_2) to G_{D} if it does not violate the degree constraint.

Performance evaluation is realized using a simulation environment of 100 nodes that are two-dimensionally Poisson distributed over a square of 100×100 base units. Results indicate that the algorithm can increase reliability and throughput when compared to setups without TC. The author also state a vertex degree $d_{\max} = 6$ as the best compromise between redundancy and interference levels. Finally, it is important to note that Hu's algorithm is initialized using a logical graph structure that in many places neglects the physical conditions (i.e., coverage and interference aspects of single-radio nodes with omni-directional antennas). Although the algorithm partially accounts for this, a differentiation between logical and physical neighbors is mandatory. Logical neighbors are these that explicitly belong to the overall topology and shall be considered by routing algorithms, whereas physical neighbors, although lying within the coverage area of current transmission power settings, are not explicitly part of the network topology and should be ignored by routing protocols.

Centralized Topology Control Using Transmission Power Assignments

The power or range assignment problem in wireless ad-hoc networks has been studied in many variants, utilizing different MST algorithms and featuring situation-specific constraints. For example, Kirousis et al. [KKKP00] included lower bounds for graph diameters, thereby effectively generating very dense graphs. Ramanathan et al. [RRH00] analyze the problem of continuously updating transmission power of nodes in wireless ad-hoc network so as to minimize energy consumption, while considering two constraints, namely connectivity and biconnectivity of the resulting graph. Therefore, they initially define the following constrained optimization problem (*Connected MinMax Power*) [RRH00]:

> Given a set $\mathcal{V}$ of network nodes, a placement function $\mathbf{p} : \mathcal{V} \to \mathbb{R}^3$ determining node positions, and the minimum transmission power function $\pi_{\min} : \mathbb{R} \to \mathbb{R}$ (giving the minimum power $\pi_{\min}$ needed to communicate over a distance of δ), find a per-node minimal assignment of transmission powers $\hat{\pi} : \mathcal{V} \to \mathbb{R}$, such that the resulting graph is connected and $\max_{v \in \mathcal{V}}(\hat{\pi}(v))$ is a minimum.

Subsequently, Ramanathan et al. extend the problem by requiring biconnectivity, i.e., there exist at least two vertex-disjoint paths between any pair of nodes (*Biconnectivity Augmentation with MinMax Power*) [RRH00]:

> Given a set $\mathcal{V}$ of network nodes, a placement function $\mathbf{p} : \mathcal{V} \to \mathbb{R}^3$, the minimum transmission power function $\pi_{\min}$ and an initial assignment of transmission powers π_0, find a per-node minimal set of power increases $\Delta_\pi(v)$, such that the resulting graph is biconnected and $\max_{v \in \mathcal{V}}(\hat{\pi}(v) + \Delta_\pi(v))$ is a minimum.

Ramanathan et al. argue for a mixed optimization strategy that minimizes the maximum transmission power instead of the total power consumption over all nodes since batteries (and their respective capacity) do not comprise a shared but rather an individual resource that cannot be redistributed among nodes. Moreover, they propose a "per-node minimum" of transmission power, meaning that no node can reduce transmission power any further without impairing (bi-)connectivity.

For solving the problem, two cases are distinguished, multi-hop networks with stationary nodes and networks formed of mobile nodes. For the former, Ramanathan et al. present two centrally executed optimization techniques for which they also prove general optimality conditions. The greedy algorithm labeled *CONNECT*, cf. Figure 6.2, starts by assuming each node to form a single cluster. Node pairs are sorted in non-decreasing order with respect to their distance. Subsequently, per iteration, transmission power of the selected pair (or the respective clusters they are part of) is increased until they become connected, given they have not been connected through previous iterations. This process is stopped once all nodes belong to a single cluster, i.e., once all network nodes are connected with each other. Finally, possible loops are eliminated.

Algorithm 1: CONNECT-AUGMENT

Data: Set of network vertices $\mathcal{V}$, placement function $\mathbf{p}(v)$, minimum power function $\pi_{\min}(\delta)$
Result: Power assignment $\hat{\pi}(v)$ for each node $v \in \mathcal{V}$ sufficient for a connected graph
begin
  sort node pairs in non-decreasing order of distance;
  initialize $|\mathcal{V}|$ clusters, one per node;
  foreach (v_1, v_2) *in sorted order* **do**
    if cluster$(v_1) \neq$ cluster(v_2) **then**
      $\hat{\pi}(v_1) = \hat{\pi}(v_2) \geq \pi_{\min}(||\mathbf{p}(v_1) - \mathbf{p}(v_2)||)$;
      merge cluster(v_1) with cluster(v_2);
      if *number of clusters is 1* **then**
        break;
      end
    end
  end
end

Figure 6.2: TC algorithm CONNECT [RRH00]

The algorithm *BICONN-AUGMENT*, cf. Figure 6.3, for generating a biconnected network graph is initialized using the result of the previous algorithm. After biconnected components have been identified and node pairs not sharing an edge have again been selected in non-decreasing order with respect to their distance, these pairs are joined only if they belong to two different biconnected components. The algorithm is terminated once

Algorithm 2: BICONN-AUGMENT

Data: Set of network vertices $\mathcal{V}$, placement function $\mathbf{p}(v)$, minimum power function $\pi_{\min}(\delta)$, initial power assignment $\pi_0(v)$ resulting in a connected graph

Result: Power assignment $\hat{\pi}(v)$ for each node $v \in \mathcal{V}$ resulting in a biconnected graph

```
begin
    sort node pairs in non-decreasing order of distance;
    compute graph G resulting from V, p, and π̂(v);
    foreach (v1, v2) in sorted order do
        if biconnectedComponent(v1) ≠ biconnectedComponent(v2)
        then
            π̂(v1) = max(π̂(v1), πmin(||p(v1) − p(v2)||));
            π̂(v2) = max(π̂(v2), πmin(||p(v1) − p(v2)||));
            add edge (v1, v2) to G;
        end
    end
end
```

Figure 6.3: TC algorithm BICONN-AUGMENT [RRH00]

the network is biconnected and per-node minimum power levels have been established.

The assignment of transmission power or range to nodes is among the most widely studied items in topology control. While some researchers aim at generating highly connected graphs [KKKP00], others argue for sparser communication graphs due to performance degradations related to 5interference [GK00]. However, both groups restrict their analyses to networks operating with single-radio nodes using the same channel.

6.1.3 Distributed Topology Control Algorithms

In many scenarios, exact node locations are not available on a global level for various reasons, e.g., node mobility. In such cases, distributed TC algorithms frequently outperform hierarchical algorithms. Although only approximating the optimum solution, they often exhibit many practical advantages, such as fast convergence and high scalability.

Distributed Topology Control Using On-Off Schemes

If node density in a wireless multi-hop network surpasses a threshold, the need for all nodes to permanently switch on their radio transceivers becomes obsolete, i.e., full connectivity and coverage for packet forwarding can be maintained with only a fraction of available nodes.

In [CJBM02], the authors present an on-off TC protocol labeled *Span* that works between link and network layer, i.e., routing builds upon topology maintenance decisions of the protocol. *Span* aims at reducing energy consumption without significantly diminishing network capacity and connectivity. It implements the idea of selecting a sufficient number of so-called coordinator nodes that stay awake permanently and are responsible for forwarding traffic within the ad-hoc network. Non-coordinator nodes remain in a power-saving mode. They only send and receive their own packets and periodically become active to check whether to become coordinator. The algorithm has to assure that the following goals are achieved: selection of a sufficiently high number of coordinators for full connectivity while at the same time limiting their number in order to increase node and network lifetime, rotation of coordinator task for a fair balancing of energy consumption, and fully distributed, local decision making. The core of the algorithm consists of a so-called *coordinator eligibility rule* that states that a non-coordinator node should become a coordinator if it discovers, analyzing only local broadcast messages, that two of its neighbors cannot reach each other directly or via one or two current coordinators.

For coordinator addition, nodes periodically leave the power-saving state and, after some back-off time has passed, check if an additional coordinator is required. If the eligibility rule applies, they announce themselves as additional coordinator. The back-off time of a node, which considers variables, such as number of neighboring nodes, energy level, or round trip delay of a probing packet, and also incorporates a random component, serves as a means of temporal coordination to prevent simultaneous announcements. In order to equally balance the coordinator task, nodes can also withdraw from it. In case the eligibility rule does not apply anymore (i.e., any two of its neighbors can reach each other directly or via one or two other coordinators) or some defined time span t_C since assuming the coordinator role has elapsed, it broadcasts its transition to a *tentative coordinator*, i.e., it will not join in subsequent coordinator selection processes.

For performance evaluation, Chen et al. configure a simulation of 120

mobile nodes equipped with IEEE 802.11g transceivers and assume a random way-point mobility model (cf. Section 4.1.1) within a 1000 m × 1000 m simulation area. On the one hand, results indicate that power consumption can be reduced by more than 50%, whereas capacity and latency are only marginally affected when compared to regular IEEE 802.11 performance (i.e., no topology control is applied). Further, lower bounds for coordinator densities are computed and compared to simulation results. On the other hand, Chen et al. admit that only limited additional power savings can be observed as node density increases. The involved control signaling of the algorithm needs to be counted as a further disadvantage of the algorithm.

Further TC mechanisms are presented in [WL99] and [WGS01], where a connected dominating set of nodes are selected to perform forwarding within a wireless ad-hoc network. A vertex subset of a graph is called a dominating set if every vertex that is not a member of the subset is adjacent to at least one vertex in the subset. Non-member nodes can return to power-saving mode. Vertices of the connected dominating set can be selected according to different rules. In [WGS01], both a selection based on battery levels to extend node lifetime as well as a selection based on node degree to reduce the size of the dominating set are proposed. Membership in the dominating subset is rotated among nodes in order to balance traffic forwarding tasks and thus energy consumption. However, Wu et al. acknowledge that a considerable additional signaling overhead, accompanied by increased power consumption, results from these methods.

Distributed Topology Control Using Specific Graphs

The inherent characteristics of specific graphs, e.g., with respect to redundancy or capacity, can be exploited to form network topologies that exhibit desired qualities. For example, k-ary tree graphs provide full connectivity with least amount of edges and limited number of child nodes, whereas planar graphs avoid any crossings of edges when drawn in a plane.

Hu et al. [Hu93] develop a distributed TC algorithm using a Voronoi graph (Voronoi diagram, [For92]) to finally form a Delaunay triangulation G_{D}, i.e., a planar triangularized graph (cf. Section 6.1.2). Assuming a set $\mathcal{V}$ of vertices (nodes), maximum node degree $d_{\max}$, a node placement function $\mathbf{p} : \mathcal{V} \rightarrow \mathbb{R}^2$, and maximum transmission range $\delta(\pi_{\max})$, the algorithm is comprised of the following steps:

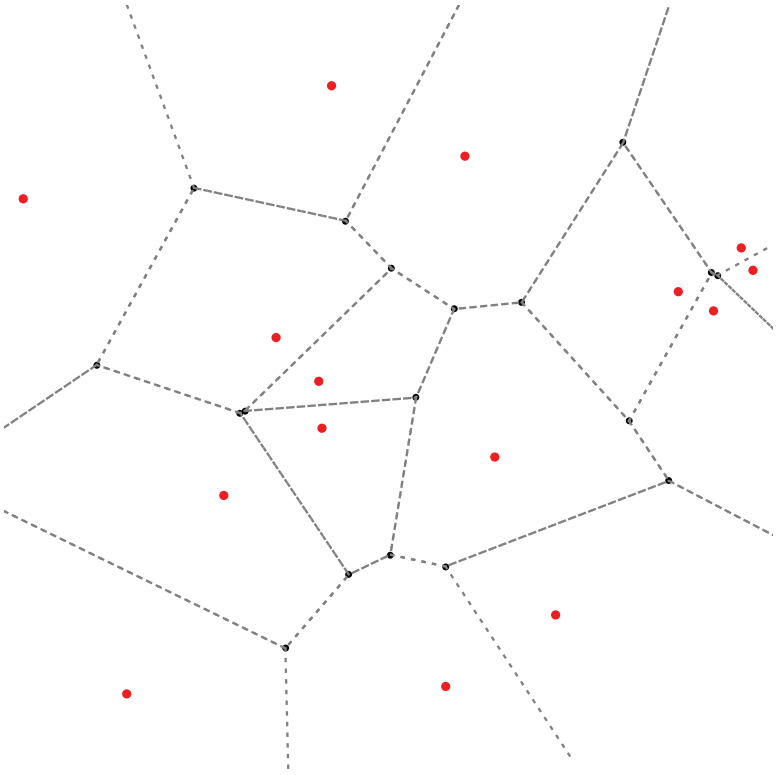
1. at specified time intervals, each node sends a maximum power broad-

cast message while in parallel listening to similar messages from adjacent nodes,

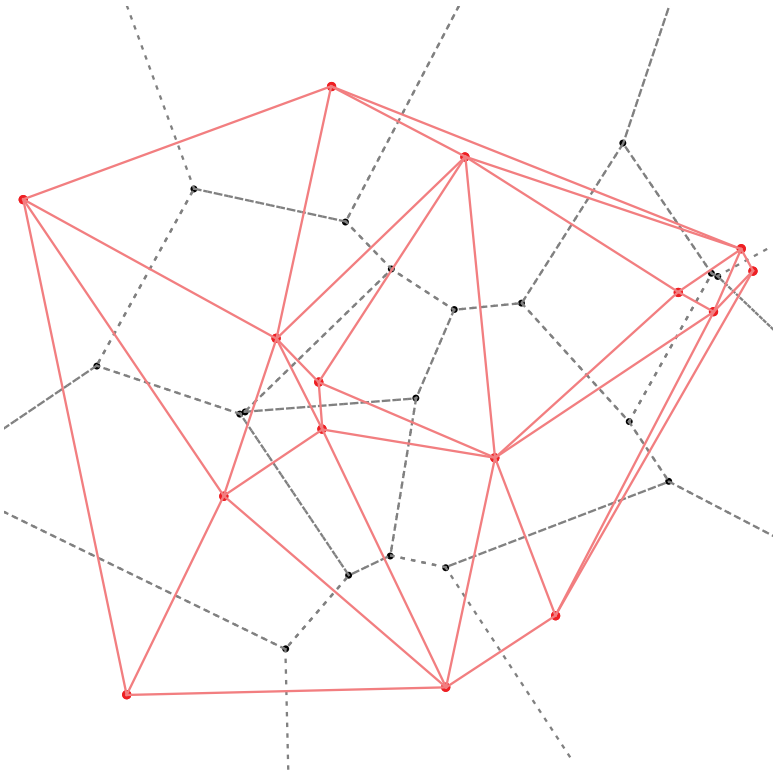
2. each node finds adjacent Delaunay triangulation neighbors within $\delta(\pi_{\max})$,
3. each node keeps only $d_{\max}$ shortest edges, informs other end nodes about deleted edges, and removes any edge deleted by a neighbor,
4. every active node (node with degree $d < d_{\max}$) starts a matching process:
   a) each active node sends a matching request to the nearest active neighbor; in turn, neighbors only acknowledge matching requests from their nearest active neighbor,
   b) if an active neighbor acknowledges the matching request, the edge is added to the topology,
   c) this process is repeated until either each node has a degree of $d = d_{\max}$ or $d < d_{\max}$ and no more active nodes within $\delta(\pi_{\max})$ are available for matching.

For step (2), which comprises the key operation of the algorithm, Hu et al. present an approach for detecting adjacent Delaunay triangulation neighbors. In a distributed manner, nodes calculate their respective Voronoi polygon which is formed by the set of perpendicular bisectors of the lines between the node and its neighbors and which demarcates the set of nearest points to a node. From this collection of Voronoi polygons (*Voronoi diagram*), the Delaunay graph or triangulation of the vertex set $\mathcal{V}$ can be deduced (cf. Figure 6.4). Although Hu et al. show the correctness of the algorithm, they do not include any quantitative evaluation with respect to effectiveness and overhead. For that matter, despite being very illustrative, the method requires a considerable amount of signaling and local processing since a high number of control messages need to be analyzed and exchanged between adjacent nodes.

MSTs provide a further example of optimized network topologies. For a given graph $G = (\mathcal{V}, \mathcal{E})$, a spanning tree $G_{\mathrm{T}} \subseteq G$ contains all vertices but only a subset of the edges of G, i.e., $G_{\mathrm{T}} = (\mathcal{V}, \mathcal{E}_{\mathrm{T}})$ and $\mathcal{E}_T \subseteq \mathcal{E}$, such that any pair of nodes is connected by exactly one path. Hence, G_{T} is free of any cycles. Since G can contain multiple trees, a tree becomes a MST if it minimizes the sum of its edge weights. Rodoplu et al. [RM99] introduce a distributed algorithm that computes, based on locally available information, among them GPS data, a so-called minimum power topology



(a) Voronoi polygons of a set of vertices



(b) Derived Delaunay triangulation

Figure 6.4: Duality of Voronoi diagram and Delaunay triangulation [Hu93]

that assures connectivity of all nodes to a selected master node. In other words, based on a transmission power range of $[0, \pi_{\max}]$ for each node, the set of possible edges is constructed, yielding G. Subsequently, they choose transmission power as edge weight and find the MST of G by minimizing the sum of transmission power over all selected edges.

Distributed Topology Control Using Sectoring

An obvious approach to achieve connectivity in wireless ad-hoc networks is to partition the range of possible transmission angles (usually 360° or 2π) into a set of sectors. Each of these sectors has a maximum angle of aperture ϑ in order to connect to at least one node per sector.

Wattenhofer et al. [WLBW01] have developed the Cone-Based Topology Control (CBTC) algorithm that realizes the described logic. The basic version of their algorithm defines an angle of aperture ϑ. Starting with a minimum level of transmission power, each network node starts sending broadcast messages and collecting respective acknowledgments from neighboring nodes. Using Angle of Arrival (AoA) techniques, nodes can determine the direction in which neighbors are located and check if at least one such neighbor is available in angular offsets of maximum ϑ. If gaps of more than ϑ occur, transmission power is incremented, a broadcast message is sent again, and received acknowledgments are analyzed. This process is repeated until angular gaps larger than ϑ do not appear any more or the maximum transmission power has been reached. Since the graph formed by this basic version of CBTC nevertheless contains numerous cycles, Wattenhofer et al. have suggested several extensions to the algorithm, e.g., [LHB+05]. The so-called *shrink-back* operation is initiated after the basic algorithm has completed. It performs transmission power reduction at nodes where offsets larger than ϑ are still present and hence send with maximum power. The basic idea is that transmission power is gradually decreased as long as the resulting sectors do not change, i.e., as long as the original cone coverage is maintained. In case of $\vartheta \leq \frac{2}{3}\pi$, the *asymmetric edge removal* constructs the largest symmetric set of edges contained in the original edge set resulting from the basic CBTC. Finally, *pairwise edge removal* further decreases transmission power by identifying cycles and eliminating the longest redundant pair of directed edges contained in each cycle. Summarizing, the graph resulting from the basic CBTC converges towards a tree graph when applying the proposed extensions. As a con-

sequence, less redundancy but also less interference can be observed, the latter being an important objective in single-channel wireless networks. In terms of effectiveness, it is shown that the CBTC algorithm including its extensions significantly reduces and, in case of node failures due to drained batteries, stabilizes transmission power levels when compared to their basic algorithm as well as to the *minimum power topology* protocol [RM99].

Distributed Topology Control Using Critical Neighbor Number

Algorithms striving to connect each node to a given number k of neighbors by setting transmission power accordingly belong to the class of neighbor-based protocols (compare Figure 6.1). They generally assume nodes to be stationary, to share a common maximum transmission power, and to use the same channel across the network. The key challenge is to derive an appropriate value for k_c ("critical neighbor number") that, depending on the given context, results in a strongly connected network with sufficiently high probability.

Based on theoretical work showing that, for a network formed of m nodes, k_c should be in the order of $\mathcal{O}(m)$ [XK04], Blough et al. [BLRS03] developed k-NEIGH, a fully distributed, asynchronous, and localized TC protocol. The following sequence of steps is performed in a timed manner. Starting with an announcement message sent by each node at full power, nodes compile a list of neighboring nodes including Time of Arrival (ToA)- and Received Signal Strength Indicator (RSSI)-based estimates of their respective distance. After a defined period of time has elapsed, each node broadcasts the list of k nearest neighbors at maximum power (if a node has less than k neighbors, it sends a list containing all neighboring nodes). Nodes can now compute a list of symmetric neighbors (i.e., a duplex channel is available) including their distance, the largest of which is being used for reducing transmission power to the appropriate level. If $k > k_c$, this results into a connected topology with high probability. Finally, an optional step of the algorithm ("pruning stage") further decreases transmission power of selected nodes without compromising connectivity requirements. More specifically, by analyzing the neighbor lists received from surrounding vertices, each node checks if more remote neighbors can be reached through an alternative multi-hop path passing through one or several closer neighbors. Using a large-scale simulation environment, Blough et al. show that k-NEIGH outperforms the widely-studied *CBTC* protocol in terms of energy

consumption and average node degree. They also present preferred values for k depending on the number of network nodes, thereby confirming the theoretically deduced computation complexity of $\mathcal{O}(m)$ for k_c.

In [RRH00], two distributed heuristics are suggested for TC in networks formed of mobile nodes. The *Local Information No Topology (LINT)* algorithm aims at maintaining a predefined vertex degree target d_{tg} (number of links to surrounding nodes) by exploiting information on the number of next hop nodes available from routing protocols. If the target value is exceeded ($d > d_{\text{tg}}$), transmission power is reduced in order to reduce coverage area. Contrary, if $d < d_{\text{tg}}$, transmission power is increased so that more nodes can be reached.

The *Local Information Link-State Topology (LILT)* algorithm on the other hand exploits information on global topology that is available from some routing protocols, e.g., Label-Switched Router (LSR), to mitigate network partitioning, which occasionally occurs if the *LINT* algorithm is used. *LILT* comprises two dedicated protocols, Neighbor Reduction Protocol (NRP) and Neighbor Addition Protocol (NAP). While NRP basically follows the *LINT* logic, thereby controlling node degree, NAP continuously checks whether the network is (bi-)connected or disconnected by analyzing according data from the routing protocol. In the latter case, all nodes set their transmission power to the maximum. In case of a biconnected network, no action is taken. If the topology is connected (but not biconnected), each node checks for the closest node that, if being removed, disconnects the network. Each node then sets a timer of value t, where t is a convex function of the measured distance. If after the expiration of t the network is still not biconnected, the node increases transmission power. This scheme allows for a limited network-wide coordination among nodes since closer nodes increase transmission power earlier, thus increasing the chance that more remote nodes do not need to change transmission power at all.

Since both *LINT* and *LILT* are heuristics, they do not guarantee an optimal power assignment. Moreover, the *LINT* protocol does not indicate how the value of d_{tg} relates to the probability of generating a connected graph, which is a core requirement in a WBN. Generally, all algorithms assume single-radio single-channel environments and only consider power control as a means for TC, thereby neglecting the beneficial impact of other possible variables, such as the availability of separately controllable radio interfaces using orthogonal channels.

6.2 Topology Control in Network Deployment

In the following, two novel iterative algorithms for multi-objective WBN deployment are presented: *Maximum Path Redundancy (MPR) Topology Control* and *Cycle-based Topology Control*.

6.2.1 Maximum Path Redundancy Topology Control

For the Quadratic Assignment Problem (QAP)-class problem of TC as outlined in Section 3.2, the author has designed two iterative approximation algorithms, *MPR* and *cycle-based TC*, that try to approximate the optimum solution for Equation (3.2). Both algorithms start with a random assignment Π_{rd} of nodes to coordinates. Subsequently, both algorithms try to establish edges in a way as to maximize the minimum node degree $\min_i(d(v_i)), v_i \in \mathcal{V}$, which can be as high as $d_{\max}$ in the optimum case. $d_{\max}$, the maximum node degree, which depicts the number of radio interfaces available at a backhaul node, is assumed to be the same for all nodes in $\mathcal{V}$.

The flowchart of Figure 6.5 depicts the sequence of operations performed in the *maximum path redundancy* algorithm. As input, it requires the set $\mathcal{V}$ of nodes, the set coordinates P ($|\mathcal{V}| = |P|$), the distance matrix $\mathbf{D}$, and the initial $m \times m$ reachability matrix $\mathbf{R} = (r_{i,j})$ depicting the set of edges under consideration for building the graph, i.e., $r_{i,j} = 1$ if v_j lies within the transmission range $\delta(\pi_{\max})$ of v_i and $r_{i,j} = 0$ otherwise. The algorithm enters the main loop by sorting nodes in ascending order of current degree d, which for all nodes equals zero in the first iteration. Nodes with the smallest occurring degree are selected for set S. If the current graph is d-regular (as is the case in the first iteration), S is formed by these nodes v_i that have the highest value $d_{\max} - d(v_i) - d_{\text{red}}(v_i)$, i.e., more descriptively, the number of additionally possible edges incident to v_i until $d(v_i) = d_{\max}$ (i.e., $d_{\max} - d(v_i)$) minus the number d_{red} of nodes reachable for (but not yet connected to) v_i (i.e., $d_{\text{red}}(v_i) = \sum_{j=1}^{m}(r_{i,j} - a_{i,j})$). Now, for each node $v_i \in \mathcal{S}$, the utility $U_{i,j} = -\alpha\,\delta_{i,j} + \beta\,\varphi_{k,i,j}$ ($\alpha, \beta \in \mathbb{R}^+$) is computed, where index j is restricted to nodes $v_j \in \mathcal{V}$ that a) lie within coverage of v_i (i.e., $\delta_{i,j} \leq \delta(\pi_{\max})$), b) are not connected to v_i yet (i.e., $a_{i,j} = 0$), and c) whose degree has not reached the maximum value yet ($d(v_j) < d_{\max}$). For $\varphi_{k,i,j}$, the algorithm selects the minimum in the set of angles between the potential edge $e_{i,j}$ connecting v_i and v_j and existing edges $e_{i,k}$, i.e., $k = 1, ..., m \,|\, a_{i,k} = 1$. Subsequently, for each $v_i \in \mathcal{S}$, edge $e_{i,k}$ is added, such that k has the highest utility, i.e.,

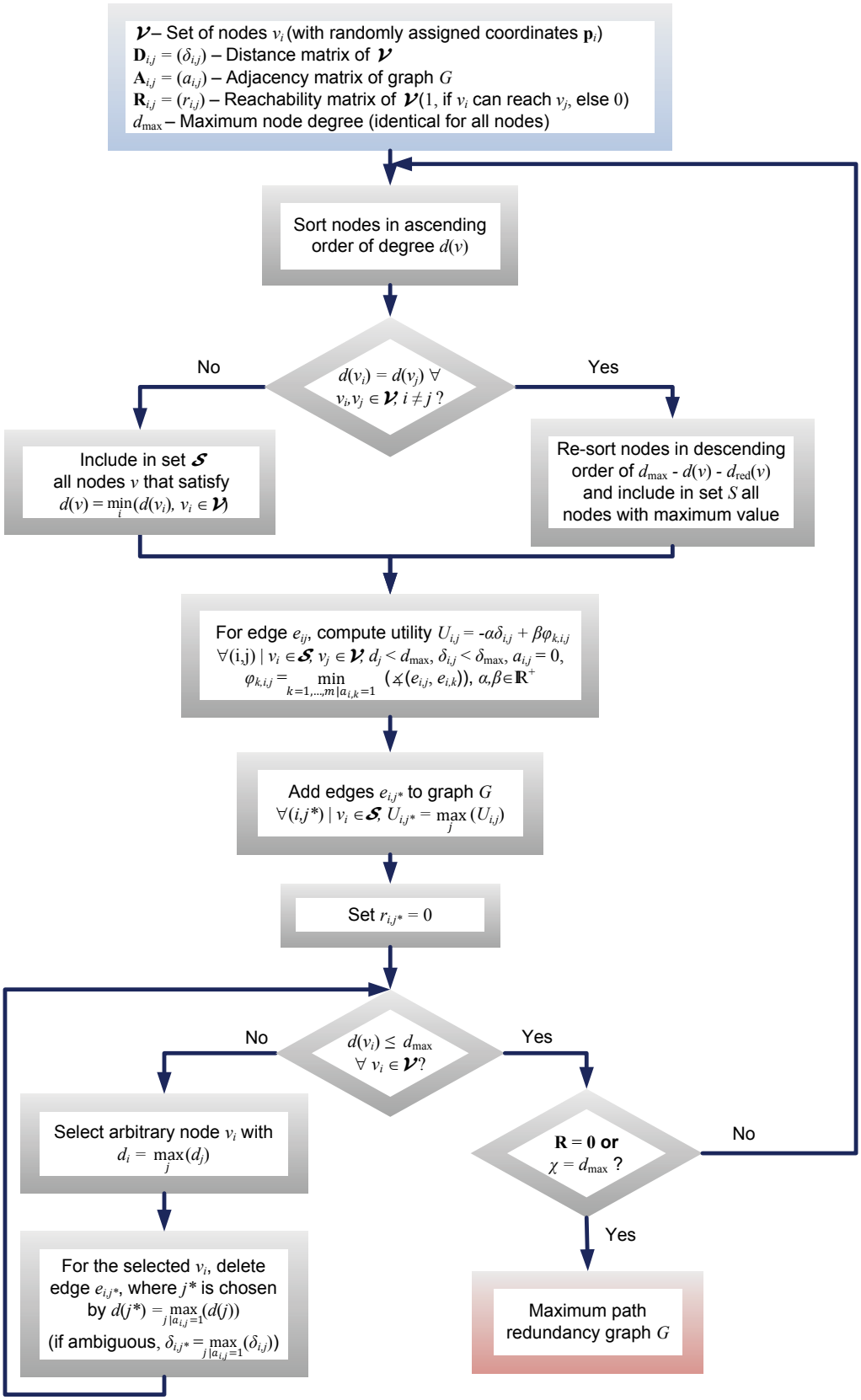


Figure 6.5: MPR TC approximation algorithm

$U_{i,k} = \max_j(U_{i,j})$, and set $r_{i,k} = 0$. This procedure of adding edges might result in nodes being incident to more edges than the maximum vertex degree $d_{\max}$, i.e., $\mathcal{S}^- = \{v_i \in \mathcal{V} \,|\, d(v_i) > d_{\max}\}$ is a non-empty set. In this case, edges are continuously deleted until S^- is empty. Here, in each step, the algorithm selects node v_i possessing the highest degree of all nodes in $\mathcal{S}^-$. From the set edges $e_{i,j}$ this node is incident to, the edge connecting to the node with the highest node degree is removed. If the edge selection is not unambiguous, then, from these edges, the one that connects to the node having the largest distance to v_i is selected. Once $d(v_i) \leq d_{\max} \; \forall \; v_i \in \mathcal{V}$, the algorithm checks if all possible edges have been considered ($\mathbf{R} = \mathbf{0}$) or $\chi = d_{\max}$. If either is true, it exits and obtains the final graph G, otherwise, the algorithm enters the main loop again.

6.2.2 Cycle-Based Topology Control

As a second novel TC approach, *cycle-based TC* has been developed. Figure 6.6 summarizes the steps of the algorithm. Conceptually, it is split into two phases. The first phase constructs a cyclic graph by analyzing the spatial distribution of nodes given by the placement function $\mathbf{p}$, i.e., a two-edge-connected spanning subgraph with minimum distance costs is created. This leaves $(d_{\max} - 2)$ vacant radio interfaces at each node. Subsequently, the second phase establishes additional edges so as to further increase edge connectivity χ, if possible up to $\chi = d_{\max}$.

More specifically, phase 1 starts by defining a center point $\mathbf{c}$ (e.g., center of gravity $\mathbf{p}_\mathrm{g} = \frac{1}{m}\sum_{i=1}^{m} \mathbf{p}_i$), which serves as the origin of a two-dimensional coordinate system. Nodes in $\mathcal{V}$ are sorted in increasing order of their angle $\arctan\left(\frac{p_2 - c_2}{p_1 - c_1}\right)$ with respect to $\mathbf{c}$, thus producing $\mathcal{V}_\mathrm{sort}$. A cyclic undirected graph G_c is constructed by applying a cyclic order on this set, i.e., the set $\mathcal{E}$ of edges becomes $\mathcal{E} = \{e_{i,i+1} \,|\, i = 1, \ldots, m-1\} \cup \{e_{m,1}\}$. If the length of an edge $e_{i,j} \in \mathcal{E}$ is larger than $\delta(\pi_{\max})$, an additional node is inserted in $\mathcal{V}_\mathrm{sort}$ at position $i+1$. Its coordinates are computed as $0.5(\mathbf{p}_i + \mathbf{p}_{i+1})$, and $e_{i,j}$ is "split" into two edges. This insertion process is repeated until the distance between all consecutive nodes in $\mathcal{V}_\mathrm{sort}$ is not larger than $\delta_{\max}$, yielding a two-edge-connected cyclic graph with m' ($m' \geq m$) nodes. At this stage, the algorithm switches into phase 2 by entering into a loop. Nodes $v_i \in \mathcal{V}_\mathrm{sort}$ are re-sorted in ascending order of their node degree $d(v)$. For nodes with the same node degree, their distance $\delta_i = \|\mathbf{p}_i - \mathbf{p}_\mathrm{g}\|$ to the center of gravity $\mathbf{p}_\mathrm{g}$, is applied as a secondary sorting criterion. In each iteration of the loop,

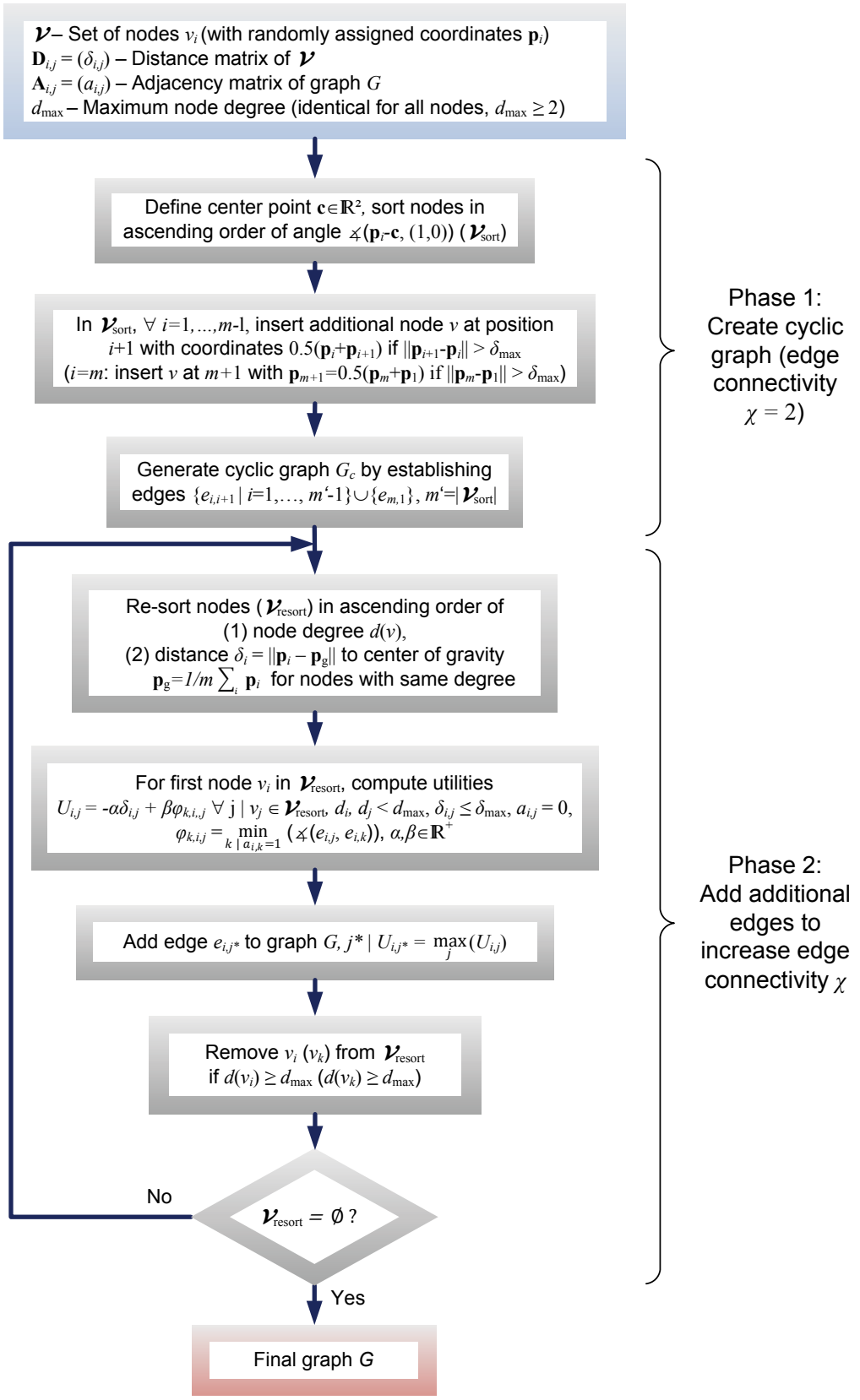


Figure 6.6: Topology control based on a cyclic graph

this yields a new, unambiguous set $\mathcal{V}_{\text{resort}}$. Subsequently, the algorithm, similar to *maximum path redundancy* TC, continues by computing utilities $U_{i,j} = -\alpha\,\delta_{i,j} + \beta\,\varphi_{k,i,j}$ $(\alpha, \beta \in \mathbb{R}^+)$ for the first node in $\mathcal{V}_{\text{resort}}$. Again, index j is restricted to nodes $v_j \in \mathcal{V}_{\text{resort}}$ that lie within coverage of v_i (i.e., $\delta_{i,j} \leq \delta(\pi_{\max})$) and that are not connected to v_i yet (i.e., $a_{i,j} = 0$). For $\varphi_{i,j}$, the algorithm selects the minimum in the set of angles between the potential edge $e_{i,j}$ connecting v_i and v_j and existing edges $e_{i,k}$, i.e., $k = 1, ..., m \,|\, a_{i,k} = 1$. Based on the utility values, the best edge e_{i,j^*} is selected and added to $\mathcal{E}$. If the degree of the according incidental nodes is exhausted (i.e., $d(v_i) \geq d_{\max}$ or $d(v_j^*) \geq d_{\max}$), the respective node is removed from $\mathcal{V}_{\text{resort}}$. In case $\mathcal{V}_{\text{resort}} \neq \varnothing$, the algorithm re-enters the described loop; otherwise, it terminates and obtains the final graph G.

In Section 7, an analytical evaluation is performed by formally deriving the approximation level of the *cycle-based* and *maximum path redundancy* TC algorithms.

6.3 Topology Control in Network Operations

The initial layout as well as the possibility for continuous adaptation of the backhaul topology of a coordinated wireless backhaul network constitute key design and operation parameters. They can be exploited to optimize a range of target metrics, among them backhaul network capacity, latency, jitter, overall Quality of Service (QoS), lifetime of a node with autarkic energy supply, and energy consumption and efficiency.

In the deployment phase, node locations and density as well as the number of available radio interfaces per node have to be chosen. Once the network has been set up, adaptations in the network topology, i.e., activation and deactivation of links between nodes or full release of a node from backhaul network have the highest impact on performance metrics.

Backhaul Topology Optimization (BTO) serves as a means to reap the benefits of such strategies. BTO algorithms optimize a given point-to-point backhaul topology according to a specified (multi-variate) objective function. During operation, when node placement cannot be easily changed anymore and the permutation function Π becomes obsolete, the manipulation of an adjacency matrix $\mathbf{A} = (a_{i_1,i_2})$ of a graph G is at the core of the described problem. The algorithms yield the optimal adjacency matrix $\hat{\mathbf{A}}$ of $T_O \subseteq G$ with respect to variables (a_{i_1,i_2}) and a given objective function. Thus, the BTO algorithm combines node on/off, interface on/off, as well as

coverage range schemes with available context information in an innovative manner.

6.3.1 Multi-Objective Backhaul Topology Optimization

Figure 6.7 depicts the fundamental steps of the BTO algorithm. The graph $G = (\mathcal{V}, \mathcal{E})$ represents the initial backhaul network that generally is $(N_i - 1)$-regular (i.e., each vertex connects to at least $N_i - 1$ other vertices), where N_i is the number of radio interfaces a node hosts (one of the interfaces usually serves as an access radio interface). In a first step, relevant input variables, depending on the structure of the optimization problem, are selected. Examples include energy consumption, battery level (in case of energy-autarkic nodes), edge capacities, or node centrality as a measure of importance of a vertex for the network as a whole (e.g., betweeness centrality or closeness centrality [Die06]). These input variables are used to compute the (initial) distance matrix $\mathbf{D} = (\delta_{i_1,i_2})$, for example,

$$\delta_{i_1,i_2} = \alpha_c \, c_{i_1,i_2}^{-\lambda_c} + \alpha_\beta \left(\beta_{i_1} + \beta_{i_2}\right)^{-\lambda_\beta}, \quad \alpha_c, \lambda_c, \alpha_\beta, \lambda_\beta \in \mathbb{R}^+, 1 \tag{6.1}$$

i.e., the link capacity matrix $\mathbf{C}$ and battery levels $\boldsymbol{\beta}$ determine inter-node distance. In graph-theoretic terms, δ_{i_1,i_2} can also be interpreted as the weight of the edge between nodes v_{i_1} and v_{i_2}.

Based on $\mathbf{D}$, the BTO algorithm can now use one of the optimization algorithms presented in Section 5.1 (e.g., Simulated Annealing) to compute the optimum graph T_O, whose set of edges $\mathcal{E}'$ only comprises a subset of all possible edges in $\mathcal{E}$. The decision of which edges (links) are to be switched off depends on the design of the optimization program (objective function and constraints) and the distance function (6.1). After the optimum graph has been found, the result is analyzed by evaluating the utility function U. In case the abort criterion is not met, the distance function is re-designed, e.g., by changing the weighting of the used input variables or substituting one variable for another. Subsequently, a new optimum graph as well as the updated value of U are computed.

This process is re-iterated until the abort criterion is met and a graph sufficiently close to the optimum is found. This flexibility of the BTO algorithm has to be counted as a major advantages. Distance and utility functions can be designed according to the requirements of the individual network configuration. Moreover, for computing the optimum graph, different optimization and search techniques can be employed.

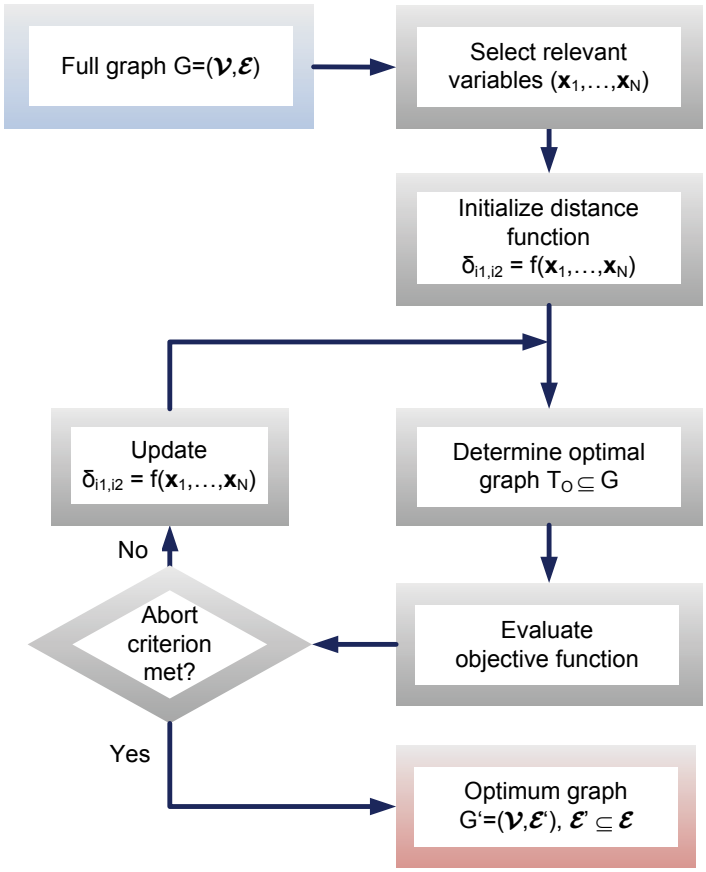


Figure 6.7: General algorithm for backhaul topology optimization [MCS13c]

In the following, the author presents an implementation suiting the requirements of coordinated, energy-autarkic WBNs. Using the graph description of the wireless backhaul network from Section 3.4, the author particularly focuses on the design of a multi-objective optimization problem that covers the most important aspects, while still being manageable in practice [MCS13b]. Therefore, the objective function as specified in Equation (4.28a) needs to be reworked. First, the QoS-related part $\alpha_Q \sum_{i_1,i_2} a_{i_1,i_2}\, \delta_{i_1,i_2}$ will focus on capacity, yielding $\alpha_Q \sum_{i_1,i_2} a_{i_1,i_2}\, \delta(c_{i_1,i_2})$, which denotes the sum of edge capacities. However, $s-t$-flow $c(\mathbf{A})$ is the more relevant quantity in the given context. Assuming that increasing $c(\mathbf{A})$ should have diminishing returns, the QoS-related summand in the objective function becomes $\alpha_Q\, e^{-c(\mathbf{A})}$. For the energy balance-related part, a similar logic applies, thus producing the updated objective function

$$\text{minimize} \quad \alpha_Q\, e^{-c(\mathbf{A})} + \alpha_E\, e^{-E(\mathbf{A})} - F. \tag{6.2}$$

Alternatively, rather than searching for the adjacency matrix $\mathbf{A}$ that results in the best energy balance, which is forced by the term $\alpha_E\, e^{-E(\mathbf{A})}$, the objective can focus on the battery level of nodes. In other words, the traffic flows in the network should preferably be switched through nodes with higher battery levels. This yields the following non-linear variation of the problem statement [CMS13]:

Program 6.3

$$\underset{\mathbf{A}=(a_{i_1,i_2})}{\text{minimize}} \quad \alpha_Q\, e^{-c(\mathbf{A})} + \alpha_E\, e^{-\boldsymbol{\beta}(\mathbf{A})} - F, \quad \alpha_Q, \alpha_E \in \mathbb{R}^+ \tag{6.3a}$$

$$\text{s.t.} \quad -\chi(\mathbf{A}) \leq -k \tag{6.3b}$$

$$\chi(\mathbf{A}) \leq d_{\max} \tag{6.3c}$$

$$\mathbf{A} - \mathbf{A}^{\mathrm{T}} = 0 \tag{6.3d}$$

$$-\sum_{i_2} a_{i_1,i_2} \leq -d_{i_1,\min} \;\; \forall\, i_1 \in \{1,\ldots,m\} \tag{6.3e}$$

$$\sum_{i_2} a_{i_1,i_2} \leq N_{i_1} \;\; \forall\, i_1 \in \{1,\ldots,m\} \tag{6.3f}$$

$$a_{i_1,i_2}\, \delta_{i_1,i_2} \leq \delta(\pi_{\text{out,max}}) \tag{6.3g}$$

$$\mathcal{P}_{i_1,i_2} \neq \emptyset \;\; \forall\, i_1, i_2 \in \{1,\ldots,m\},\; i_1 \neq i_2 \tag{6.3h}$$

$$\pi_{i,\text{out}}(x) \leq \pi_{\text{out,max}} \;\; \forall\, i \in \{1,\ldots,m\},\, x \in \{1,\ldots,N_i\} \tag{6.3i}$$

$$-\beta_i \leq 0 \;\; \forall\, i \in \{1,\ldots,m\} \tag{6.3j}$$

$$a_{i,i} = 0 \;\; \forall\, i \in \{1,\ldots,m\} \tag{6.3k}$$

$$a_{i_1,i_2} \in \{0,1\} \;\; \forall\, i_1, i_2 \in \{1,\ldots,m\} \tag{6.3l}$$

Given the input $G = (\mathcal{V}, \mathcal{E})$ and $\mathbf{D}$, the program finds the Pareto-optimal adjacency matrix $\hat{\mathbf{A}}$ (and the according graph $T_O \subseteq G$) that maximizes the objective function in Equation (6.3a). $c(\mathbf{A})$ denotes the capacity of the network, which is computed, according to [EK72], as the maximum $s-t-$flow from the gateway vertices $v_k \in \mathcal{V}_{\text{GW}}$ to all other vertices $v \in \mathcal{V} \backslash \mathcal{V}_{\text{GW}}$ of G. Constraints (6.3e) and (6.3f) restrict the degree of any vertex to a defined interval, whereas constraint (6.3d) guarantees that the adjacency matrix is symmetric, reflecting the undirected nature of the graph. Further, constraint (6.3h) assures that $\mathbf{A}$ describes a connected graph by forcing the set $\mathcal{P}_{i_1,i_2}$ of paths from vertex i_1 to vertex i_2 to be non-empty, i.e., there is at least one path from each vertex to any other vertex in G.

In the given form, optimization program 6.3 is of non-linear nature since both objective function and constraints contain non-linear functions. How-

ever, in order to allow for efficient solving in environments with real-time requirements, particularly in testbeds, operational systems, but also in simulation tools, the linearization of the program is essential. In particular, alternative expressions for the objective function (6.3a) and the constraint (6.3h) have to be found. Therefore, the adjacency matrix $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_{m-1}\ \mathbf{a}_m]^{\mathrm{T}}$ is converted to a vector $\mathbf{a}'$ by successively appending each row $\mathbf{a}_{i_1} = (a_{i_1,1}, \ldots, a_{i_1,m})$ of $\mathbf{A}$, i.e., $\mathbf{a}' = (\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_{m-1}\ \mathbf{a}_m)^{\mathrm{T}}$. For example, $\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$ is transformed to $\mathbf{a}' = (a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2})^{\mathrm{T}}$. With respect to the objective function, the capacity-related term is restrained to the links (edges) incident to gateway nodes. For the energy balance-related summand, a linear function of battery levels $\boldsymbol{\beta}$ is chosen. Further, several heuristics exist for linearizing constraint (6.3h) that reliably guarantee a connected graph. One option is to enforce a sufficiently high minimum degree d_{min} in constraint (6.3e). However, an approach with less effect on the solution is to enforce a minimum number $j_{\min}$ of links (edges) in the network, e.g., $\sum_{i_1,i_2} a_{i_1,i_2} \geq m - 1$. Finally, constraints (6.3c) and (6.3h) will be cut due to their negligible impact on the optimum solution. These modifications yield the (linear) Program 6.4.

Program 6.4

$$\underset{\mathbf{a}'}{\text{minimize}} \quad -\alpha_Q \sum_{k|v_k \in \mathcal{V}_{\mathrm{GW}}} \sum_{i|v_i \notin \mathcal{V}_{\mathrm{GW}}} \frac{a_{k,i}}{\rho_{k,i}}\, c_{k,i} - \alpha_E \sum_{i_1=1}^{m} \beta_{i_1} \sum_{i_2=1}^{m} a_{i_1,i_2},$$

$$\alpha_Q, \alpha_E \in \mathbb{R}^+ \tag{6.4a}$$

$$\text{s.t.} \quad -\chi(\mathbf{A}) \leq -k \tag{6.4b}$$

$$\mathbf{A} - \mathbf{A}^{\mathrm{T}} = 0 \tag{6.4c}$$

$$-\sum_{i_1,i_2 \in \{1,\ldots,m\}} a_{i_1,i_2} \leq 1 - m \tag{6.4d}$$

$$\sum_{i_2} a_{i_1,i_2} \leq N_{i_1} \ \ \forall\, i_1 \in \{1, \ldots, m\} \tag{6.4e}$$

$$a_{i_1,i_2}\, \delta_{i_1,i_2} \leq \delta(\pi_{\mathrm{out,max}}) \tag{6.4f}$$

$$\pi_{i,\mathrm{out}}(x) \leq \pi_{\mathrm{out,max}} \ \ \forall\, i \in \{1, \ldots, m\},\ x \in \{1, \ldots, N_i\} \tag{6.4g}$$

$$-\beta_i \leq 0 \ \ \forall\, i \in \{1, \ldots, m\} \tag{6.4h}$$

$$a_{i,i} = 0 \ \ \forall\, i \in \{1, \ldots, m\} \tag{6.4i}$$

$$a_{i_1,i_2} \in \{0, 1\} \ \ \forall\, i_1, i_2 \in \{1, \ldots, m\} \tag{6.4j}$$

Here, $c_{k,i}$ is the capacity of the edge formed between v_k and v_i. In order to flexibly incorporate different optimization goals, the objective function (6.4a) is designed as a weighted sum of selected parameters. The coefficients α_Q and α_E, besides normalizing performance variables to a common range, allow for a flexible weighting of the individual objectives. Program 6.4 is an exemplary implementation of the topology optimization as generically denoted in Section 3.2. In the particular example, the goal is twofold, first, to maximize the capacity of the network and second, to particularly increase vertex degree at these nodes where batteries exhibit high charging levels.

6.3.2 Extended Backhaul Topology Optimization

Connectivity Relaxations

Constraints (6.3h) ($\mathcal{P}_{i_1,i_2} \neq \emptyset$) and (6.4d) ($-\sum_{i_1,i_2} a_{i_1,i_2} \leq 1-m$), respectively, assure that the resulting sub-graph is a connected graph, i.e., that each node can be reached from any other node. Depending on the current characteristics of the network (e.g., current topological configuration, average distance between nodes, distribution of gateway nodes, battery status), this can be a considerably restrictive constraint. In order to allow for a reasonable relaxation, it is sensible to drop it and replace it with an alternative constraint that requires that each node v_i needs to be connected to at least one gateway node v_k. In light of a typical backhaul network structure, where gateway nodes are connected with each other through a some form of a wired connection, this is completely sufficient to ensure overall connectivity. Consequently, constraint (6.3h) can be replaced by the following expression:

$$\bigcup_k \mathcal{P}_{i,k} \neq \emptyset \quad \forall\, i | v_i \in \mathcal{V} \backslash \mathcal{V}_{\mathrm{GW}}, k | v_k \in \mathcal{V}_{\mathrm{GW}}. \tag{6.5}$$

In case a linear program is required, constraint (6.5) is transformed to

$$\sum_{i,k} a_{i,k} \leq -1 \quad \forall\, i \in \{1,\ldots,m\}, k | v_k \in \mathcal{V}_{\mathrm{GW}}. \tag{6.6}$$

Fairness Considerations

The objective function of the original problem statement (3.3) includes a summand F reflecting the fairness of resource allocation. The BTO algorithm has not considered this aspect yet, although it can impact fairness

significantly. For example, changes in backhaul network topology can result in throughput reductions for selected users. Fairness can be measured using appropriate metrics that compare the distribution of uplink and downlink throughput experienced by users within and across different service level agreement (SLA) groups. Readily available metrics for measuring the uniformity of such discrete distribution include mean, variance, min-max-ratio, or third standardized moment (usually referred to as skewness). Within the scope of this work, the author will use Jain's fairness index [JCH84] to introduce the notion of fair allocation of throughput among users. Jain's index has been widely used as a scale-independent, continuous, and particularly intuitive indicator to evaluate the fairness of throughput allocations in networks, e.g., in [BGY12]. For a given vector $\boldsymbol{\nu} \in \mathbb{R}_+^N$ of user data rates, where N is the number of users, Jain's fairness index $J : \mathbb{R}_+^N \rightarrow]0, 1]$ is computed according to

$$J(\boldsymbol{\nu}) = \frac{\left(\sum_{i=1}^{N} \nu_i\right)^2}{N \sum_{i=1}^{N} \nu_i^2}, \tag{6.7}$$

where $J(\boldsymbol{\nu}) \in]0, 1]$. Hence, higher values of J imply a fairer (i.e., more uniform) distribution of data rates among users, while lower quantities represent an increasingly unequal distribution.

In order to incorporate fairness considerations into the BTO program, F (cf. Equation (6.3a)) is replaced by $J(\boldsymbol{\nu})$, producing

$$\underset{\mathbf{A}=(a_{i_1,i_2})}{\text{minimize}} \quad \alpha_Q \, e^{-c(\mathbf{A})} \; + \; \alpha_E \, e^{-\boldsymbol{\beta}(\mathbf{A})} - \alpha_J J(\boldsymbol{\nu}), \quad \alpha_Q, \alpha_E \in \mathbb{R}^+, \tag{6.8}$$

where α_J denotes the weight (importance) that is given to the fairness objective. Of course, Equation (6.7) can also be used in Program (6.4).

6.4 Summary

In Chapter 6, the author has motivated the relevance and advantages of TC algorithms in the deployment and operation of point-to-point radio networks, discussed the benefits and limitations of centralized and distributed TC methods, and argued for hierarchical TC solutions to allow for an efficient implementation of QoS guarantees and SLAs. Based on this fundamental analysis, two novel TC algorithms for network deployment, *maximum path redundancy* and *cycle-based TC*, have been developed. While

the first one particularly focuses on maximizing path redundancy levels and capacity by optimizing edge connectivity of the network graph, the second algorithm has a priority on balancing node degree and hop distance between node pairs. Further, the author has evolved Program (3.3) (that has been further extended to Program (4.28)) into the final Backhaul Topology Optimization (BTO) problem for WBNs. This has yielded different variations of the original problem, including a linear version of the BTO. Moreover, the author has suggested a general approach for efficiently solving the BTO problem, thus allowing for execution in time-constrained environments as well as an extended variation that takes into account the distribution of resources among users (extended Backhaul Topology Optimization (eBTO)). In the following chapters, both an analytical and a quantitative, simulation-based evaluation are performed to quantify the gains achieved by the algorithms with respect to different KPIs.

Chapter 7

Analytical Evaluation

Chapters 7 and 8 perform an evaluation of the concepts and algorithms developed in the scope of this work. The chapters cover two distinct aspects:

- *General analytical evaluation* (Chapter 7) - This part derives exact analytical results with respect to performance guarantees of the designed algorithms. For example, optimum PhotoVoltaic (PV) panel surface areas and battery capacities are derived and expressed as a function of parameters known to the system designer. Further, for topology design and control algorithms, minimum values for graph connectivity levels are formally derived. Moreover, upper and lower bounds for energy consumption are determined. This also serves as the basis for further analytical benchmarking of network lifetime and backhaul connectivity.
- *Scenario-driven numerical evaluation* (Chapter 8) - In this part, the evaluation returns to the sample scenarios depicted in earlier chapters. By explicitly defining the relevant scenario parameters as well as specific restrictions within a flexible, model-based simulation framework, numerical results emphasize both the suitability of a Wireless Backhaul Network (WBN) for many deployment scenarios as well as the effectiveness of the algorithms developed in this work. Further, they are used to verify analytical results of Chapter 7.

7.1 Bounds for Power Consumption and Node Lifetime

In this section, the author analytically derives upper and lower bounds for energy consumption and node lifetime by analyzing minimum and maximum vertex connectivity situations.

7.1.1 Estimation of Power Consumption

As introduced in Section 3.4, $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_m)$ denotes the m-dimensional vector of instantaneous power consumption of the nodes v_i $(i = 1, \ldots, m)$ of G at a given time instant t. A common approach to model power consumption uses the state of the node, described by the status of its radio interfaces, cf. Table 8.1. Based on this, the author derives upper (π_u) and lower bounds (π_l) for power consumption of the entire network by estimating extreme values for power consumption per interface and the amount of active nodes (for the moment excluding other sources of power consumption). Accordingly, this yields the following expression for the upper bound π_u:

$$\begin{aligned} \pi_\mathrm{u} = \sum_{v_i \in \mathcal{V}} \pi_i &\leq \sum_{v_i \in \mathcal{V}} \left(\pi_\mathrm{on} + \sum_k \pi_{i,k} \right) \\ &\leq \sum_{v_i \in \mathcal{V}} (\pi_\mathrm{on} + d(v_i)\pi_\mathrm{Tx}) \\ &\leq m(\pi_\mathrm{on} + d_\mathrm{max}\pi_\mathrm{Tx}). \end{aligned} \tag{7.1}$$

Here, π_on and $\pi_{i,k} \in \{\pi_\mathrm{Tx}, \pi_\mathrm{Rx}, \pi_\mathrm{idle}\}$ depict base power consumption of node v_i and power consumption of radio interface k of v_i, respectively. These values are summed over all m nodes in $\mathcal{V}$ and, per node, over all interfaces. Since d_max is the upper bound for node degree $d(v_i)$ and the maximum power consumption of a single interface usually occurs in the Tx mode, $\pi_\mathrm{on} + d_\mathrm{max}\pi_\mathrm{Tx}$ constitutes an upper bound for a single node's power consumption. Similarly, a lower bound π_l for power consumption can be identified:

$$\begin{aligned} \pi_\mathrm{l} = \sum_{v_i \in \mathcal{V}} \pi_i &\geq \sum_{v_i \in \mathcal{V}} \left(\pi_\mathrm{on} + \sum_k \pi_{i,k} \right) \\ &\geq \sum_{v_i \in \mathcal{V}} (\pi_\mathrm{on} + d(v_i)\pi_\mathrm{idle}) \\ &\geq \tilde{m}(\pi_\mathrm{on} + \pi_\mathrm{idle}), \end{aligned} \tag{7.2}$$

where $\tilde{m}$ describes the minimum number of active nodes. Comparing upper and lower power consumption levels, the absolute difference amounts to

$$\pi_\mathrm{u} - \pi_\mathrm{l} = m(\pi_\mathrm{on} + d_\mathrm{max}\pi_\mathrm{Tx}) - \tilde{m}(\pi_\mathrm{on} + \pi_\mathrm{idle}), \tag{7.3}$$

$$= m(d_\mathrm{max}\pi_\mathrm{Tx} - \pi_\mathrm{idle}), \text{ if } m = \tilde{m}. \tag{7.4}$$

Further, the maximum relative saving in power consumption can be approximated according to

$$\frac{\pi_{\mathrm{u}} - \pi_{\mathrm{l}}}{\pi_{\mathrm{u}}} = 1 - \frac{\tilde{m}}{m} \frac{\pi_{\mathrm{on}} + \pi_{\mathrm{idle}}}{\pi_{\mathrm{on}} + d_{\max}\pi_{\mathrm{Tx}}}. \tag{7.5}$$

Inserting the power consumption figures of a backhaul node as presented in Section 4.1.3 and assuming $\tilde{m} = m$ and $d_{\max} = 3$ yields a maximum reduction in power consumption of approximately 54%.

7.1.2 Estimation of Node Lifetime

Continuing the analysis performed in Section 7.1.1, it is assumed that at least one interface of a node needs to remain active for receiving control information, but for most of the time remains in idle state. Although this assumes the extreme case of no user traffic, it serves as good basis for approximating maximal lifetime t_u. Accordingly, bounds for lifetime of a node v_i with discharging voltage U_{dis}, capacity Q_i, and battery charging level β_i can now be estimated. Computing

$$t_{\mathrm{l}} = \frac{\beta_i \, Q_i \, U_{\mathrm{dis}}}{\pi_{\mathrm{on}} + d_{\max}\pi_{\mathrm{Tx}}} \tag{7.6}$$

and

$$t_{\mathrm{u}} = \frac{\beta_i \, Q_i \, U_{\mathrm{dis}}}{\pi_{\mathrm{on}} + \pi_{\mathrm{idle}}}, \tag{7.7}$$

a lower (t_{l}) and upper (t_{u}) bound for remaining lifetime is derived, assuming that no recharging will occur (e.g., during night time operation). A schematic example for upper and lower bounds of lifetime is depicted in Figure 7.1.

The difference between maximal and minimum operation time $t_{\mathrm{u}} - t_{\mathrm{l}}$ cab be derived as follows:

$$t_{\mathrm{u}} - t_{\mathrm{l}} = \frac{\beta_i \, Q_i \, U_{\mathrm{dis}} \, (d_{\max}\pi_{\mathrm{Tx}} - \pi_{\mathrm{idle}})}{(\pi_{\mathrm{on}} + \pi_{\mathrm{idle}}) \, (\pi_{\mathrm{on}} + d_{\max}\pi_{\mathrm{Tx}})}, \tag{7.8}$$

which is equivalent to a relative improvement in operation time of

$$\frac{t_{\mathrm{u}} - t_{\mathrm{l}}}{t_{\mathrm{l}}} = \frac{d_{\max}\pi_{\mathrm{Tx}} - \pi_{\mathrm{idle}}}{\pi_{\mathrm{on}} + \pi_{\mathrm{idle}}}. \tag{7.9}$$

Inserting the power consumption figures presented in Section 4.1.3 and

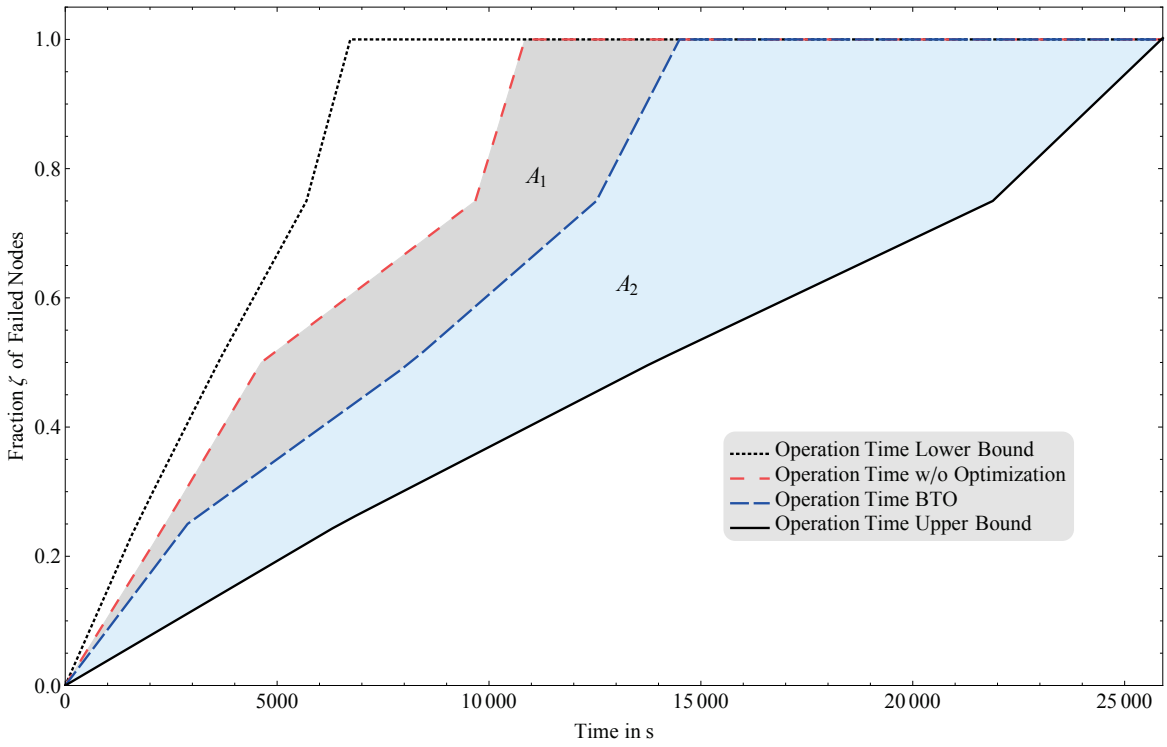


Figure 7.1: Illustration of the Lifetime Extension Measure (LEM) [MCS13a]

setting $d_{\max} = 3$ yields a difference of approximately 54% between t_u and t_l. For a network policy that requires two radio interfaces at a backhaul node to remain available at any time in order to handle incoming and outgoing traffic in parallel (as does the BTO algorithm), the difference reduces to 35%.

7.2 Dimensioning Battery Capacity and Photovoltaic Panel Area

When designing energy-autarkic systems, one of the key challenges becomes the estimation of the energy balance in the given environment. Required system components have to be dimensioned accordingly. In the case of wireless backhauling with solar-powered radio routers, the components of major interest are buffer battery capacity, the PV panel size, and, to a limited extent, the efficiency of a PV module. This work assumes the latter to be given. The major questions therefore are [MS14b]:

- What are sufficient dimensions for battery capacity and panel surface to assure sufficient reliability (e.g., $< 5\,\%$ probability of node failure)?

- What are longest periods without solar irradiation and how can battery capacity Q be chosen accordingly?
- What power yield (in other words, what panel area A) is required to fully (re-)charge battery and, in parallel, provide power for network operation, assuming typical local irradiation patterns?

7.2.1 Estimation of Battery Capacity

According to the author's work in [MS14a], required battery capacity is primarily determined by the power demand of a backhaul node. On the power consumption side, the author utilizes the model presented in Section 4.1.3 that assumes a power demand proportional to the number N_x of activated radio interfaces on top of a basic power demand π_0:

$$\pi = \pi_0 + N_\mathrm{x}\,\pi_\mathrm{x}. \tag{7.10}$$

For simplification, the author restricts the analysis to differentiate between day and night time operation of the backhaul network. For either time period, a number $N_\mathrm{x,d}$ and $N_\mathrm{x,n}$, respectively, of active radio interfaces is assumed, with $N_\mathrm{x,d} > N_\mathrm{x,n}$. Hence, power consumption follows

$$\pi_\mathrm{d} = \pi_0 + N_\mathrm{x,d}\,\pi_\mathrm{x}, \quad \text{for day time operation,} \tag{7.11a}$$
$$\pi_\mathrm{n} = \pi_0 + N_\mathrm{x,n}\,\pi_\mathrm{x}. \quad \text{for night time operation.} \tag{7.11b}$$

A further factor that determines required battery capacity is the length of time periods without any power supply from the PV module, during which energy supply falls back on battery power. The time from dusk until dawn on the shortest winter day, labeled discharging period Δt_dis, serves as a reference. During this time, the battery has to supply the amount of energy ΔE_dis, which in turn requests a capacity $\Delta Q^{(-)}$ of at least

$$\Delta Q^{(-)} = \frac{\Delta E_\mathrm{dis}}{U_\mathrm{dis}} \tag{7.12}$$

Similarly, ΔE_ch depicts the required energy for (re)charging the battery, yielding

$$\Delta Q^{(+)} = \frac{\Delta E_\mathrm{ch}}{U_\mathrm{ch}}. \tag{7.13}$$

Generally, output voltage suffers from two effects. First, it is considerably lower than input voltage, i.e., $U_\mathrm{ch} > U_\mathrm{dis}$. Second, output voltage declines

with falling charging level of a battery. To avoid deep discharge, which can irreversibly damage a battery, further discharging is usually discontinued at the so-called cut-off voltage. Depending on the battery type, cut-off voltage is reached at 50 to 80 % of discharging. Further, reaching full charging level usually requires a higher charge than what has been supplied during discharging, i.e., $\Delta Q^{(+)} > \Delta Q^{(-)}$. In order to account for these effects, batteries are characterized by their Wh-efficiency η_{Wh}, which summarizes the differences in both voltage and charge between discharging and charging periods:

$$\eta_{\mathrm{Wh}} = \frac{\Delta Q^{(-)}\, U_{\mathrm{dis}}}{\Delta Q^{(+)}\, U_{\mathrm{ch}}}. \tag{7.14}$$

Using Equations (7.11b) as well as (7.12) and assuming that cut-off voltage is reached at a charging level of β_{cut} (e.g., $\beta_{\mathrm{cut}} = 25\,\%$), required battery capacity Q_{max} can be estimated according to

$$\begin{aligned} Q_{\mathrm{max}} = \frac{1}{1-\beta_{\mathrm{cut}}}\,\Delta Q^{(-)} = \frac{1}{1-\beta_{\mathrm{cut}}}\,\frac{\Delta E_{\mathrm{dis}}}{U_{\mathrm{dis}}} &= \frac{1}{1-\beta_{\mathrm{cut}}}\,\frac{\pi_{\mathrm{n}}\,\Delta t_{\mathrm{dis}}}{U_{\mathrm{dis}}} \\ &= \frac{1}{1-\beta_{\mathrm{cut}}}\,\frac{(\pi_0 + N_{\mathrm{x,n}}\,\pi_{\mathrm{x}})\Delta t_{\mathrm{dis}}}{U_{\mathrm{dis}}}. \end{aligned} \tag{7.15}$$

7.2.2 Estimation of Panel Surface Area

In order to calculate the required panel area A, the analysis needs to identify the energy demand E during the time period t_{ch} of charging as well as typical yield from PV panels [MS14a]. The calculation of E is split into two parts: energy required for network operation and energy required for battery recharging. The following equation is obtained by substituting Equations (7.11a) and (7.13) for the respective parts:

$$\begin{aligned} E &= \pi_{\mathrm{d}}\,\Delta t_{\mathrm{ch}} + \Delta E_{\mathrm{ch}} \\ &= (\pi_0 + N_{\mathrm{x,d}}\,\pi_{\mathrm{x}})\,\Delta t_{\mathrm{ch}} + \Delta Q^{(+)}\,U_{\mathrm{ch}}. \end{aligned} \tag{7.16}$$

On the power supply side, daily yield Y_d in Wh can be computed as

$$Y_{\mathrm{d}} = \eta\, A \int_{t_1}^{t_2} I(t)\,dt = \eta\; A\overline{I}(t_2 - t_1) = \eta\, A\,\overline{H}, \tag{7.17}$$

where $I(t)$ and $\overline{I}$ are instantaneous and mean global irradiation in $\mathrm{W/m^2}$, respectively. $\overline{H}$ depicts the mean daily sum of global irradiation in $\mathrm{kWh/m^2}$. Inserting Equation (7.16) into Equation (7.17) as a substitute for Y and

solving for A yields

$$A = \frac{(\pi_0 + N_{\mathrm{x,d}}\,\pi_{\mathrm{x}})\Delta t_{\mathrm{ch}} + \Delta Q^{(+)} U_{\mathrm{ch}}}{\eta\,\overline{H}}. \tag{7.18}$$

Further, inserting Equations (7.14) and (7.12) produces

$$A = \frac{(\pi_0 + N_{\mathrm{x,d}}\,\pi_{\mathrm{x}})\Delta t_{\mathrm{ch}} + \frac{\Delta Q^{(-)}\,U_{\mathrm{dis}}}{\eta_{\mathrm{Wh}}}}{\eta\,\overline{H}}, \tag{7.19}$$

and finally

$$A = \frac{(\pi_0 + N_{\mathrm{x,d}}\,\pi_{\mathrm{x}})\Delta t_{\mathrm{ch}} + \frac{1}{\eta_{\mathrm{Wh}}}\,(\pi_0 + N_{\mathrm{x,n}}\,\pi_{\mathrm{x}})\Delta t_{\mathrm{dis}}}{\eta\,\overline{H}}. \tag{7.20}$$

The resulting panel size A of Equation (7.20) is a function of expected energy consumption at different times of the day and the respective node configuration. For night time consumption, a weighting factor depending on Wh-efficiency of the battery is included, reflecting the inefficiencies of storing energy in a battery. Further, daily global irradiation is a major variable. Here, different values are possible, such as median, mean, or selected p-quantiles. The smaller p is chosen, the less sensitive to irradiation fluctuations (e.g., due to unstable weather) power generation becomes. Finally, it is important to mention that the analysis does not account for aging effects of the battery.

7.3 Approximation Level of Topology Control Algorithms

In the following, the author formally derives the approximation level of the cycle-based TC algorithm when compared to the optimum solution of Equation (3.2). For the Maximum Path Redundancy (MPR) algorithm, an according chain of evidence can be derived.

Proposition 7.1

Under the assumption of a placement function $\mathbf{p} : \mathcal{V} \to P \subseteq \mathbb{R}^2$ *assigning coordinates to a set of nodes* $\mathcal{V}$ *and assuring that, for any particular node* $v_i \in \mathcal{V}$*, there are at least two other nodes* v_j*,* v_k *in the transmission range of* v_i*, i.e., more formally,* $\exists\ j, k \in \{1, 2, \ldots, m\}$ *so that* $\delta_{i,j}, \delta_{i,k} \leq$

$\delta(\pi_{max})\ \forall\ i \in \{1, 2, \ldots, m\}$, $j, k \neq i$, the cycle-based TC algorithm yields a $\frac{2}{d_{max}}$-approximation for the Quadratic Assignment Problem (QAP)-class problem of topology control in WBNs.

Proof: Due to the placement of nodes according to p and the chance for insertion of additional nodes, there is the possibility that each node is incident to at least two edges in an undirected graph G containing all nodes. Obviously, this is also guaranteed by the algorithm's initial phase that constructs a cyclic graph, which in turn has the inherent property of being two-edge-connected, i.e., there are exactly two paths from any arbitrarily selected node to any other node in G, namely, either clockwise or counter-clockwise. More precisely, for the $m \times m$ incidence matrix $\mathbf{\Psi}_{i,k} = (\psi_{i,k})$ of the interim cyclic graph G_c with m nodes and m edges, $\sum_{k=1}^{m} \psi_{i,k} = 2\ \forall\ i \in \{1, 2, \ldots, m\}$ holds. For estimating the approximation level γ of the algorithm, the author considers the objective function value U^* of the optimum solution (assignment) Π^* and the objective function value U^A of an arbitrary solution Π_{arb} of the algorithm, according to Equation (3.1). Now, the approximation level γ is defined as $\gamma := \frac{U^A}{U^*}$, with $\gamma \leq 1$ in case of a maximizing objective function. Assuming the most extreme scenario that the algorithm does not add any further edges to the interim cyclic graph (which per definition contains m edges), it can be deduced that

$$U^A = \sum_{i,j \in \{1,2,\ldots,m\}, i \neq j} a_{i,j}\ \delta_{\Pi_{\text{arb}}(i), \Pi_{\text{arb}}(j)} \geq m \min_{i,j}(\delta_{i,j}) \tag{7.21}$$

and

$$U^* = \sum_{i,j \in \{1,2,\ldots,m\}, i \neq j} a_{i,j}\ \delta_{\Pi^*(i), \Pi^*(j)} \leq \frac{m\ d_{\max}}{2} \max_{i,j}(\delta_{i,j}), \tag{7.22}$$

where Π^* is the optimum permutation and $\frac{m\ d_{\max}}{2}$ the upper bound for the number of edges in G when $d(v_i) = d_{\max}\ \forall\ i \in \{1, 2, \ldots, m\}$. $\mathbf{A} = (a_{i,j})$ denotes the $m \times m$ adjacency matrix of G. Hence, the (distance-dependent) approximation level of the algorithm becomes

$$\gamma \geq \frac{2\ \delta(\pi_{\min})}{d_{\max}\ \delta(\pi_{\max})}. \tag{7.23}$$

In the more generic case, the analysis does not include the distance factor, i.e., the author only considers the ratio of the number of edges established by the algorithm and the theoretic maximum. More specifically, U^A and U^* simplify to

$$U^A = \sum_{i,j \in \{1,2,\ldots,m\}, i \neq j} a_{i,j} \geq m \tag{7.24}$$

and

$$U^* = \sum_{i,j \in \{1,2,\ldots,m\}, i \neq j} a_{i,j} \leq \frac{m\ d_{\max}}{2}, \tag{7.25}$$

respectively, and γ becomes

$$\gamma \geq \frac{2}{d_{\max}}. \tag{7.26}$$

Most notably, neither the distance-dependent nor the distance-independent bound are

a function in either n (number of edges) or m (number of nodes). Rather, the only determinant of approximation quality is the maximum node degree $d_{\max}$. ■

For a value of $d_{\max} = 3$, the approximation level of the *cycle-based TC* algorithm becomes as good as $\gamma = \frac{2}{3}$. Due to technological constraints (e.g., power consumption of a backhaul node, number of orthogonal channels), the number of radio interfaces of a backhaul node usually does not exceed four or five. Therefore, in the worst case (cyclic graph with no further edges being established), approximation levels of the algorithm become $\gamma = \frac{1}{2}$ for $d_{\max} = 4$ or $\gamma = \frac{2}{5}$ for $d_{\max} = 5$. However, Monte Carlo simulations realizing a sufficiently large number of graph instances shows that the algorithm usually ($> 95\%$) generates a $(d_{\max} - 1)$-approximation. Section 8.3 presents the results of this quantitative evaluation of both algorithms for selected WBN instances, thus supporting the preceding formal analytical assessment.

7.4 Summary

Chapter 7 has presented an analytical evaluation of the algorithms developed in this thesis. More specifically, the author has derived upper and lower bounds for power consumption and derived according analytical expressions for minimum and maximum node lifetime. In the context of self-sufficient energy supply for network nodes, required minimum battery capacities and solar panel areas have been determined depending on battery cut-off level, charging and discharging voltages, equipment efficiency, as well as solar irradiation levels. Moreover, the symbolic approximation level for the cycle-based TC algorithm has been derived. Comparing the respective values of the defined objective function, the algorithm approaches the optimal solution by at least $\frac{2}{d_{\max}}$. In the following chapter, a simulation-based, numerical evaluation of the algorithms will help to confirm the results obtained in this chapter.

Chapter 8

Numerical Evaluation

The numerical analysis, which is carried out by means of a multi-domain simulation framework, accounts for several performance criteria. Besides deriving the requirements on battery capacity and PhotoVoltaic (PV) panel size for both the Tanzania and the Alpine scenario, the author also evaluates the impact of the topological deployment and operation algorithms developed in Chapter 6 on network performance. Eventually, the evaluation shall clarify the general suitability of an energy-autarkic coordinated Wireless Backhaul Network (WBN) for the defined (and similar) scenarios. Therefore, the analysis is guided by the following fundamental questions:

- What are the relevant parameters to configure the evaluation framework in the simulation tool?
- How can scenario-specific settings be reflected?
- How does equipment for PV energy supply have to be dimensioned for assuring reliable network operation?
- How can topology optimization in the deployment phase improve network performance?
- What are network performance figures of coordinated energy-autarkic WBNs under the given restrictions of non-permanent energy supply and fluctuating network capacity?
- How suitable are energy-autarkic, coordinated WBNs for selected geographical regions and weather conditions?
- How effective are the newly developed optimization algorithms, in particular with respect to trading off energy savings for Quality of Service (QoS)?

In order to answer these and related questions, the author first introduces the simulation framework in further detail.

8.1 Simulation Environment and Scenario Configuration

8.1.1 Simulation Environment

A quantitative evaluation of the algorithms described in Chapter 6 as well as the effectiveness of the designed multi-objective optimization programs is performed in a simulation framework for coordinated wireless backhaul networks implementing the evaluation methodology depicted in Chapter 3. The simulator is implemented using the *Wolfram Mathematica* technical computing environment [Wol14], which has been chosen due to its versatile capabilities in the areas of computer algebra and numerics. The software uses own symbolic language based on C++ and covers, among others, mathematical (symbolic) computation, numerics, algebraic manipulation, number theory, and graph computation. For example, *Mathematica* provides several techniques for solving linear and non-linear optimization problems both symbolically and numerically (e.g., interior point methods, genetic algorithms, or simulated annealing). Further, it implements numerous concepts from graph theory.

When designing a WBN reference deployment, numerous restrictions, such as upper and lower bounds for the number of gateway nodes m_{GW} and access nodes m_{AP} as well as for node density $\frac{m_{\mathrm{AP}}}{A_{\mathrm{cov}}}$ have to be considered. Moreover, minimum distance δ_{min} between two nodes, minimum number of neighboring nodes in coverage range of a node, sufficient centrality of each node, as well as a set of rules assuring a backhaul network with minimum edge connectivity χ and according redundancy and capacity describe further deployment constraints (graph-theoretical quantities are further explained in Appendix A). Modeling the operation of a WBN requires the incorporation of a diverse set of models and parameters, many of which have been presented and developed in Chapters 3 through 6. Tables 8.1 and 8.2 summarize the most important wireless networking models and deployment restrictions, most of them are based on [SZJ+08], [Int97], [3GP10], and [Irm08]. A detailed description of the modeling approach can also be found in [MCKS13]. In order to assure the comparability and representativeness of evaluation results, the algorithms to be analyzed were applied to different backhaul network configurations that, in conjunction with the scenarios to be defined in Section 8.1.2, cover various constellations of the parameter space. This includes location-specific and weather-

Table 8.1: Simulation models and parameters for WBN operation

Feature	**Implementation**
Maximum Tx power	Backhaul: 23 dBm (access: 20 dBm)
Frequency band	Backhaul: 5 GHz (access: 2.4 GHz)
System bandwidth	Backhaul: 40 MHz (access: 20 MHz)
Net channel capacity	According to IEEE 802.11n specification (backhaul: assuming orthogonal channels and advanced interference mitigation, access: accounting for CSMA characteristics in case of medium congestion)
User population	3-7 stationary users per cell (uniformly distributed over total area), varying mobile user population
Mobility model	Extended Gauss-Markov mobility (varying mean velocities)
Traffic models	Full buffer (stationary terminals), best effort (FTP) and voice (VoIP) (mobile terminals)
Traffic prioritization	(1) Voice services, (2) best effort, (3) user priority levels
Backhaul scheduling	Proportional fair within priority classes
Channel model	Typical large open space environment for Non-Line Of Sight (NLOS) and LOS conditions and 150 ns average rms delay spread [EG04, KMH+07]
Propagation model	Break point model with $\delta_0 = 100$ m, path loss $PL\,\vert_{dB} = 80 + 37.6 \log_{10}(\frac{\delta}{\delta_0})$, where δ is the distance from Tx to Rx in m
Handover scheme	Best SINR with *ping-pong* mitigation scheme
Energy consumption	State-dependent power consumption: $\{\pi_{\mathsf{Tx}}, \pi_{\mathsf{Rx}}, \pi_{\mathsf{idle}}\}$ for radio interfaces, $\{\pi_{\mathsf{on}}, \pi_{\mathsf{off}}\}$ for nodes [EV07]
Energy supply	75% of nodes solar supply (based on irradiation model, accounting for seasonal variations, weather, and other parameters of technical equipment), 25% of nodes connected to power grid
Battery capacity	80 Ah, with uniformly distributed initial charging levels (between 40 and 90%)
Simulation length	Variable, from several minutes up to a few days
Time increment	10, 100, or 1,000 ms

Table 8.2: Simulation models and parameters for WBN deployment

Feature	**Implementation**
Backhaul node density $\frac{m}{A_{\text{tot}}}$	Depending on topography and link capacity requirements
Number of gateways m_{GW}	Percentage of number m of backhaul nodes, $0.1 \leq \frac{m_{\text{GW}}}{m} \leq 0.5$
Number of access nodes m_{AP}	Multiple of number m of backhaul nodes, $2 \leq \frac{m_{\text{AP}}}{m} \leq 9$
Coverage ratio $\frac{A_{\text{cov}}}{A_{\text{tot}}}$	Service-dependent minimum threshold, $\frac{A_{\text{cov}}}{A_{\text{tot}}} \geq 0.95$, A_{tot}: total area, A_{cov}: total area covered by access nodes
Node centrality	Lower bounds for node degree (≥ 2) and betweenness centrality (≥ 3)
Minimum backhaul node distance δ_{min}	Function of backhaul node density and centrality
Grid area	Rectangular layout with varying side lengths, torus-like wrap around
Access network topology	Hexagonal grid with IEEE 802.11n access nodes, $500\,\text{m} \leq ISD \leq 900\,\text{m}$
Backhaul network topology	Predefined locations for IEEE 802.11n backhaul and gateway nodes (varying node densities)

dependent irradiation data for different time of day and time of the year, e.g., irradiation levels of overcast winter days, foggy fall days, or sunny summer days. Network configurations have been monitored on how Key Performance Indicator (KPI) (such as energy consumption, throughput, power consumption, yield of PV modules, battery levels, user outage) improve or deteriorate when the methods developed in this work are compared to state-of-the-art reference algorithms. According results are presented in the following sections of Chapter 8. The operative perseverance of backhaul nodes with autarkic energy supply, which can be undermined by either a completely discharged battery at the node itself or a failure of nodes lying on the path to a gateway node, was of particular interest since network and node availability are the most crucial aspects in the analysis of energy-autarkic WBNs.

8.1.2 Configuration of Considered Scenarios

The evaluation of WBN operation strategies depends on numerous input parameters from various areas. In order to allow for an appropriate classification and comparison of evaluation results, there is the necessity to define a basic set of of so-called scenarios that fix a certain portion of input parameters. Typical deployment areas for coordinated, energy-autarkic WBNs include Alpine regions with difficult topography in order to connect remote villages as well as rural areas in Sub-Saharan Africa, e.g., in Tanzania [AHT13]. Other possible deployment regions are rural areas in Scandinavian countries. While there are some commonalities in these scenarios, such as remoteness of installation sites, difficult energy supply, and rural communities, there are also some major differences, among them total amount of solar irradiation and temporal distribution of its availability. In this context, characterizing scenario parameters and their impact are listed in the following.

- Mobility patterns and service requests of users determine network utilization and congestion levels,
- time of day and weather conditions determine solar panel yield, node availability, and user activity,
- availability and reliability of power grid determine design choices for power supply,
- geographical location (particularly, geographical latitude) determines irradiation levels and solar power yield,
- network topology determines network capacity and QoS levels,
- node configuration (such as number of radio interfaces) determines network lifetime and path redundancy levels (edge connectivity of graph),
- availability of qualified field engineers determines to what extent autonomous networking functions (e.g., self-healing) are required.

Table 8.3 lists the aspects that characterize either deployment scenario. Most important factors with respect to solar energy supply using PV modules are geographical latitude as well as typical weather conditions. Moreover, sun declination also varies with seasonal periodicity. Here, Tanzania, due to its proximity to the equator, exhibits less fluctuations than Europe, where PV yield is substantially higher in summer months but very limited in winter.

Table 8.3: Characteristics of deployment scenarios

Feature	**Alpine scenario**	**Tanzania scenario**
Geographical latitude	$47°$ northern latitude	$6°$ southern latitude
User mobility patterns	Gauss-Markov mobility with low to medium user density	Gauss-Markov mobility with low to medium user density
Prevailing service classes	high prevalence of both voice and data services	high (medium) prevalence of voice (data) service
Weather conditions	temperate climate with varying sunshine levels	tropic climate with virtually daily sunshine
Power supply	largely from PV modules, most important nodes from power grid	from PV modules and (unreliable) power grid
PV panel surface area	$0.5 \ldots 2.5\,\mathrm{m}^2$	$0.5 \ldots 1.5\,\mathrm{m}^2$
PV module efficiency	0.25	0.25
Capacity of buffer battery	80 Ah	40 Ah
Fundamental network topology	MPR topology	MPR topology
Typical configuration of backhaul node	$d_{\mathrm{max}} = 3$	$d_{\mathrm{max}} = 3$
Qualified maintenance personnel	available, but rather expensive	hardly available

As a further paramount criterion, time of day and year have to be taken into account. Their direct effect on subscriber behavior (e.g., very few service requests at night) and on solar irradiation (e.g., lower sun declination angle in winter months) is depicted in Sections 7.2 and 8.2 and has to be considered carefully when interpreting results.

8.2 Evaluation of Battery Capacity and Solar Panel Size

In this section, the objective is to evaluate the requirements on battery capacity and PV panel size in order to allow for a stable and reliable operation of a backhaul node. Variations in energy supply and consumption are analyzed depending on important input parameters, such as geographical latitude, weather conditions (i.e., solar irradiation) over the course of several days, time of year, as well as network load [MCS14]. For each parameter constellation, sufficient dimensions for battery capacity and solar panel size are derived. Figures 8.1 and 8.2 depict global horizontal irradiation maps for the regions of the scenarios.

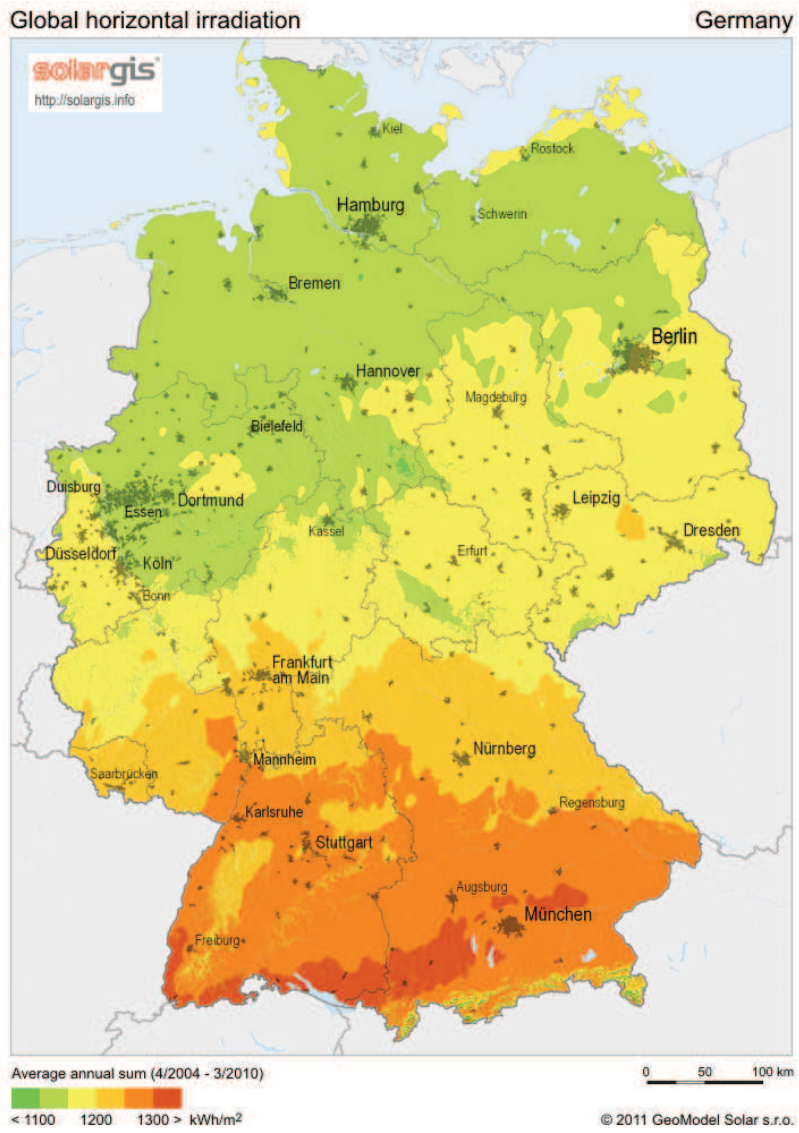


Figure 8.1: Map of global horizontal irradiation in Germany [Geo14]

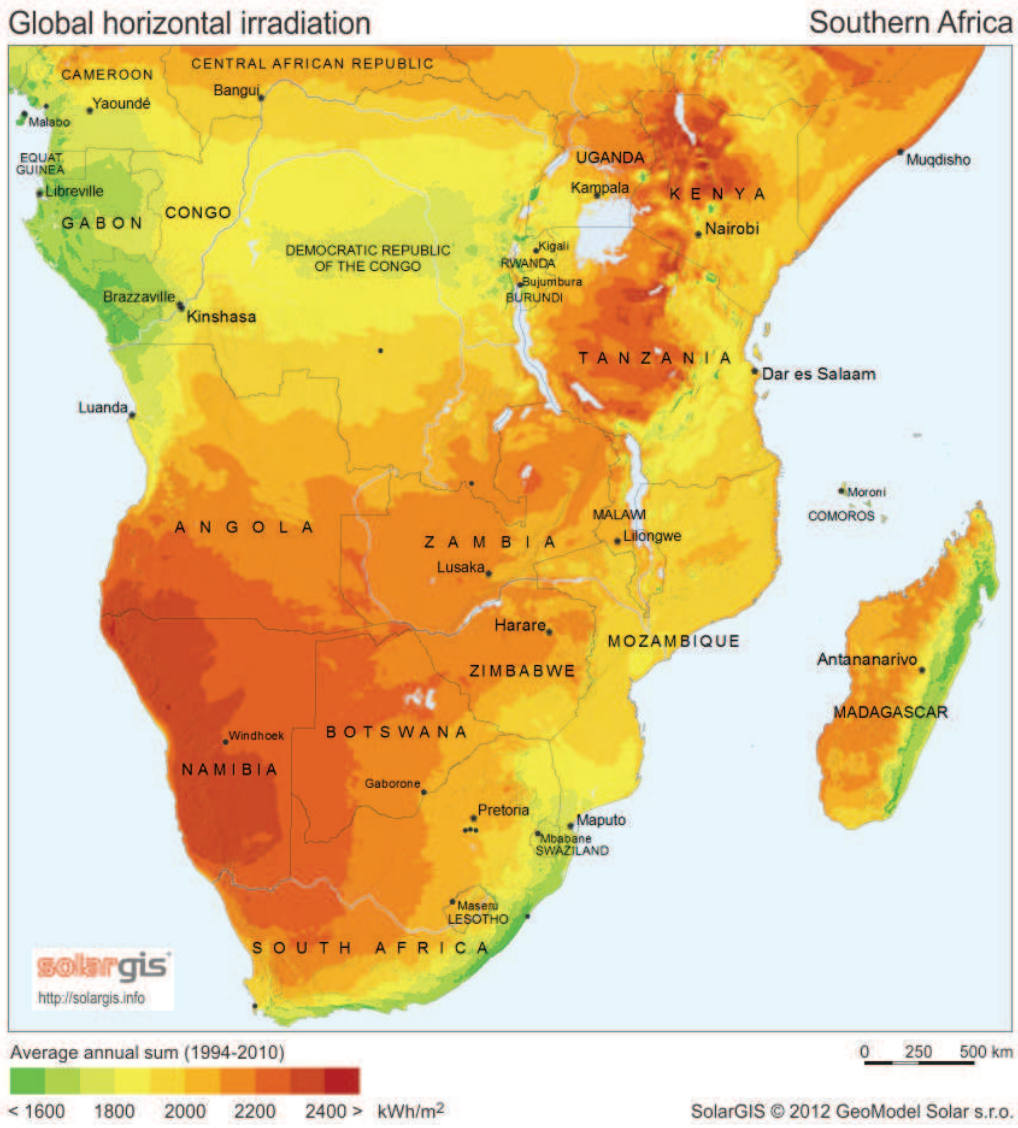
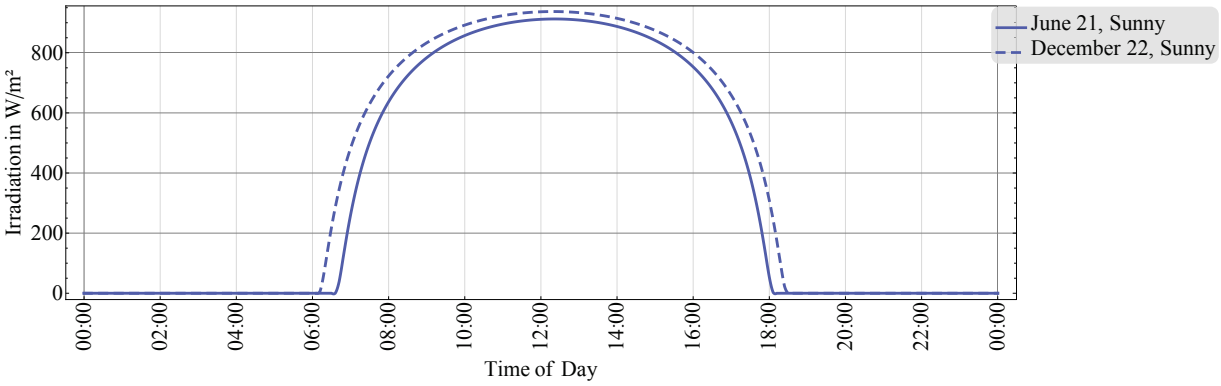


Figure 8.2: Map of global horizontal irradiation in Tanzania [Geo14]

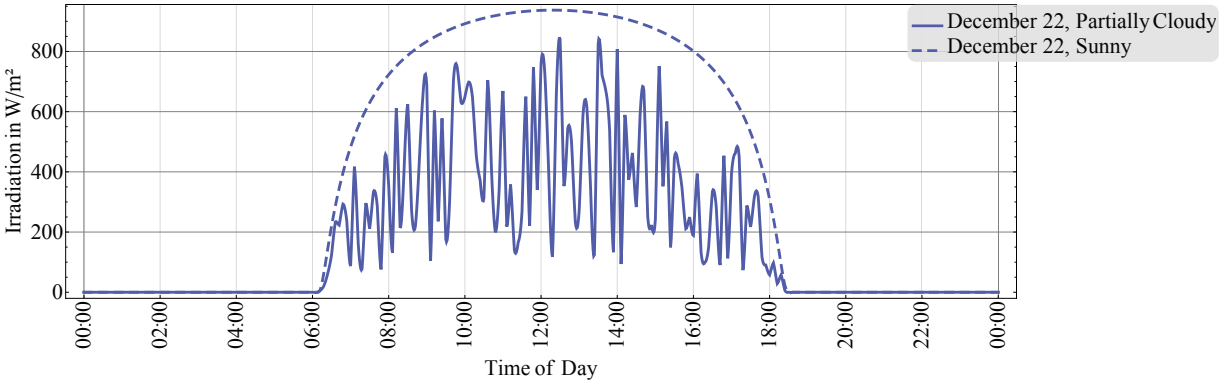
8.2.1 Tanzania Scenario

The availability of electric power from solar sources forms the starting point of this analysis. Therefore, the spatial and temporal variation of solar irradiation, which is composed of direct and diffuse components, needs to be quantified. Figure 8.3 depicts representative sample patterns of instantaneous solar irradiation $I(t)$ in $\frac{\mathrm{W}}{\mathrm{m}^2}$ for different weather conditions (sunny, partially cloudy, overcast) in Tanzania. More specifically, Figure 8.3(a) compares irradiation levels between winter and summer solstice days. The country's proximity to the equator (in the concrete case, a latitude of approximately $-7°$ is assumed) results in an almost negligible difference between summer (December 22) and winter (June 21) solstice. During sunny days, maximum irradiation is reached at around noon with levels of

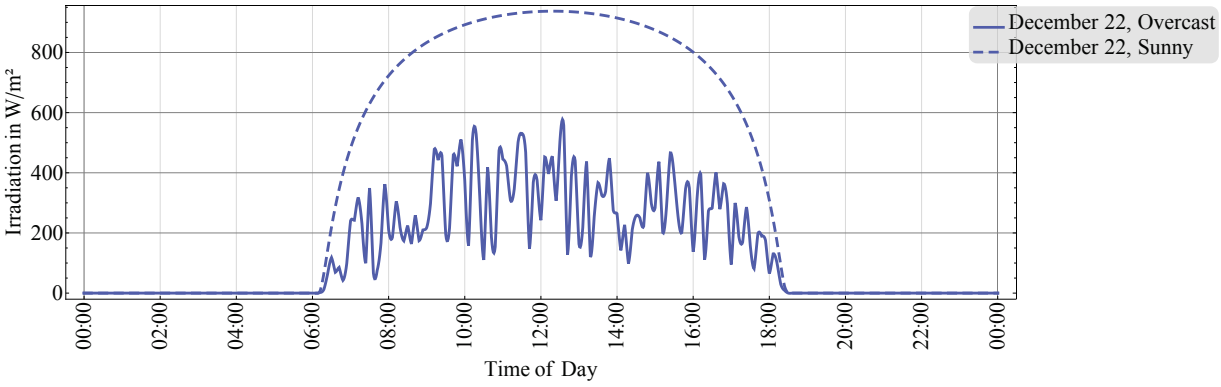
937 $\frac{\mathrm{W}}{\mathrm{m}^2}$ and 913 $\frac{\mathrm{W}}{\mathrm{m}^2}$, respectively [Joi12]. Sunrise and sunset take place at approximately 6 am and 6 pm. Figures 8.3(b) and 8.3(c) show solar irradiation characteristics for partially cloudy and overcast weather. Here, solar power yield has to cope with substantial fluctuations. However, even in the presence of compact clouds, diffuse irradiation can contribute a significant amount of energy. Irradiation patterns during rainy weather are similar to those during strong overclouding.



(a) Clear sky (solstice days)



(b) Partially cloudy sky



(c) Overcast Sky

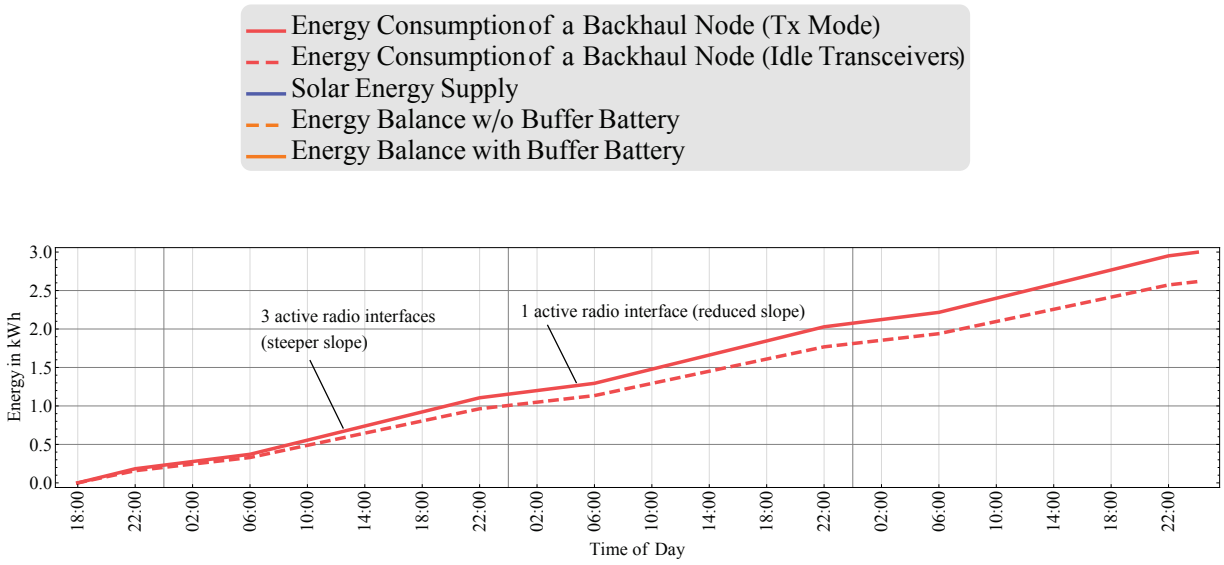
Figure 8.3: Intraday global irradiation levels in Tanzania

Reference Analysis with Continuously Sunny Weather

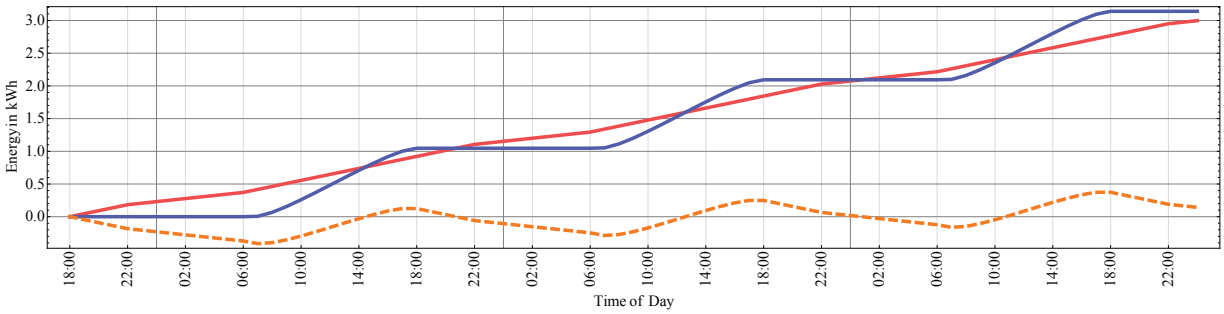
Generally, irradiation levels have to be mapped to energy supply, which in turn are compared to the demand of network nodes. For example, at noon, a solar panel of area $A = 0.5\,\mathrm{m}^2$ and exhibiting a typical PV module efficiency of $\eta = 0.25$ of modern mass produced PV modules yields a peak output power of $937\,\frac{\mathrm{W}}{\mathrm{m}^2} \cdot 0.5\,\mathrm{m}^2 \cdot 0.25 \approx 117\,\mathrm{W}$, which is more than sufficient for operating a radio node. However, in order to conduct a meaningful analysis, a sequence of several days needs to be considered. Figure 8.4(a) illustrates the cumulative energy consumption of a backhaul node in a WBN during a period of 78 hours, where the power consumption model as described by Equations (7.10), (7.11a), and (7.11b) is employed. Reflecting typical network capacity and redundancy requirements, $N_{\mathrm{x,n}} = 1$ active radio interface between 10 pm and 6 am and $N_{\mathrm{x,d}} = 3$ active radio interfaces during the rest of the time are assumed. This results in a steeper slope of the cumulative energy consumption curve during day time. While the solid line in Figure 8.4(a) depicts energy consumption for permanent Tx mode for all active radio interfaces, the dashed line assumes permanently idle radio interfaces. Actual energy consumption will hence lie in between these two bounds. Further analyses will only consider the (extreme) case of permanent Tx mode since the difference in power consumption is comparatively small. Further, as a baseline reference, the evaluation initially assumes a sequence of days with sunny weather, i.e., with maximum solar energy yield [MCS14]. The solid line in Figure 8.4(b) illustrates the resulting cumulative energy yield.

$$Y_{\mathrm{PV,cum}} = \eta\, A \int_{t_1}^{t_2} I(t)\, dt = \eta\, A\, \overline{H}, \tag{8.1}$$

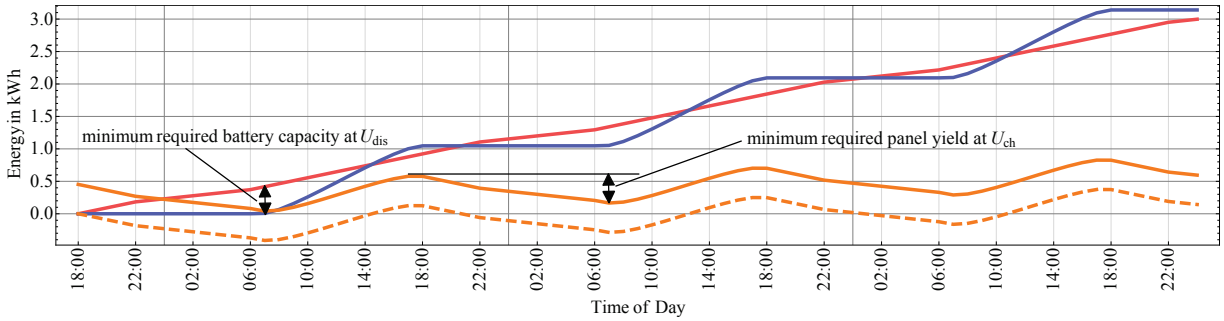
of a PV module with $A = 0.5\,\mathrm{m}^2$ and module efficiency $\eta = 0.25$, $(t_2 - t_1) = 78\,\mathrm{h}$, and $\overline{H} = \overline{I}(t_2 - t_1)$. The dashed line depicts the difference between energy supply and demand, in other words, the energy balance of a node. Obviously, this curve has a negative slope during darkness and a positive slope if power supply exceeds consumption, which is the case when the yield of a PV module is sufficiently high. In general, the objective is to appropriately dimension panel surface A and battery capacity Q to accommodate for the observed fluctuations. Particularly, the time period of declining energy balance shall be buffered by an according battery. Moreover, PV panel area needs to be chosen so as to yield enough power for



(a) Energy consumption



(b) Energy supply



(c) Energy balance

Figure 8.4: Cumulative energy supply, consumption, and balance of a backhaul node with autarkic power supply in Tanzania

both node operation (π) and recharging drained batteries (ΔE_{ch}). These aspects are illustrated in Figure 8.4(c), where the dashed line is lifted above the axis of abscissae, in the given example by $\Delta E_{\text{T1}} \approx 0.37\,\text{kWh}$. Assuming a discharging voltage of $U_{\text{dis}} = 12\,\text{V}$ and a cut-off battery level of $\beta_{\text{cut}} = 0.25$, required battery capacity Q_{T1} is computed according to Equation (7.15):

$$Q_{\text{T1}} = \frac{1}{1-\beta_{\text{cut}}} \frac{\Delta E_{\text{T1}}}{U_{\text{dis}}} \approx 41.2\,\text{Ah}.$$

Further, the yield of a PV panel shall satisfy the energy demand $E_{\text{op}} + E_{\text{ch}} = \int \pi(t)dt + \frac{\Delta E_{\text{T1}}}{\eta_{\text{Wh}}}$. The required panel area $A_{\text{T1,min}}$ is calculated according to

$$A_{\text{T1}} = \frac{\overline{\pi}\,\Delta t_{\text{ch}} + \frac{\Delta E_{\text{T1}}}{\eta_{\text{Wh}}}}{\eta\,\overline{H}}. \tag{7.19}$$

Assuming a Wh-efficiency of $\eta_{\text{Wh}} = 0.75$, an average power consumption of $\overline{\pi} = \pi_{\text{d}} = 45.2\,\text{W}$ for node operation, a daily sum of solar irradiation of $\overline{H}_{\text{sunny}} = 8.9\,\frac{\text{kWh}}{\text{m}^2}$, and $\Delta t_{\text{ch}} = 12\,\text{h}$, this results in a minimum panel size of $A_{\text{T1}} \approx 0.47\,\text{m}^2$. In other words, the panel area of $A = 0.5\,\text{m}^2$, as selected in the example of Figure 8.4, is sufficient for autarkic energy supply during sunny weather.

Sensitivity Analysis

Although sunny weather has a fair share in Tanzania throughout the year, periods of bad weather occur and a complete analysis hence has to consider a sequence of days with overcast or rainy weather. On average, the country enjoys a yearly sum of global irradiation of $H_y \approx 2350\,\frac{\text{kWh}}{\text{m}^2}$. However, in some regions, H_y can be as high as $2700\,\frac{\text{kWh}}{\text{m}^2}$ [Geo14]. Figure 8.5 depicts the mean daily sum of global irradiation $\overline{H}$ throughout the year. The *clear-sky-peak* line indicates the maximum energy per m^2 arriving on the (horizontal) ground in case of sunny weather. Since Tanzania is located in the southern hemisphere, lowest values occur in June. However, due to varying weather conditions and rainy seasons, the actual mean sum of horizontal irradiation deviates significantly, as depicted by the *mean – horizontal* curve. It exhibits a minimum in May, while the maximum is reached in the month of October. Assuming a fixed panel surface (no sun tracking), the annual mean value of solar irradiation is maximized for a surface tilt of 9° and facing to the north, as illustrated by the the *mean*

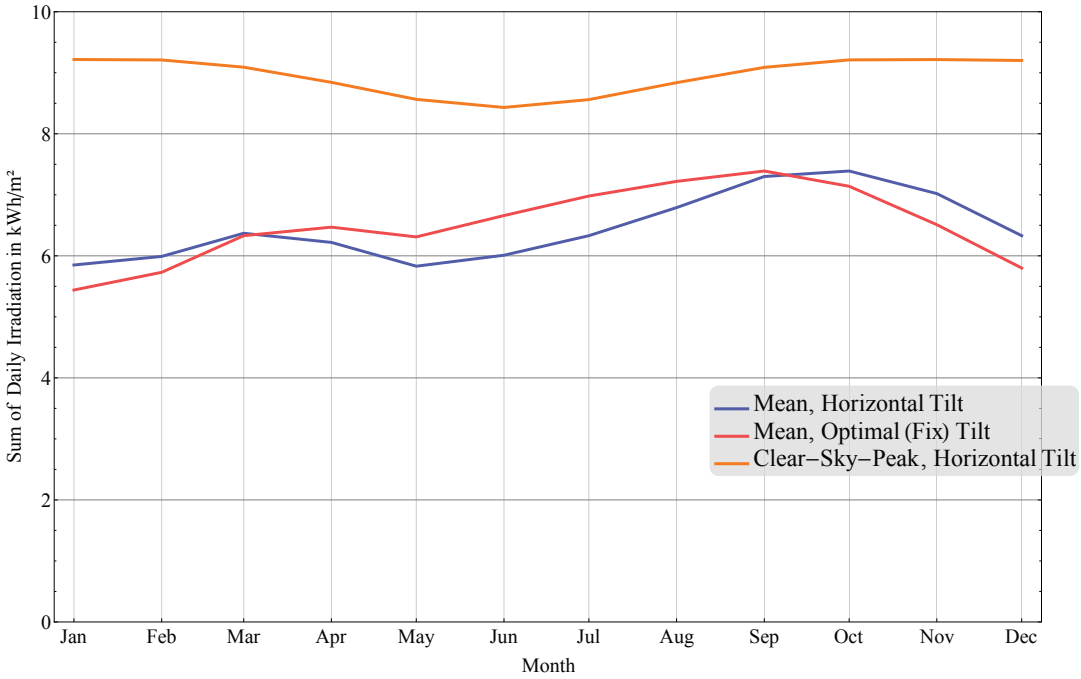
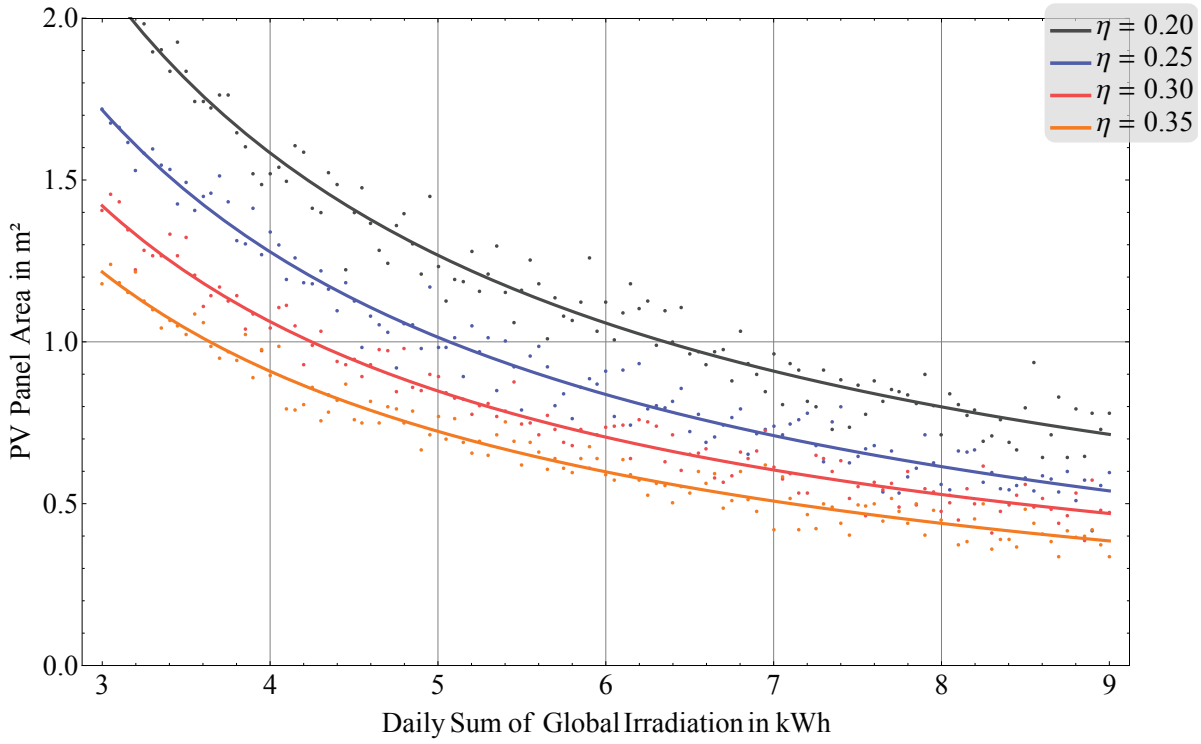


Figure 8.5: Sum of daily global irradiation in Tanzania

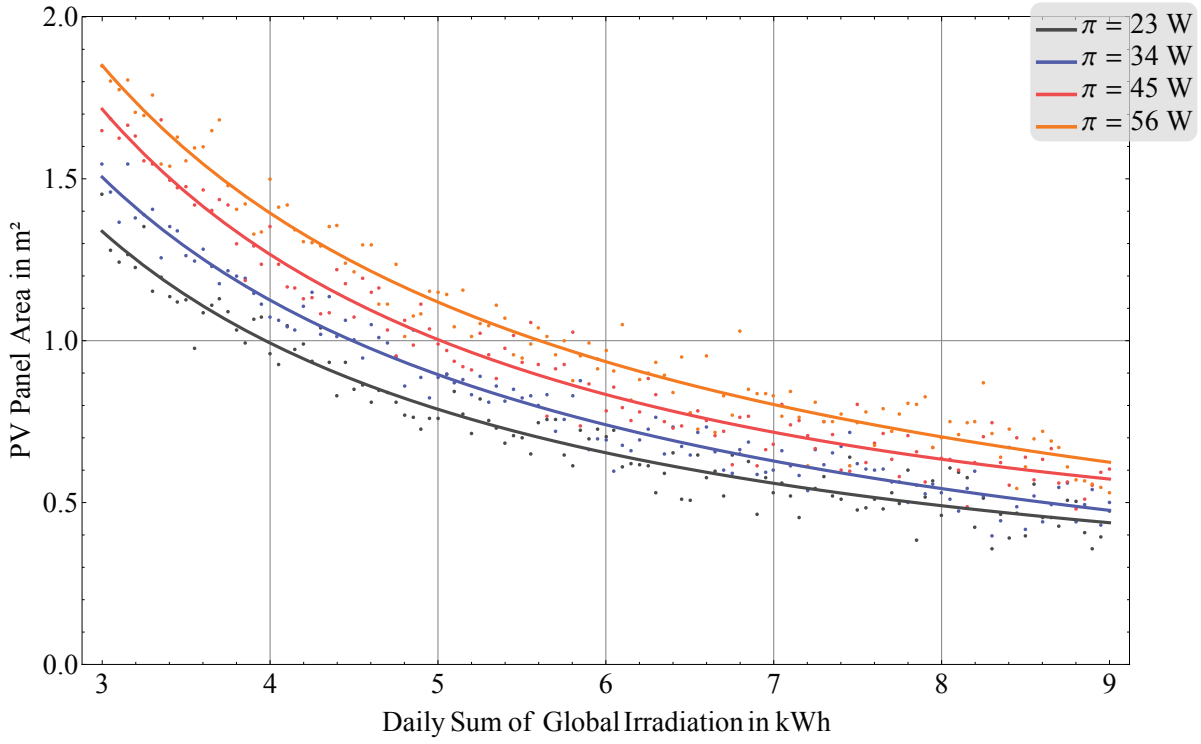
– *optimal tilt* curve. While a horizontal surface receives a mean value of $\overline{H} \approx 6.45\,\text{kWh}$ per day and m^2, the tilted surface captures $\overline{H} \approx 6.5\,\text{kWh}$. Tracked solar panels (i.e., tracked azimuth and tilt) can even capture higher amounts of irradiation by accounting for daily and seasonal variations.

Figure 8.6(a) illustrates the sensitivity of PV panel size to daily sum of global irradiation as well as PV module efficiency η. Similar to the following figures, rational functions have been used to fit simulation data. The first observation is that increasing levels of sunshine only yield regressive reductions of required panel size. For example, for $\eta = 0.25$, an increase from 4 to 5 kWh yields a panel reduction of approximately $0.3\,\text{m}^2$, an identical increase from 8 to 9 kWh only saves about $0.07\,\text{m}^2$. Moreover, panel size remains below $1\,\text{m}^2$ for at least $5\,\frac{\text{kWh}}{\text{m}^2}$ of daily irradiation, while mean irradiation is $6.5\,\frac{\text{kWh}}{\text{m}^2}$ in Tanzania. Generally, improvements in PV module efficiency can result in significant panel size reductions. Expected future developments, such as reaching $\eta = 0.35$, would reduce the required irradiation for $1\,\text{m}^2$ by 26 % to approximately $3.7\,\frac{\text{kWh}}{\text{m}^2}$, although further increases in η only yield diminishing returns. Moreover, energy-efficient radio equipment will also contribute to reducing panel size.

Figure 8.6(b) illustrates required surface areas for varying energy consumption levels π. Here, a power reduction of 45 W to 23 W (e.g., during night time operation) reduces irradiation requirements for a panel of $1\,\text{m}^2$ by approximately 20 % from 5 to $4\,\frac{\text{kWh}}{\text{m}^2}$. Power reduction can also be achieved

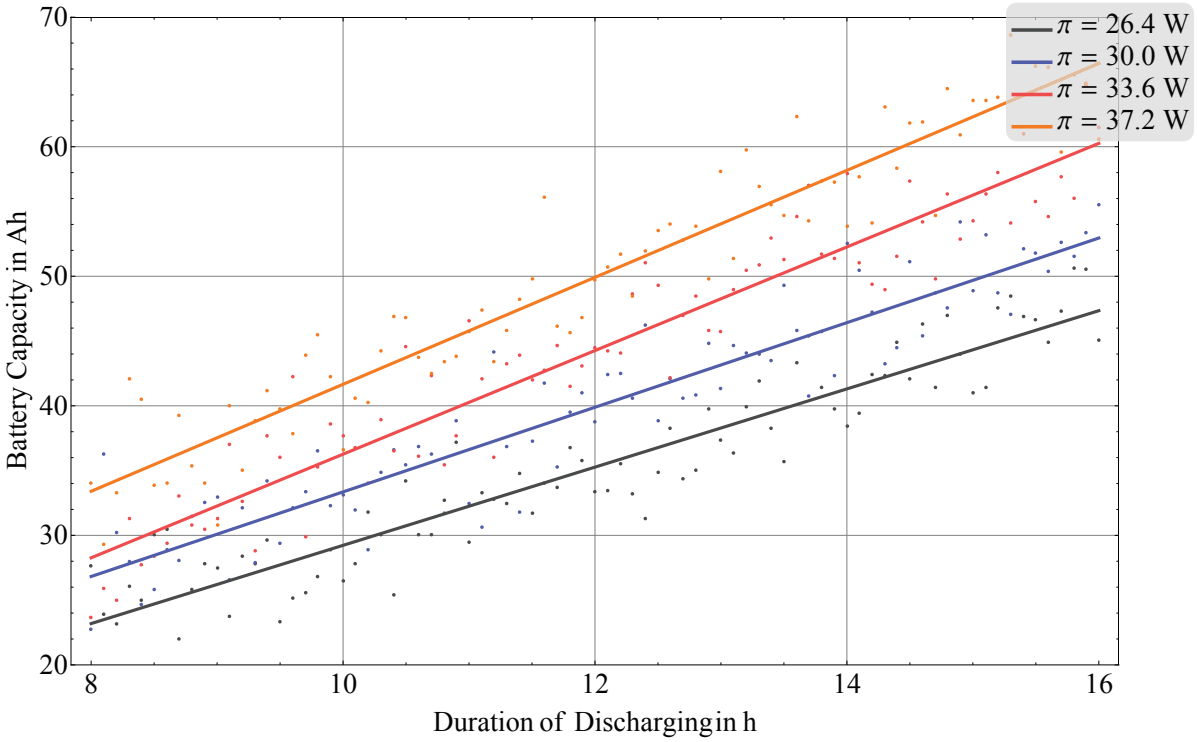


(a) Impact of PV module efficiency

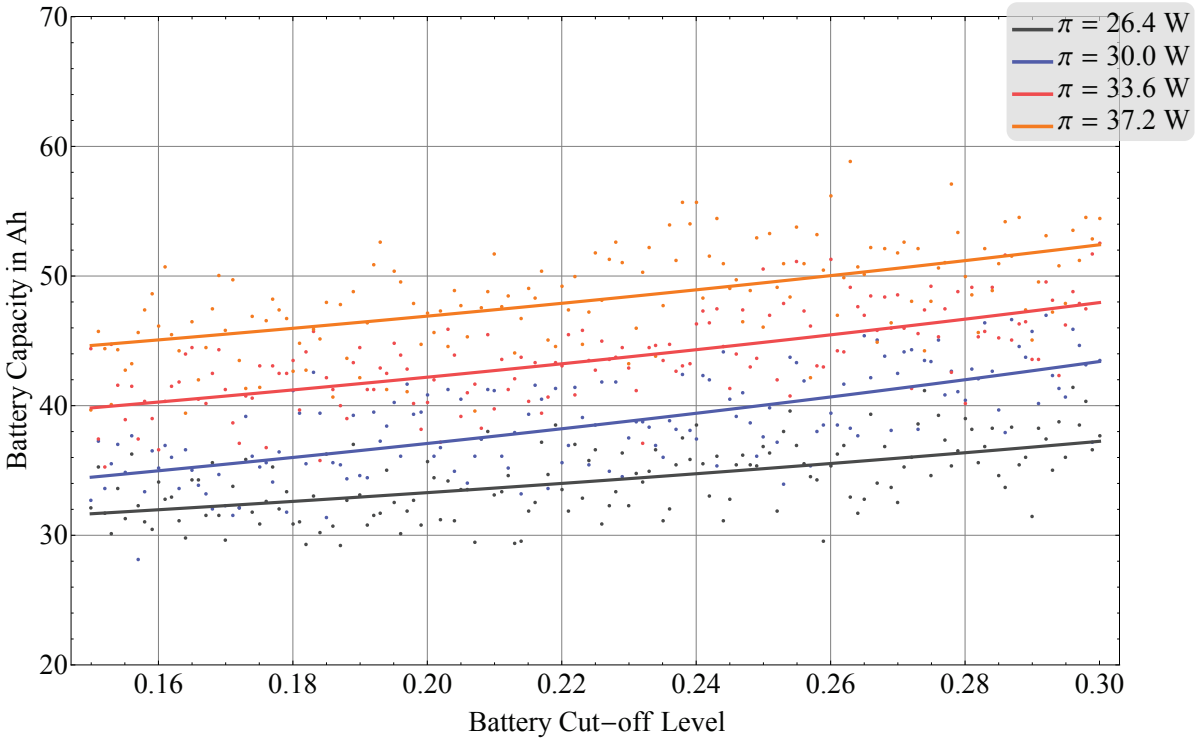


(b) Impact of mean power consumption

Figure 8.6: Sensitivity of PV panel size to different parameters



(a) Impact of discharging period



(b) Impact of battery cut-off level

Figure 8.7: Sensitivity of PV battery capacity to different parameters

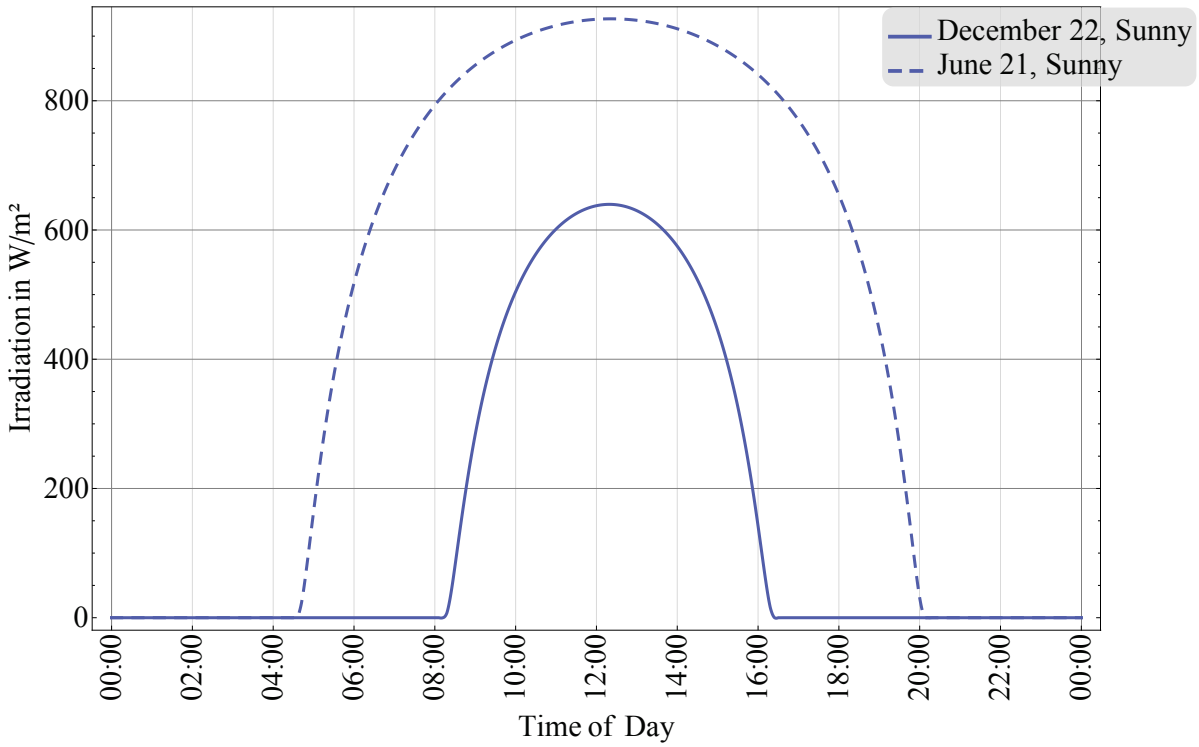
by appropriately controlling the number of active radio interfaces of a backhaul node, e.g., by means of topology control as analyzed in Section 8.4.

The sensitivity of battery capacity Q to the duration of discharging periods (usually nighttime) is illustrated in Figure 8.7(a). Q varies between 20 and 70 Ah for 8 to 16 hours of discharging. Required battery size also scales linearly with average power consumption π_{dis} during discharging. More specifically, $\frac{\partial Q}{\partial \pi} = \frac{1}{1-\beta_{\text{cut}}} \frac{\Delta t_{\text{dis}}}{U_{\text{dis}}}$ varies, depending on parameters, between $0.78\,\frac{\text{Ah}}{\text{W}}$ and $1.9\,\frac{\text{Ah}}{\text{W}}$. In Figure 8.7(b), minimum battery capacity as a function of battery cut-off level β_{cut} with parameter π is depicted. Although $\frac{\partial Q}{\partial \pi} = \frac{1}{1-\beta_{\text{cut}}} \frac{\Delta t_{\text{dis}}}{U_{\text{dis}}}$, Q increases almost linearly with cut-off level in the relevant range of $0.15 \leq \beta_{\text{cut}} \leq 0.3$. In summary, a battery capacity of 50 to 60 Ah (i.e., adding reasonable safety margin to Q_{T1}) is sufficient to allow for a continuous operation under normal conditions. In order to further decrease the likelihood of node failures due to drained energy sources, the deployment of a conventional car battery, which can exhibit capacity levels up to 80 to 90 Ah, is a viable alternative.

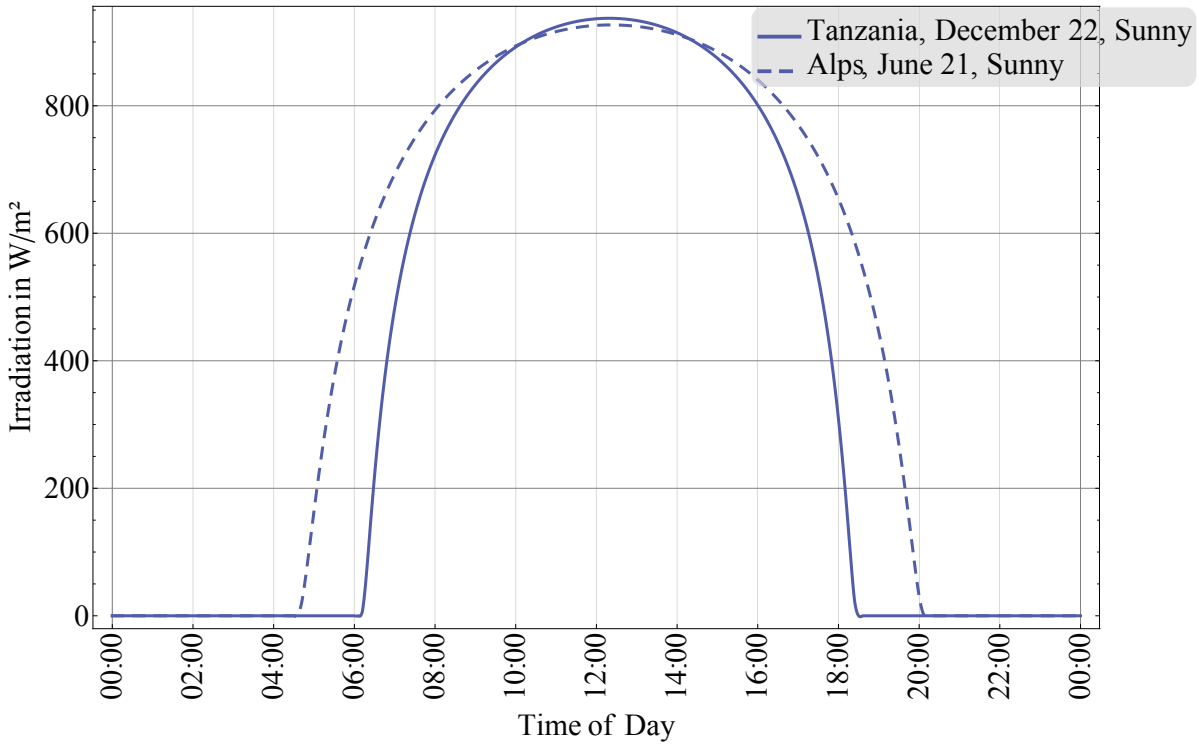
8.2.2 Alpine Scenario

The Alpine scenario exhibits a few fundamental deviations from the Tanzania scenario. Most importantly, weather conditions are not as optimal as in Sub-Saharan Africa, i.e., a considerably smaller number of sunny days is to be expected. Due to the northern latitude, irradiation levels fluctuate more significantly across the seasons. Finally, users' telecommunication behavior is more traffic-intensive, requiring higher network capacity. In total, this results in a less favorable overall balance of energy supply and demand.

On the energy supply side, Figures 8.8 and 8.9 illustrates the daily irradiation pattern depending on time of the year and weather conditions. Seasonal difference is depicted in Figure 8.8(a) that shows irradiation on a sunny winter and summer day, respectively. At summer solstice, the sun shines about twice as long (16 h in contrast to 8 h), global irradiation peaks at approximately $925\,\frac{\text{W}}{\text{m}^2}$ ($637\,\frac{\text{W}}{\text{m}^2}$), and daily sum of irradiation amounts to $11.2\,\frac{\text{kWh}}{\text{m}^2}$ ($3.8\,\frac{\text{kWh}}{\text{m}^2}$). Compared to Tanzania, the higher irradiation sum on summer days is achieved because of earlier sunrise and later sunset, as shown in Figure 8.8(b). Of course, this is reversed in winter. Figures 8.9(a) and 8.9(b) illustrate typical irradiation patterns for partially cloudy and overcast weather in summer. Due to fluctuations and lower peak values, daily sum of irradiation is reduced considerably.

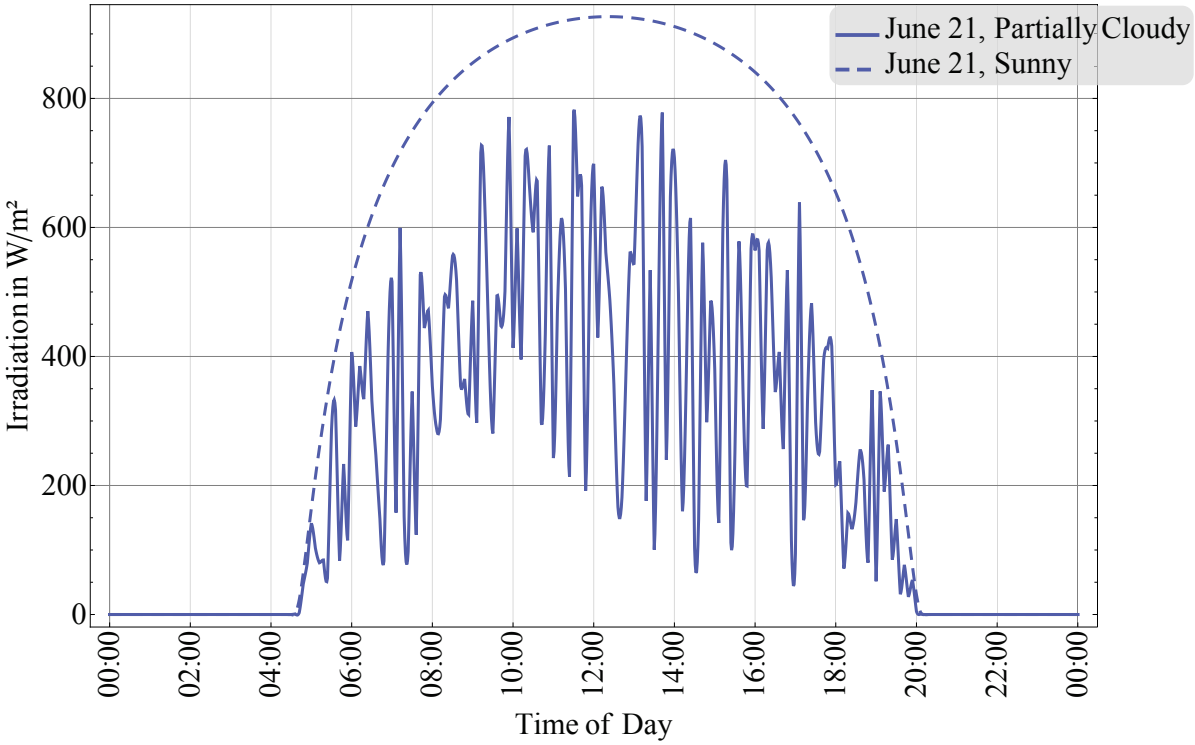


(a) Summer vs. winter solstice

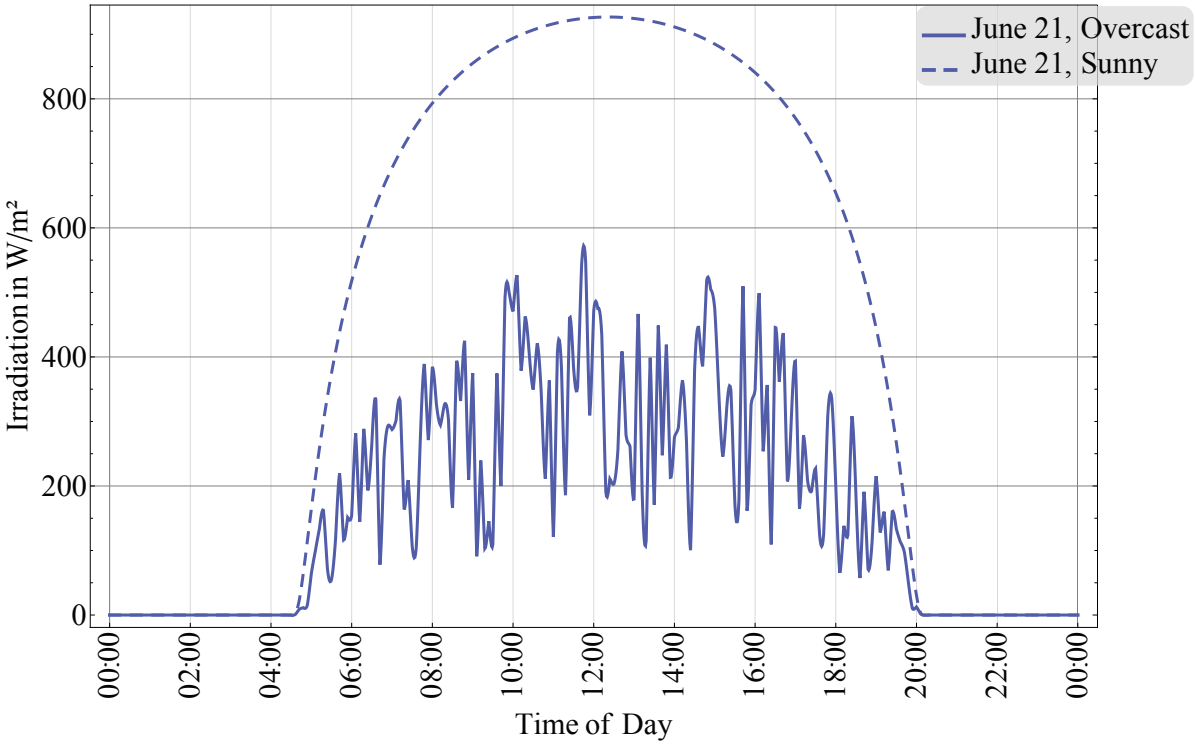


(b) Comparison with Tanzania

Figure 8.8: Intraday global irradiation levels in the Alps



(a) Partially cloudy sky



(b) Overcast sky

Figure 8.9: Intraday global irradiation levels in the Alps depending on weather conditions

Reference Analysis with Continuously Sunny Weather

An initial analysis that assumes a sequence of sunny days yields a first reference with respect to the technical parameters of the PV module components. For example, at winter solstice at 12 pm, a solar panel of area $A = 1\,\mathrm{m}^2$ and a typical PV module efficiency of $\eta = 0.25$ of modern mass produced PV modules yields a peak output power of $637\,\frac{\mathrm{W}}{\mathrm{m}^2} \cdot 1\,\mathrm{m}^2 \cdot 0.25 \approx 159\,\mathrm{W}$, which is sufficient for operating a radio node. Further, assuming a power demand of 50 W, the required irradiation level is $200\,\frac{\mathrm{W}}{\mathrm{m}^2}$, a threshold that is achieved during approximately 7 h (between 8:45 am and 3:45 pm) on a sunny winter day. More specifically, assuming, similar to the Tanzania scenario, $N_{\mathrm{x,n}} = 1$ active radio interface between 10 pm and 6 am and $N_{\mathrm{x,d}} = 3$ active radio interfaces during the rest of the time, power demand π_{op} for network operation can be computed according to Equations (7.10), (7.11a), and (7.11b). While this power demand is met by autarkic power supply from a PV module, a buffer battery has to help out during periods of darkness or extremely bad weather. For example, during night time in winter, the battery has to supply power for approximately 17 h (between 4 pm and 9 am). During this time, a node has a total energy demand of $E_{\mathrm{A1}} = 8\,\mathrm{h} \cdot 23\,\mathrm{W} + 9\,\mathrm{h} \cdot 45\,\mathrm{W} = 589\,\mathrm{Wh}$. Assuming an average discharging voltage of $U_{\mathrm{dis}} = 12\,\mathrm{V}$ and a cut-off battery level of $\beta_{\mathrm{cut}} = 0.25$, the required battery capacity Q_{A1} can be computed according to Equation (7.15)

$$Q_{\mathrm{A1}} = \frac{1}{1-\beta_{\mathrm{cut}}} \frac{\Delta E_{\mathrm{A1}}}{U_{\mathrm{dis}}} \approx 65.4\,\mathrm{Ah},$$

which is equivalent to 157 % of capacity Q_{T1} required in the Tanzania scenario. At summer solstice, when nights are considerably shorter, the battery has to provide $E_{\mathrm{A2}} = 8\,\mathrm{h} \cdot 23\,\mathrm{W} + 2\,\mathrm{h} \cdot 45\,\mathrm{W} = 274\,\mathrm{Wh}$, which results in capacity Q_{A2}

$$Q_{\mathrm{A2}} = \frac{1}{1-\beta_{\mathrm{cut}}} \frac{\Delta E_{\mathrm{A2}}}{U_{\mathrm{dis}}} \approx 22.8\,\mathrm{Ah}.$$

Particularly in winter, reduced sunshine levels are common in central Europe. Consequently, at day time, batteries cannot always be recharged to the full extent, since all PV energy is fully taken for node operation. This means that, occasionally, batteries have to provide energy for longer phases than a single night, resulting in more excessive capacity requirements, as discussed in Section 8.2.2. Further, as noted before, the yield of a PV module shall satisfy the energy demand $E_{\mathrm{op}} + E_{\mathrm{ch}} = \int \pi(t)dt + \frac{\Delta E_{\mathrm{A1}}}{\eta_{\mathrm{Wh}}}$. The re-

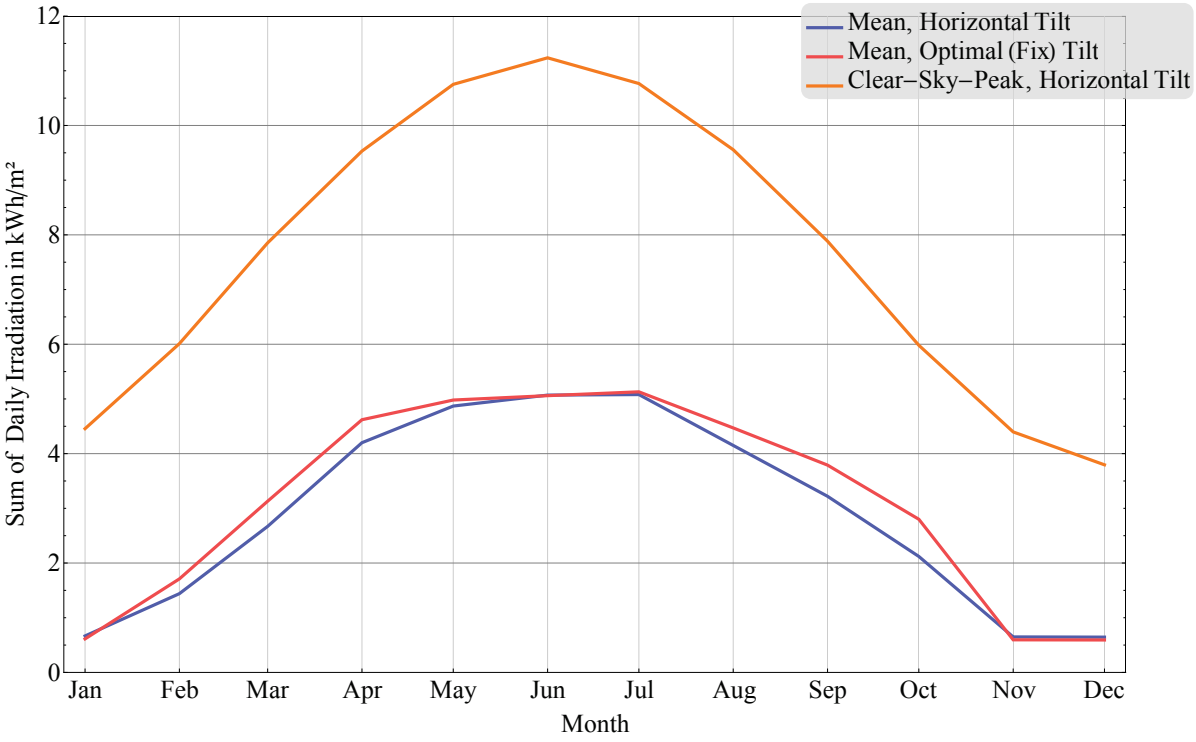


Figure 8.10: Sum of daily global irradiation in the Alps

quired panel areas A_{A1} and A_{A2} are calculated according to Equation (7.19). Assuming a Wh-efficiency of $\eta_{\mathrm{Wh}} = 0.75$, an average power consumption of $\overline{\pi} = \pi_{\mathrm{d}} = 45.2\,\mathrm{W}$ for node operation, a daily sum of solar irradiation of $\overline{H}_{\mathrm{A1}} = 3.8\,\frac{\mathrm{kWh}}{\mathrm{m}^2}$ and $\Delta t_{\mathrm{ch}} = 7\,\mathrm{h}$ at winter solstice and $\overline{H}_{\mathrm{A2}} = 11.2\,\frac{\mathrm{kWh}}{\mathrm{m}^2}$ and $\Delta t_{\mathrm{ch}} = 14\,\mathrm{h}$ at summer solstice, the following minimum panel sizes of A_{A1} and A_{A2} are required:

$$A_{\mathrm{A1}} \approx 1.16\,\mathrm{m}^2 \text{ and}$$
$$A_{\mathrm{A2}} \approx 0.36\,\mathrm{m}^2.$$

The following section performs a sensitivity analysis of the considered quantities with respect to selected parameters, thus, among other, accounting for fluctuating, unsettled weather conditions or technological advancements.

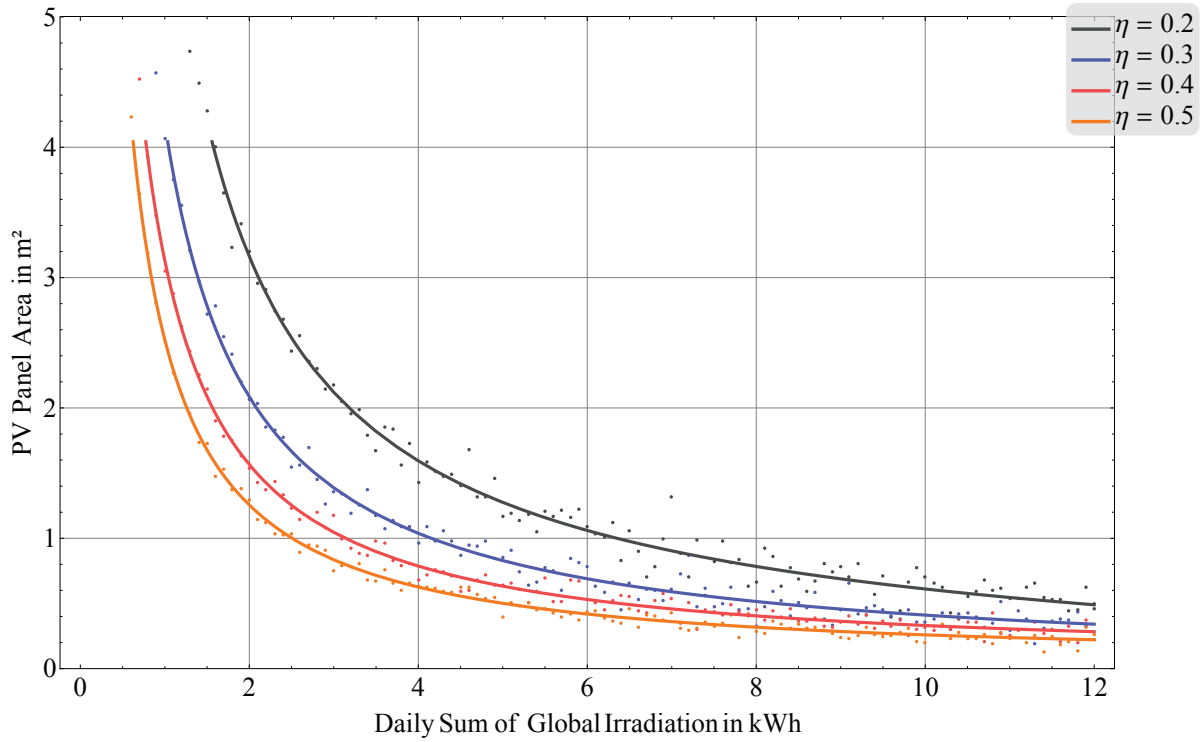
Sensitivity Analysis

The Alpine region, for most part, is characterized by a moderate climate. Higher elevations exhibit so-called mountain climate where temperatures are lower throughout the year and vegetation becomes scarcer. On average, the region has a yearly sum of global irradiation in the range of $850 < H_y < 1400\,\frac{\mathrm{kWh}}{\mathrm{m}^2}$, depending on local conditions [Joi12]. Figure 8.10 depicts the mean daily sum of global irradiation $\overline{H}$ throughout the year.

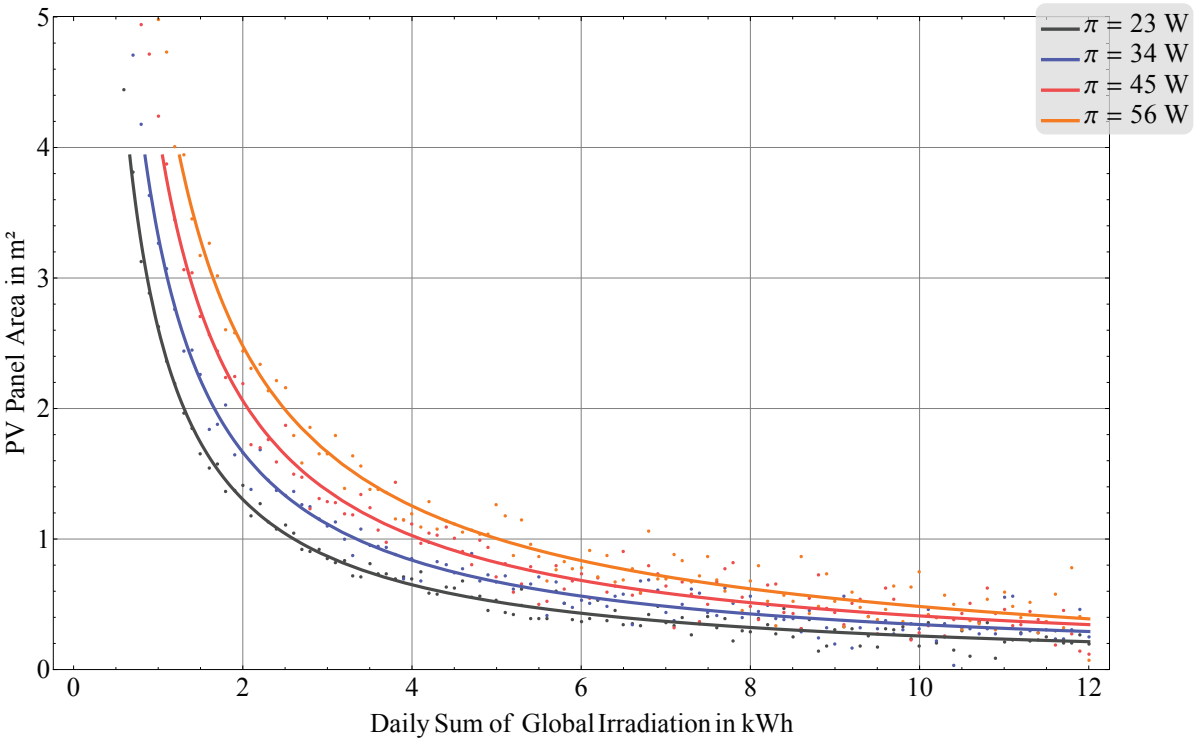
The *clear-sky-peak* line indicates the maximum energy per m^2 arriving on the (horizontal) ground in case of sunny weather. Since the Alps are located in the northern hemisphere, highest values occur in June. However, due to unsettled weather conditions, the actual mean sum of horizontal irradiation is significantly lower, as depicted by the *mean – horizontal* curve. It exhibits its maximum in July, reaching more than $5\,\frac{\text{kWh}}{\text{m}^2}$, while the minimum in winter months is as low as $0.65\,\frac{\text{kWh}}{\text{m}^2}$. Assuming a fixed surface (no sun tracking), the annual mean value of solar irradiation is maximized for a southbound panel with a tilt of 27°, as illustrated by the *mean – optimal tilt* curve. While a horizontal surface of $1\,\text{m}^2$ receives a mean value of $\overline{Y} \approx 2.91\,\text{kWh}$ per day, the tilted surface captures $\overline{Y} \approx 3.13\,\text{kWh}$ per day. As mentioned, panels with tracked tilt and azimuth could capture higher amounts of irradiation.

Figure 8.11(a) depicts the required panel area as a function of daily sum of global irradiation. Increasing levels of sunshine only yield diminishing returns in terms of reduction of panel size. Similarly, the effect of efficiency gains continuously attenuates with higher efficiency levels. Further, and most importantly, mean irradiation sum in November, December, and January, which is in the range of 0.6 to $0.7\,\frac{\text{kWh}}{\text{m}^2}$ per day, would theoretically require a comparatively large panel surface of at least $8\,\text{m}^2$, depending on PV module efficiency. In addition, a reduction of mean power consumption does not significantly counteract the problem of low irradiation levels, as depicted in Figure 8.11(b). Even a scaled-down power consumption of 23 W per node would require a panel size of approximately $5\,\text{m}^2$ in case daily yield amounts to $1\,\frac{\text{kWh}}{\text{m}^2}$. In summary, if a panel size of 1 to 1.5m^2 is considered as the upper bound, the analysis demonstrates that autarkic energy supply only guarantees reliable operation in the months of April through September, where mean daily sum of irradiation exceeds $3.5\,\frac{\text{kWh}}{\text{m}^2}$, which is a lower limit considering the efficiency of today's PV equipment.

As for required battery capacity, Figure 8.12(a) depicts a linear dependency on the duration of discharging. Obviously, the longer the night, the higher the required capacity becomes. However, it also shows that an adaption of network capacity to night time demand, which results in less power consumption, can help to significantly decrease required capacity, a fact that is exploited in the TC algorithms developed in this work. Eventually, Figure 8.12(b) shows that a lower cut-off level q of a battery only yields minor gains with respect to battery capacity reduction. Overall, battery dimension stays within feasible intervals of capacity for the Alpine scenario

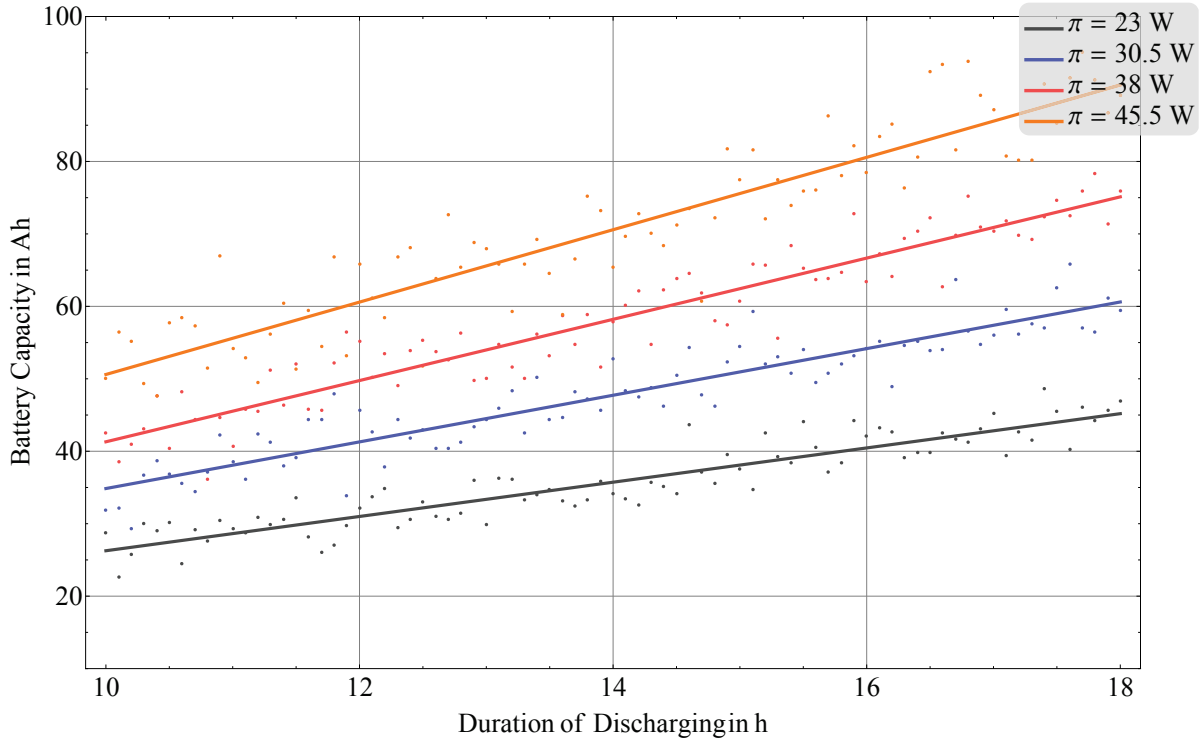


(a) Impact of PV module efficiency

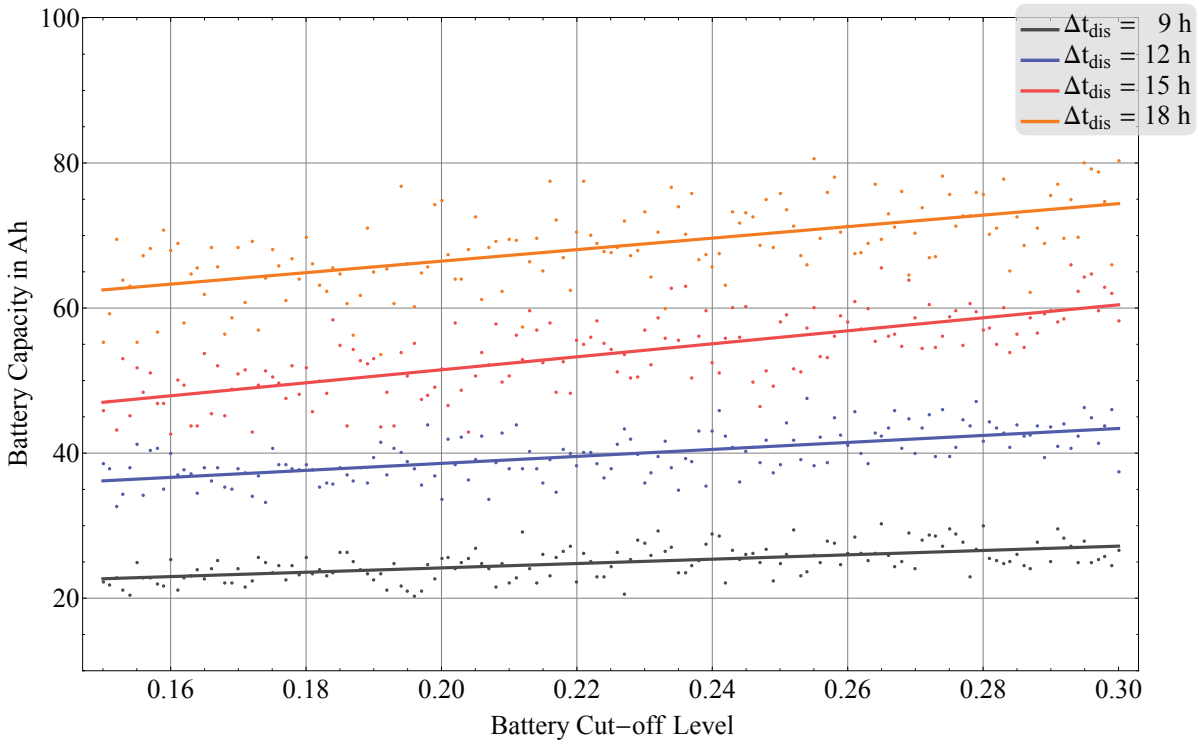


(b) Impact of mean power consumption

Figure 8.11: Sensitivity of PV panel size to different parameters



(a) Impact of discharging period



(b) Impact of battery cut-off level

Figure 8.12: Sensitivity of PV battery capacity to different parameters

and is mainly determined by power demand and time span between charging periods.

In summary, the Alpine scenario is significantly more challenging than the Tanzania scenario. For both cases, battery capacity is within reasonable boundaries, e.g., car batteries would be a reasonable choice. However, while reasonably sized PV modules in Tanzania can reliably produce sufficient energy to operate backhaul nodes and recharge drained batteries, Alpine irradiation levels in winter months do not allow for such a setup. Here, in case no power grid is available due to remote locations, other energy supply solutions have to be found.

8.3 Performance Evaluation of Deployment Algorithms

In Section 6.2, the author has presented two approximation algorithms (*maximum path redundancy* and *cycle-based TC*) for the problem of initial topology in WBNs as expressed in Equation (3.2). Further, the author has shown that the approximation level γ is a function of node distance $\delta_{i,j}$ and maximum node degree $d_{\max}$. More specifically, for the given case, approximation levels have been proven to not fall below $\gamma = \frac{2}{d_{\max}}$. The following evaluation demonstrates the exceptionally good approximation levels of the designed algorithms by computing initial graph topologies for numerous node configurations that vary with respect to density, number, and maximum degree of nodes. In total, several thousands of different network instances have been evaluated. The results clearly highlight the capability of the algorithms of constructing a graph that is not just 2-edge-connected, but generally rather $(d_{\max} - 1)$-connected, thus frequently reaching an approximation level of $\gamma = \frac{d_{\max}-1}{d_{\max}} > \frac{2}{d_{\max}}$. Network capacity has been used as a further metric for network topology evaluation.

8.3.1 Maximum Path Redundancy Algorithm

The Maximum Path Redundancy (MPR) TC algorithm aims at maximizing path redundancy within the backhaul network. This is achieved by maximizing the minimum vertex degree of backhaul nodes. In the optimum case (upper bound), all nodes have a vertex degree of $d_{\max}$ and edge connectivity also equals $d_{\max}$. For a more detailed analysis of the resulting WBN structure, the computation of network capacity follows the *Max-Flow-Min-*

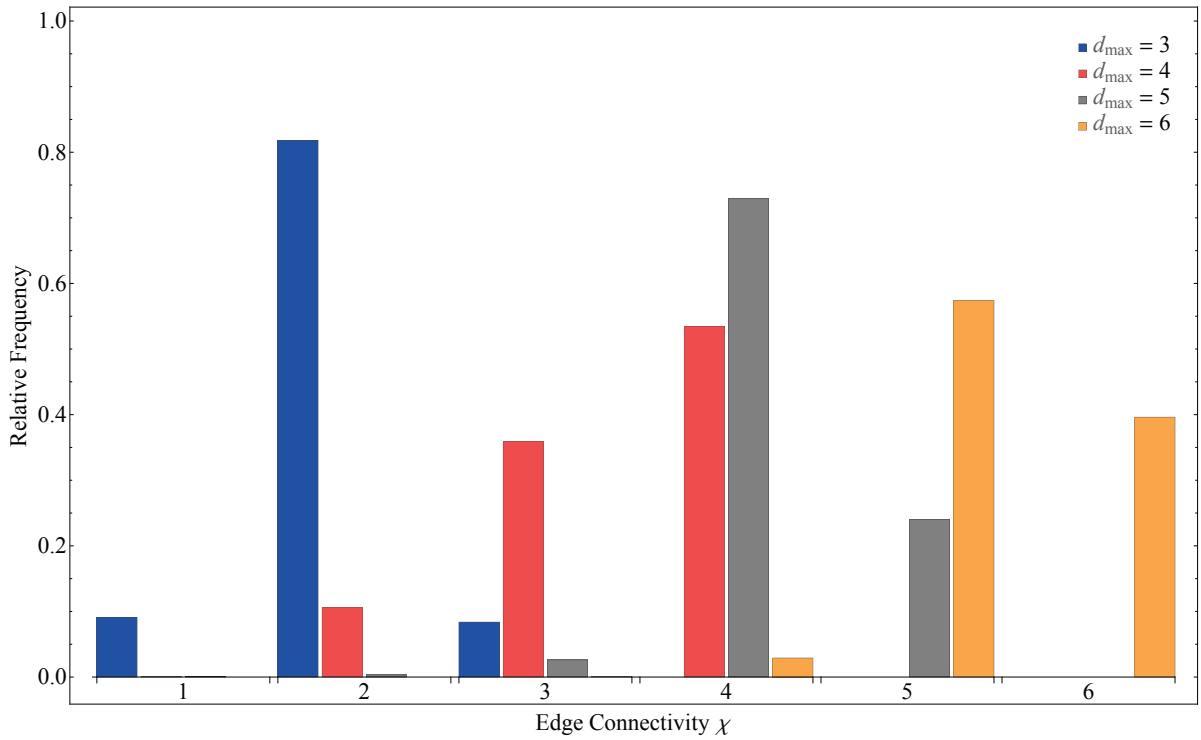


Figure 8.13: Relative frequency of edge connectivity in MPR TC

Cut theorem ([FF55], cf. Section 3.5.1) that maximizes the network flow between a set of source nodes on the one side and a set of target nodes on the other. Gateway nodes $\mathcal{V}_{\text{GW}}$ are chosen as source nodes and remaining nodes $\mathcal{V} \setminus \mathcal{V}_{\text{GW}}$ as target nodes. Figures 8.13 through 8.16 illustrate the evaluation results for the MPR topology control algorithm.

In Figure 8.13, relative frequency of edge connectivity χ is depicted for four different values of parameter $d_{\max}$ (maximum node degree). As elucidated in previous sections, $d_{\max}$ serves as an upper bound for χ. For $d_{\max} = 3$, the mode of edge connectivity becomes $\chi = 2$; for approximately 9%, MPR topology control yields $\chi = d_{\max}$. Another 9% of networks remain at a level of $\chi = 1$, i.e., only one edge needs to be removed to disconnect the network. For increasing $d_{\max}$, mode of edge connectivity is either $\chi = d_{\max} - 1$ or even $\chi = d_{\max}$ for $d_{\max} = 4$. Moreover, with the exception of $d_{\max} = 3$, the critical bound of $\gamma = \frac{2}{d_{\max}}$ is surpassed reliably, i.e., the MPR algorithm does not create one-edge-connected networks. Complementing Figure 8.13, Figure 8.14 shows the average edge connectivity χ achieved by MPR topology control as a function of number m of backhaul nodes. The chart further includes upper bounds (dashed lines) for each realization of parameter $d_{\max}$. Generally, it can be observed that, on average, the MPR algorithm approaches the optimum edge connectivity as much as 0.5 to 1 base units, thus confirming that $\chi = d_{\max} - 1$ or $\chi = d_{\max}$ in the majority

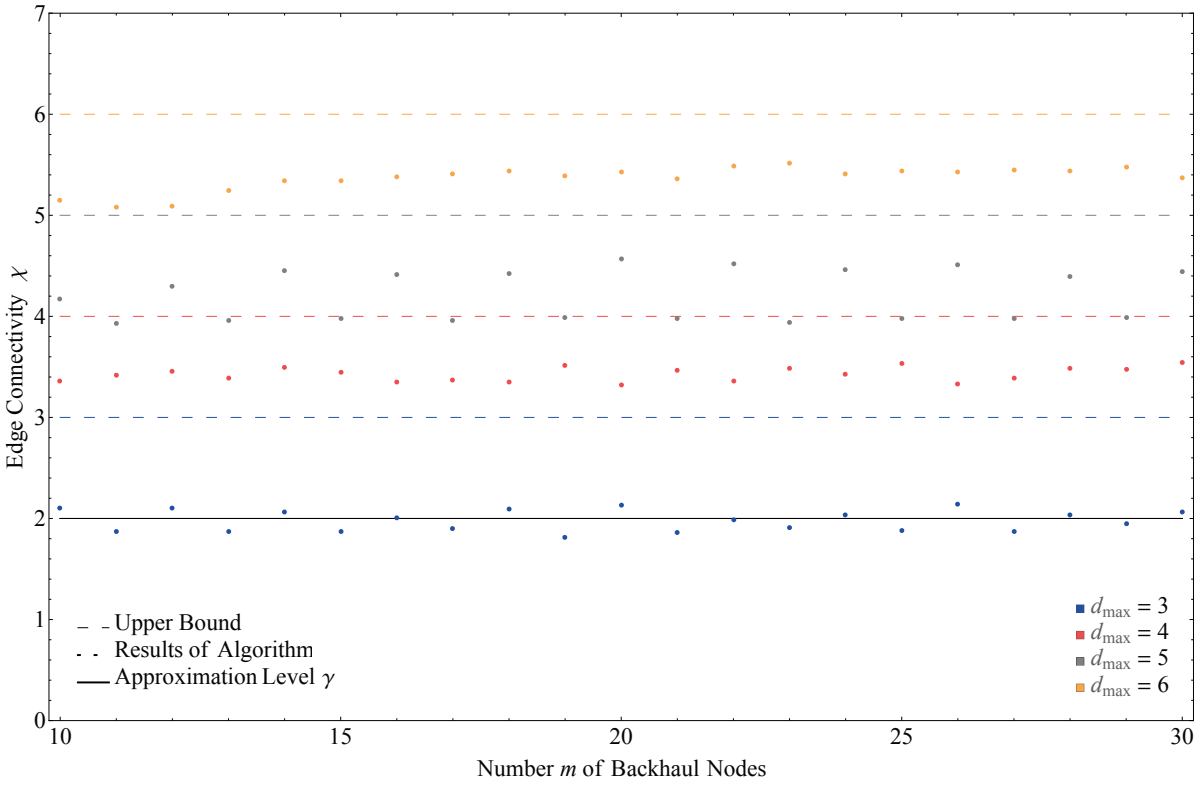


Figure 8.14: Average edge connectivity in MPR TC

of realizations. Further, there is a relatively constant performance with respect to size m of the network. More specifically, the redundancy level does not significantly degrade if the number of nodes is reduced. A common phenomenon is also mirrored in Figure 8.14: in case of an odd value for maximum node degree, the realized average edge connectivity of the networks is significantly lower for an odd number than for an even number of vertices. This can be observed for $d_{\max} = 3$ or $d_{\max} = 5$, where the points follow a zig-zag-pattern, i.e., lower edge connectivity for an odd number of backhaul nodes, higher values of χ for even numbers of backhaul nodes. This is a direct consequence of the fact that, formally, a graph with an odd number of vertices cannot be constructed to be $d_{\max}$-regular if $d_{\max}$ (i.e., the maximum permitted vertex degree) is odd, too. Generally, the maximum number $n_{\max}$ of undirected edges that can be established in a graph is equal to $\left\lfloor \frac{m\, d_{\max}}{2} \right\rfloor$ ($d_{\max}$ edges at m nodes and a factor of $\frac{1}{2}$ since each edge connects two nodes). If $\frac{m\, d_{\max}}{2}$ is an integer, then, in case of no additional restrictions, $\chi = d_{\max}$. However, if $\frac{m\, d_{\max}}{2}$ is not an integer, then at least one node v remains with degree $d(v) = d_{\max} - 1$, i.e., $n_{\max} = \frac{m\, d_{\max} - 1}{2}$. Therefore, coercively, $\chi \leq d_{\max} - 1$ and the according values of mean edge connectivity in Figure 8.14 immediately drop significantly.

Figure 8.15 depicts network capacity of a WBN as a function of number

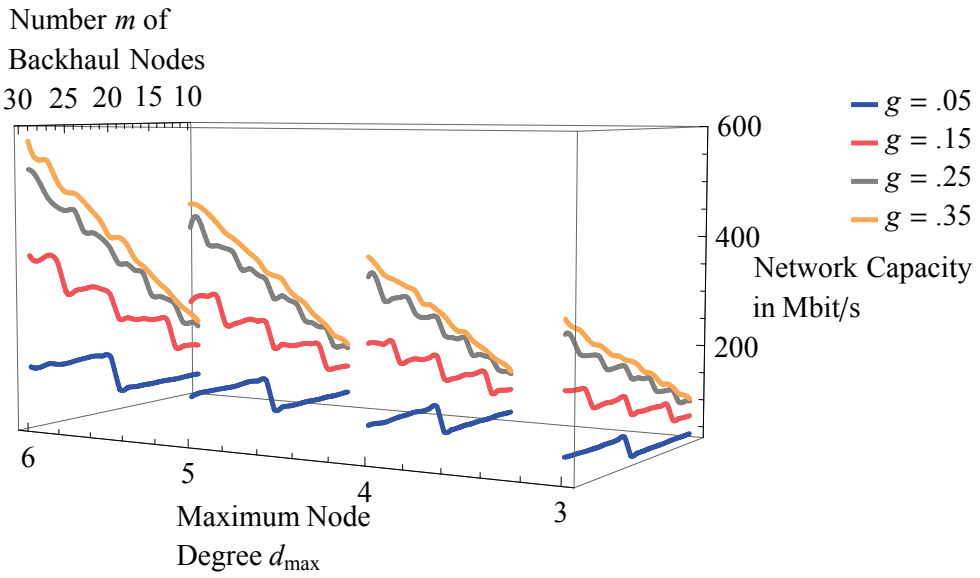


Figure 8.15: Network capacity of MPR topologies

m of backhaul nodes and maximum node degree $d_{\max}$. The author has varied the fraction g of gateway nodes among all backhaul nodes from 0.05 up to 0.35. Network capacity varies between approximately 55 and close to 600 Mbit/s for the given parameter space ($3 \leq d_{\max} \leq 6$, $10 \leq m \leq 30$, $0.05 \leq g \leq 0.35$). It immediately becomes visible that an increasing fraction of gateway nodes yields a higher network capacity if other variables are kept constant. However, the comparison of individual gains reveals diminishing marginal utility of increasing the share of gateway nodes, i.e., $\frac{\partial c_G}{\partial g} > 0$, but $\frac{\partial^2 c_G}{\partial g^2} < 0$. Furthermore, a larger network, i.e., additional nodes, does not result in higher backhaul capacity unless a gateway node is added. In that case, WBN capacity makes an immediate jump to a higher level, where the curve again levels until an additional gateway node is added. In Figure 8.15, the effect is particularly illustrated for $g = 0.15$. At large, the effect also justifies the chosen approach of approximating network capacity as a function of capacity of gateway edges, as proposed in Section 6.3.1.

Finally, Figure 8.16 again plots backhaul network capacity against m and $d_{\max}$, this time using a second-order interpolation creating a series of surfaces of parameter g. Once more, the step-wise increase when a gateway node is added as well as the decreasing benefit of higher shares of gateway nodes are clearly visible. The gray plane for $g = 0.25$ at times even surpasses the yellow plane for $g = 0.35$. Additionally, the figure shows the effect of increasing $d_{\max}$. While for a small number of nodes, additional edges at nodes do not help to boost network capacity (c_G) significantly, a

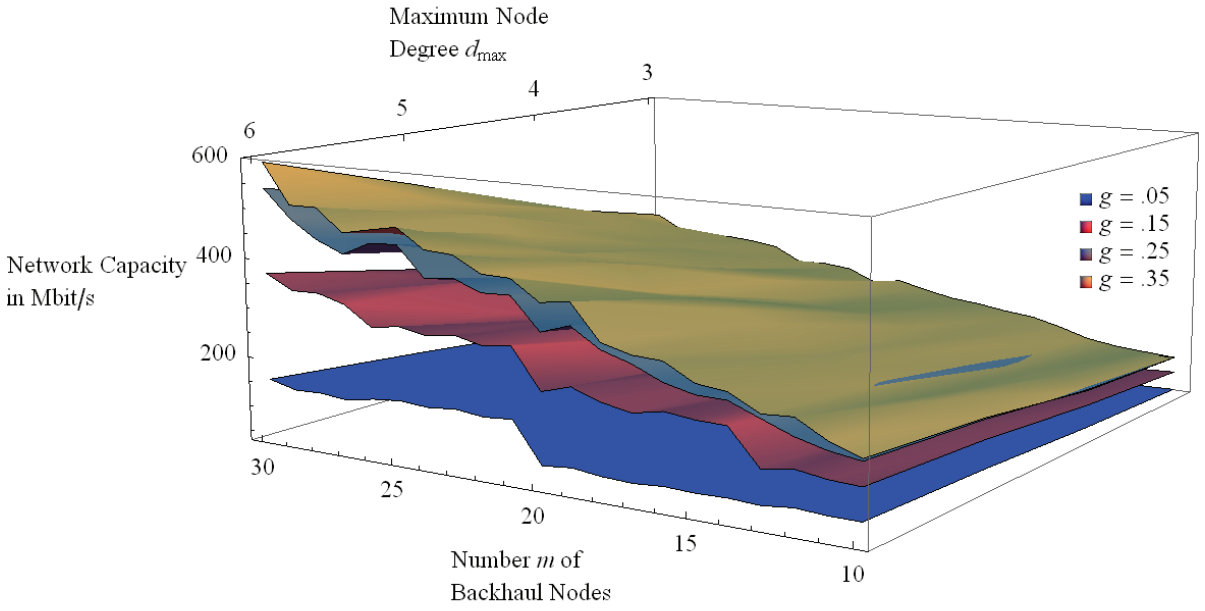


Figure 8.16: Capacity in MPR TC depending on number of deployed gateway nodes

positive effect of $d_{\max}$ can be observed if the network consists of a larger number of backhaul nodes, i.e., $\frac{\frac{\partial c_G}{\partial d_{\max}}}{\partial m} > 0$.

8.3.2 Cycle-based Topology Control

The main motivation for designing the cycle-based TC algorithm lies in the fact that the MPR approach yields some one-edge-connected backhaul networks, which is an insufficient level of redundancy. Therefore, cycle-based TC particularly emphasizes the surpassing of a lower threshold, which in turn is accompanied by disadvantages in approximating optimum redundancy levels.

Akin to the MPR topology control approach, the performance of cycle-based TC has been evaluated on several dimensions. The availability of redundant paths is shown in Figure 8.17. For any $d_{\max}$, the mode of edge connectivity is $\chi = d_{\max} - 1$, whereas the optimum value is never reached in more than 12% of cases. However, a comparatively high percentage of graphs is designed with an edge connectivity of $\chi \leq d_{\max} - 1$, e.g., in the case of $d_{\max} = 5$, the relative frequency of $\chi = 3$ is larger than 20%.

As an inherent characteristic of cycle-based topology control, the algorithm completely avoids the construction of one-edge-connected networks. Figure 8.18 reveals the correlation between edge connectivity and the num-

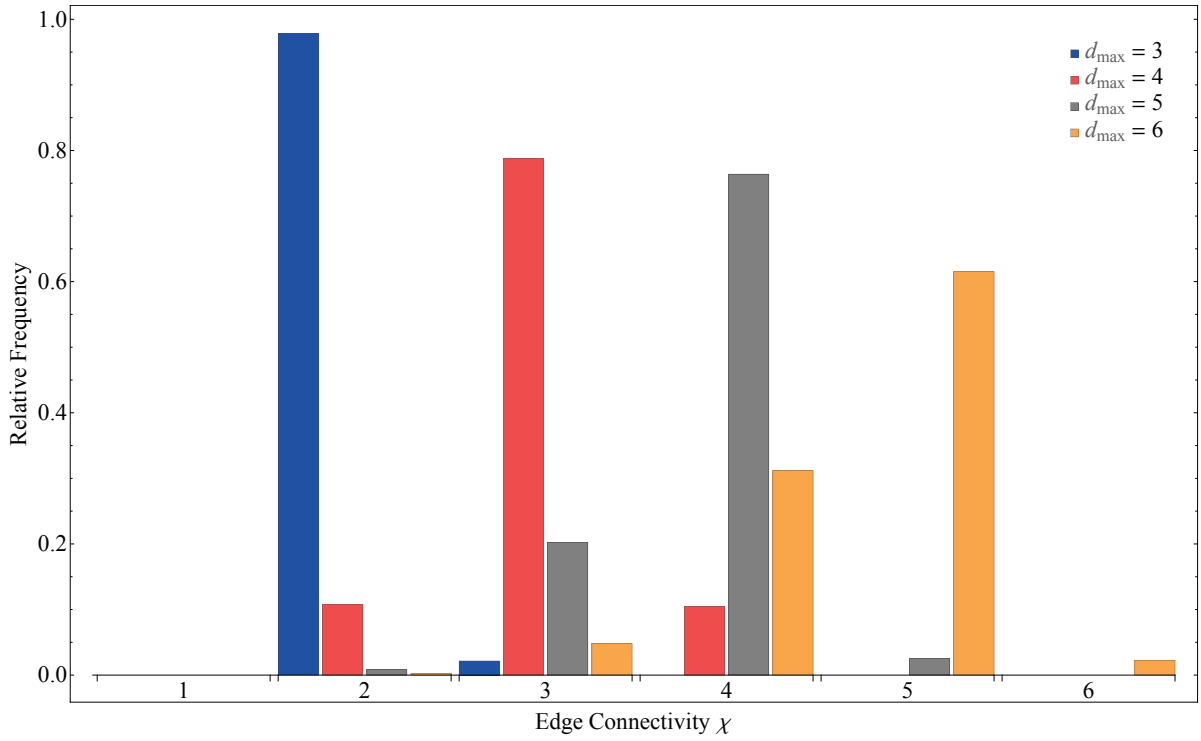


Figure 8.17: Relative frequency of edge connectivity in cycle-based TC

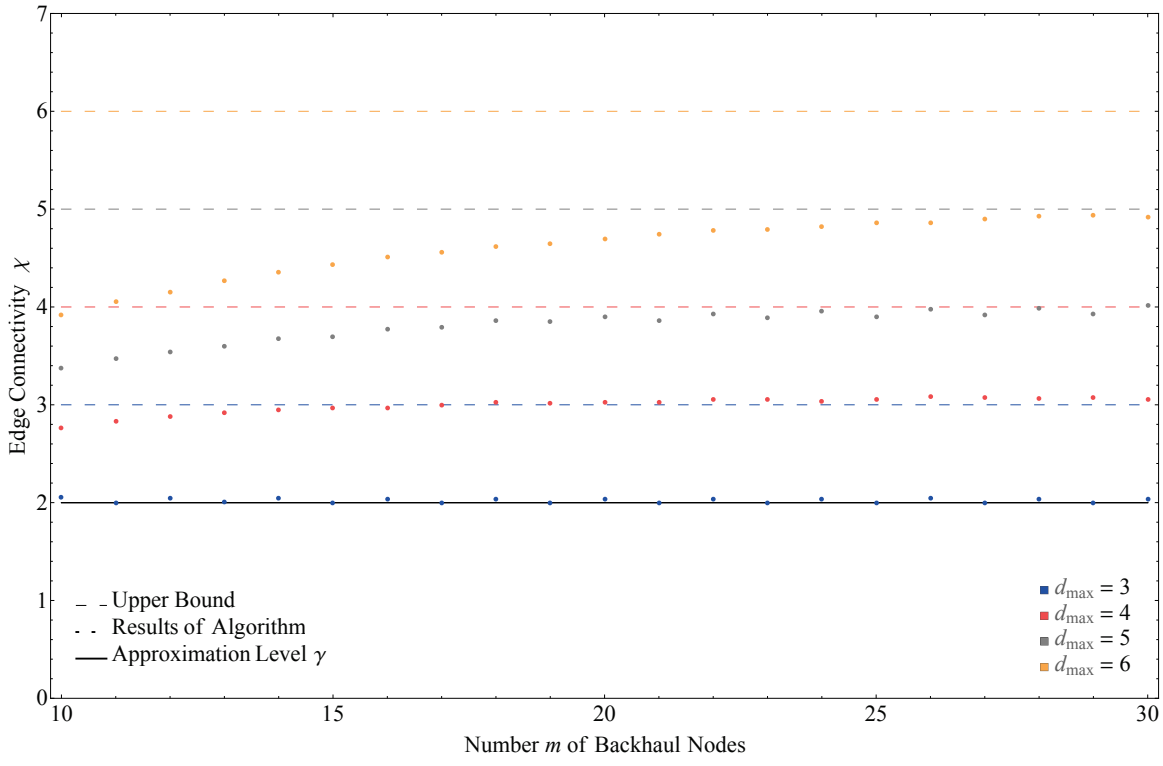


Figure 8.18: Average edge connectivity in cycle-based TC networks

ber of nodes in a backhaul network. While there is a slightly positive correlation for a smaller number of nodes, edge connectivity levels out for a higher number of nodes, i.e., $\frac{\partial \chi}{\partial m} > 0$ and $\frac{\partial^2 \chi}{\partial m^2} < 0$. Moreover, since cycle-based TC has fewer degrees of freedom than the MPR approach, the benefit of an increased $d_{\max}$ is not as significant and the realized values for χ do not approximate the optimum value as well as MPR topology control.

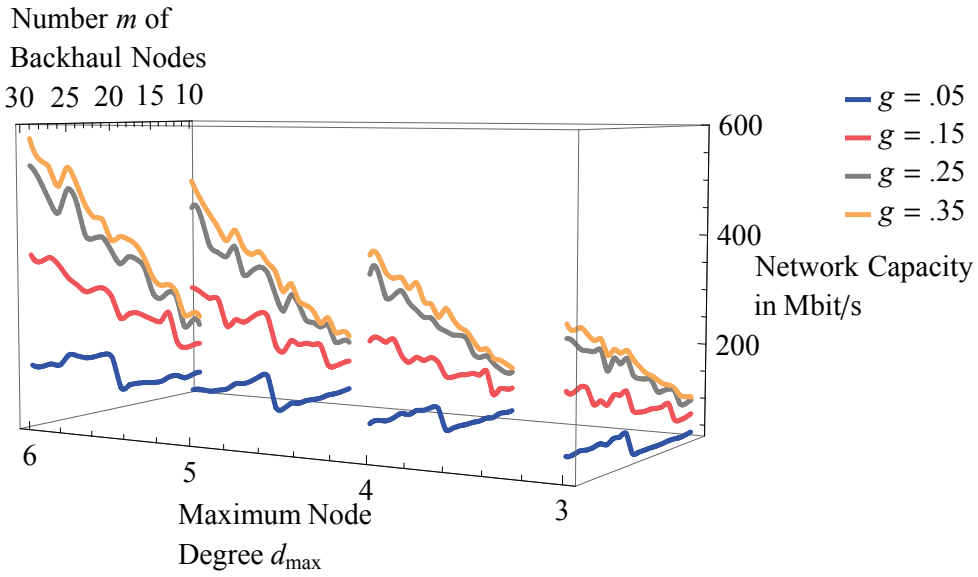


Figure 8.19: Network capacity of cycle-based topologies

Figures 8.19 and 8.20 plot network capacity against $d_{\max}$ and number of backhaul nodes. The parameter space is chosen identical to Section 8.3.1 ($3 \leq d_{\max} \leq 6$, $10 \leq m \leq 30$, and $0.05 \leq g \leq 0.35$). Both figures indicate that network capacity ranges from approximately 60 Mbit/s to almost 600 Mbit/s. The former figure shows that an increased share of gateway nodes does not yield constant but rather diminishing gains in network capacity, akin to the MPR approach. In fact, the switch from $g = 0.25$ to $g = 0.35$ induces an average improvement of network capacity of less than 10 Mbit/s. Further, an increasing number of backhaul nodes does only boost network capacity if the added node is a gateway node. The latter figure, which applies a second-order interpolation on the simulation results, not only confirms these observations, but also shows that a growing $d_{\max}$ is more beneficial for networks consisting of higher number of nodes.

In summary, Table 8.4 compares the evaluation results of MPR and cycle-based topology control. It focuses on the two major evaluation metrics, redundancy level and and network capacity.

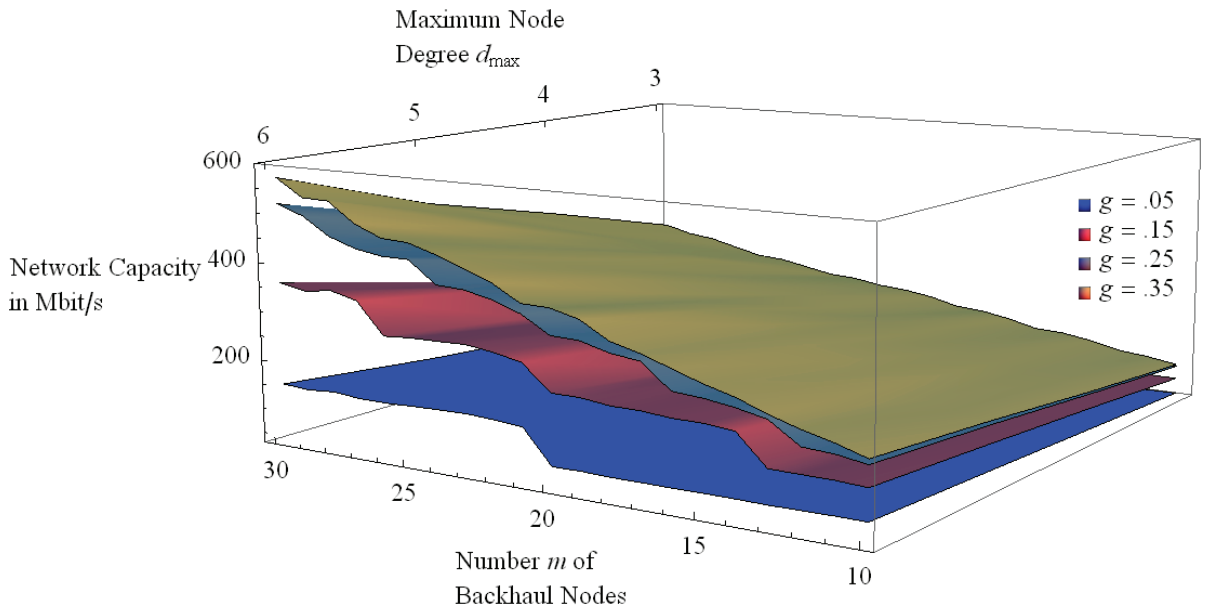


Figure 8.20: Capacity in cycle-based TC depending on number of deployed gateway nodes

Table 8.4: Comparison of MPR and cycle-based TC in WBNs

	MPR Topology Control	**Cycle-based Topology Control**
Redundancy level (Edge connectivity χ of network)	formation of one-edge-connected networks is possible, histogram of χ has highly negative skew, higher mean value of χ, more frequent occurrence of optimum value $\chi = d_{\max}$, $\left(\frac{\partial \chi}{\partial d_{\max}}\right)_{\text{MPR}} > 0$, volatile results w.r.t. m in specific, unfavorable parameter settings	formation of one-edge-connected networks is excluded, histogram of χ has moderately negative skew, lower mean value of χ, $\chi = d_{\max} - 1$ is mode value, $\left(\frac{\partial \chi}{\partial d_{\max}}\right)_{\text{MPR}} > \left(\frac{\partial \chi}{\partial d_{\max}}\right)_{\text{cycle-based TC}} > 0$, stable results with $\frac{\partial \chi}{\partial m} > 0$, $\frac{\partial^2 \chi}{\partial m^2} < 0$
Network capacity c_G	no significant difference between MPR and cycle-based topology control, number and placement of gateway nodes is main determinant of capacity, $\frac{\partial c_G}{\partial g} > 0$, $\frac{\partial^2 c_G}{\partial g^2} < 0$, $\frac{\partial c_G}{\partial d_{\max}}$ is larger for higher values of m	

8.4 Performance Evaluation of Operations Algorithms

One of the major objectives when applying Backhaul Topology Optimization (BTO) algorithms to an energy-autarkic point-to-point radio network with non-permanent energy supply is to extend the lifetime of individual nodes as well as to provide carrier-grade Quality of Service (QoS) levels. Therefore, the optimum degree of connectedness of network nodes among each other (edge connectivity) needs to be maintained. The author has collected according simulation data and compared it to state-of-the-art TC algorithms, such as *PlainTC* [MNNA09]. In cases where TC algorithms have originally been designed for single-channel networks, minor adaptations were necessary to allow for a sensible application in multi-channel multi-radio environments. The analysis concentrates on the Alpine scenario since WBN deployments in this region are more prone to node failures due to drained batteries.

8.4.1 Context-aware Backhaul Topology Optimization

In this part of the analysis, the simulation configuration of the Alpine scenario, depicted in Table 8.3, is selected. As described in earlier sections, the critical aspect with respect to autarkic energy supply is whether network nodes can operate throughout the night. Further, in order to appropriately evaluate the performance gains of the BTO algorithm, the author selects a setup where batteries are likely to reach their cut-off charging level, i.e., a battery capacity of 20 Ah, a PV panel size of $0.5\,\mathrm{m}^2$, and $d_{\min} = 2$ is chosen for all nodes. Moreover, simulations cover a time span from dusk until dawn on a winter date. This allows for optimally measuring the lifetime extension effect achieved through topological adaptations [MCS13d].

Starting with the evaluation of physical domain aspects, Figure 8.21(a) illustrates the sequential failure of nodes if (adapted) *PlainTC* is applied to a point-to-point radio network. More specifically, the abscissa depicts simulated time in s, while the ordinate shows the fraction ζ of failed nodes, i.e.,

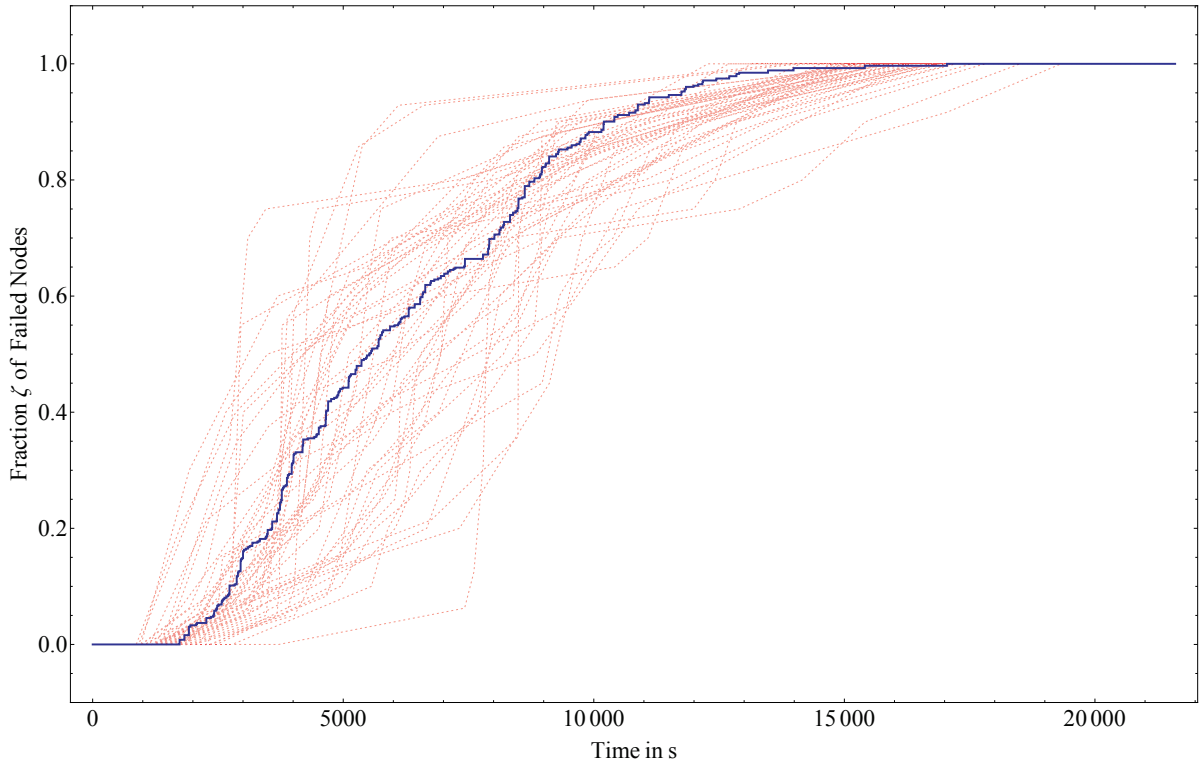
$$\zeta = \frac{\text{number of nodes out of operation due to drained batteries}}{\text{number of nodes with autarkic energy supply}}. \quad (8.2)$$

The evaluation has been carried out for 50 different network topology configurations, each of the dashed lines indicating the result for one of the con-

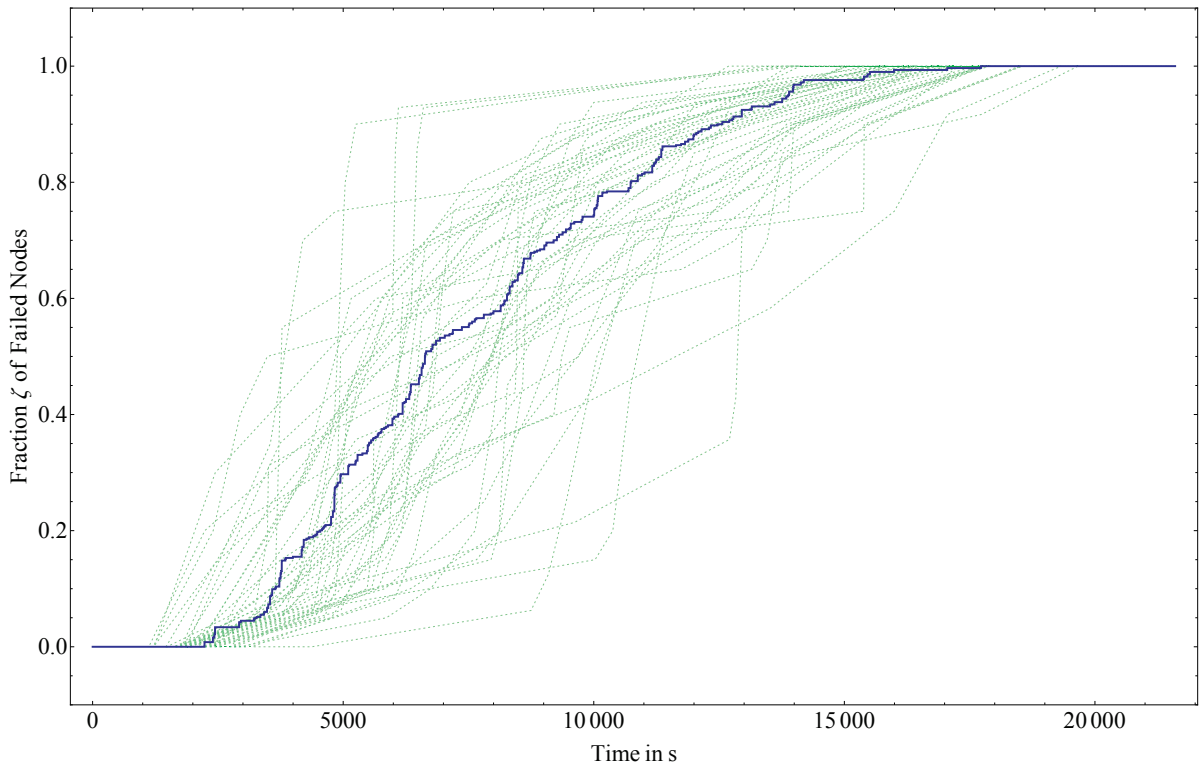
figurations. The bold blue line shows the average over all configurations. Horizontal deviation from that curve indicates to what extend networks vary in terms of reaching a given failure level. This so-called *temporal deviation* from the average is rather symmetrical and can be as high as approximately $6,000$ s. Vertical deviation denotes the difference in failure level between individual network configurations at a given time. In Figure 8.21(a), it can be as high as $\zeta = 0.6$ in either direction. Further, on average, first nodes fail at $t \approx 1,900$ s (with individual configurations varying between $1,900$ and $3,900$ s) and the fraction ζ of failed nodes reaches 1 at time $t \approx 15,500$ s (with individual configurations varying between $12,000$ and $19,000$ s).

Figure 8.21(b) depicts the according quantities for the *BTO* algorithm, based on the same set of 50 network topology configurations. Similar to the results of *PlainTC*, deviation from the average failure behavior is quite symmetrical: maximum vertical deviation amounts to approximately $\zeta = 0.6$, while temporal deviation is as high as approximately $7,000$ s. Moreover, first nodes fail in a window of [1,100 s, 4,500 s] (average value: $t \approx 2,200$ s), and $\zeta = 1$ is reached between $12,700$ s and $19,500$ s, depending on the network configuration (average value: $t \approx 17,800$ s).

Figure 8.22 illustrates a direct comparison of the average node failure CDFs. Further, it includes upper and lower bounds for node failure according to the power consumption analysis performed in Section 7.1. Due to the uniform distribution of initial battery charging level, the bounding curves exhibit a linear slope, indicating a successive failure of nodes. The slope (node failures per time interval) is steeper for the upper bound curve since, in this case, nodes are assumed to form the maximum possible number of links to neighboring backhaul nodes. The network operated according to BTO shows a consistently lower level of node failures at any point in time and is able to approach the lower bound curve. A closer approximation could be realized by further trading-off network capacity in favor of energy efficiency. Nevertheless, for a time span of more than 160 min, $\Delta\,\zeta$ falls into an interval of $[0.10, 0.21]$. In other words, with the exception of phases with almost completely drained batteries, BTO outperforms *PlainTC* by at least $\zeta = 0.10$. Accordingly, mean lifetime of a node improves from 104.1 min for a *PlainTC* network to 126.9 min for a BTO network. This is equivalent to an average improvement of node lifetime of 21.8%. Another important observation is that the average temporal improvement of lifetime increases over time, a result attributed to the fact that the BTO operation mode ex-



(a) Adapted *PlainTC*



(b) BTO

Figure 8.21: Empirical CDF of node failures [MCS13d]

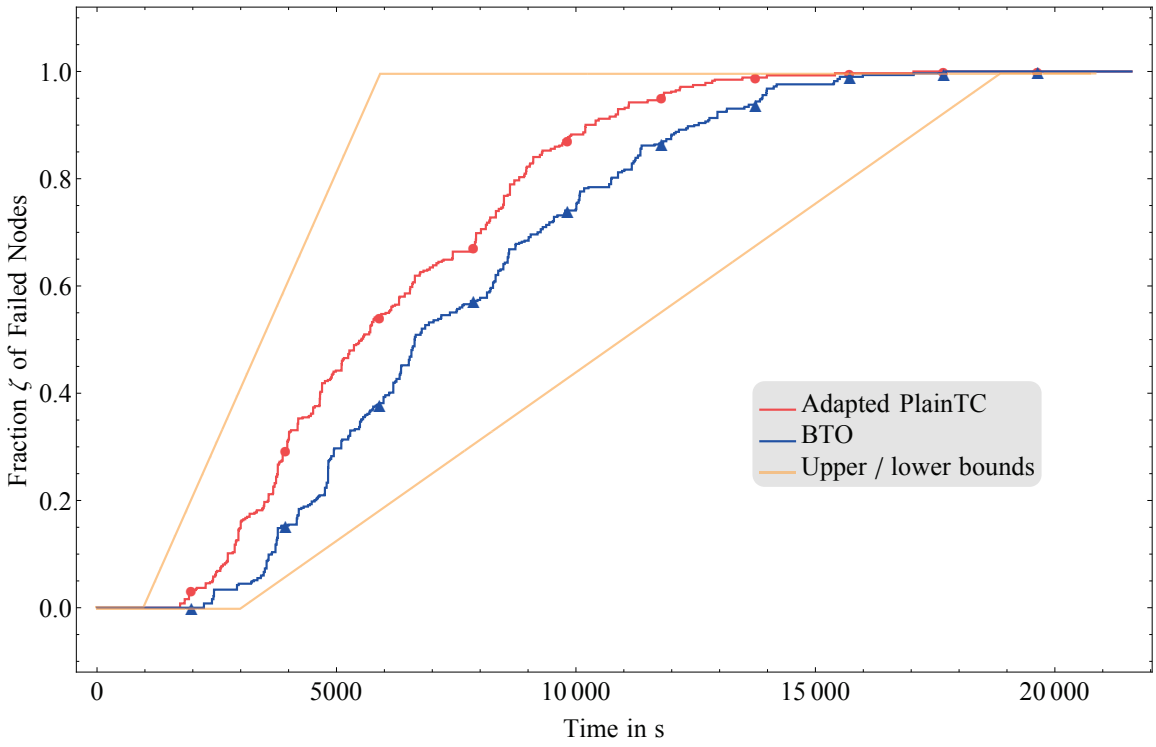


Figure 8.22: Comparison of empirical CDF of node failures [MCS13d]

hibits an accumulative effect by allowing nodes to continuously save energy for consumption in later instants of time.

For a further analysis of this accumulative effect, Figure 8.23 depicts the temporal difference Δt (labeled Lifetime Extension Measure (LEM), cf. Section 3.5.2) between a *PlainTC*-operated and a BTO-operated network of reaching a given node failure fraction ζ. Using a zeroth order of approximation, a given failure level ζ is reached almost $1,400$ s (i.e., more than 20 minutes) earlier in the *PlainTC* operation mode. Applying a first-order approximation, the temporal gap steadily increases, starting from around 500 s up to more $2,000$ s for $0.8 \leq \zeta \leq 1.0$. More specifically, the linear regression (red line) reveals that the benefit in lifetime extension Δt_{LEM} due to the BTO algorithm can be approximated by Equation (8.3)

$$\Delta t_{\mathrm{LEM}}/\mathrm{min} = 11.7 + 21.8\,\zeta, \quad \zeta \in [0,1]\,. \tag{8.3}$$

The described accumulative effect of the BTO operation results in a further advantage: besides extending the lifetime of nodes in general, it achieves an above-average lifetime extension in situations where only a low percentage of connected nodes is left. This is extraordinary beneficial if an elementary network infrastructure has to be provided as long as possible, e.g., in disaster relief scenarios. For an extended evaluation of the effectiveness of

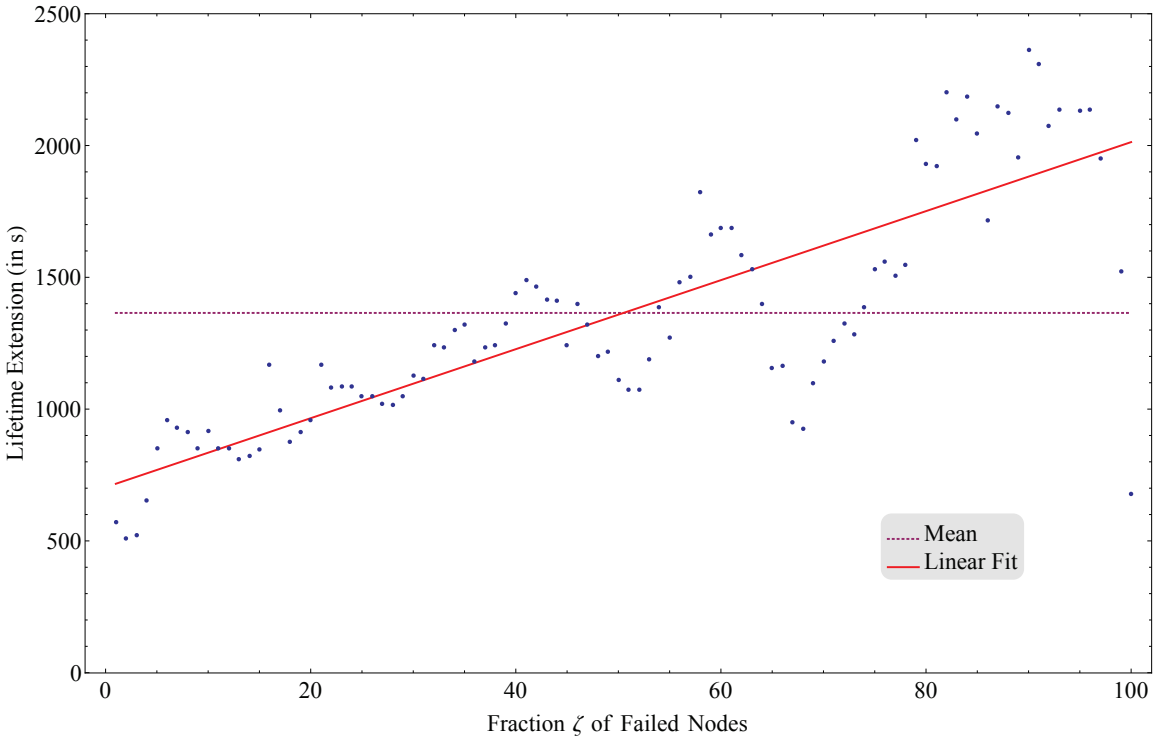


Figure 8.23: Lifetime extension of nodes as a function of fraction of failed nodes ζ [MCS13d]

BTO in the context of lifetime extension, the author has introduced upper and lower bounds for energy consumption and according lifetime extension bounds, cf. Section 7.1. Figure 8.24 depicts the gain G_{BTO} achieved by BTO and the gain $G_{\max}$ achieved by a (hypothetical) operation mode with minimal capacity levels and accordingly low power consumption. Again, the aforementioned 50 network configurations have been considered. As can be observed in Figure 8.24, G_{BTO} varies from a few hundreds up to approximately $3,000\,\mathrm{s}$ of improvement in network lifetime. Further, values for $G_{\max}$ lie between $5,000$ and $15,000\,\mathrm{s}$. From these figures, $\mathrm{LEM} = \frac{G_{\mathrm{BTO}}}{G_{\max}}$ can be calculated. Figure 8.25 outlines the result for each network topology configuration, sorted by increasing LEM. While the values vary between 0.02 to 0.27, the average improvement amounts to $\mathrm{LEM}_{\mathrm{avg}} \approx 14\,\%$, i.e., BTO on average achieves 14% of (theoretic) maximum lifetime extension $G_{\max}$.

For the network domain performance analysis, the impact of the proposed energy-conserving strategy on network performance is evaluated. The author focuses on analyzing user outage and average data throughput. For a VoIP service, a user is defined in outage if the average packet call throughput is less than the minimum average throughput requirement of 128 kbit/s [SZJ+08]. Figure 8.26 illustrates a histogram of aggregated outage prob-

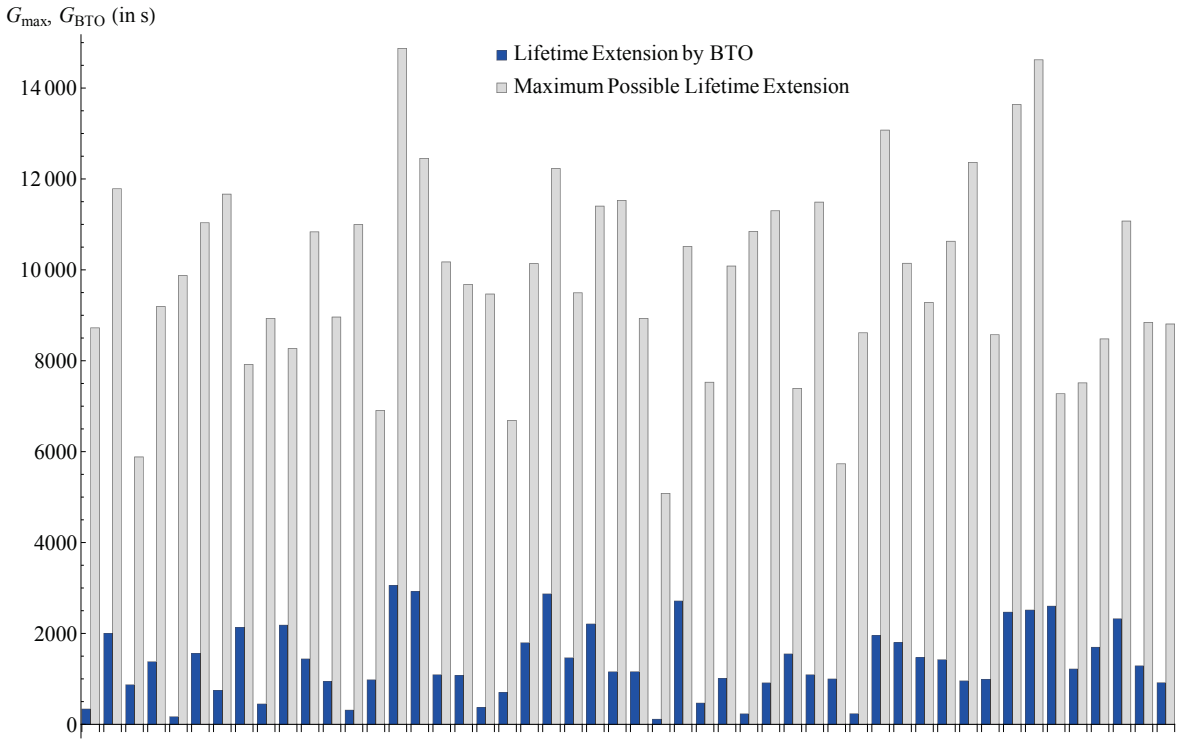


Figure 8.24: Maximum possible and realized lifetime extension [MCS13b]

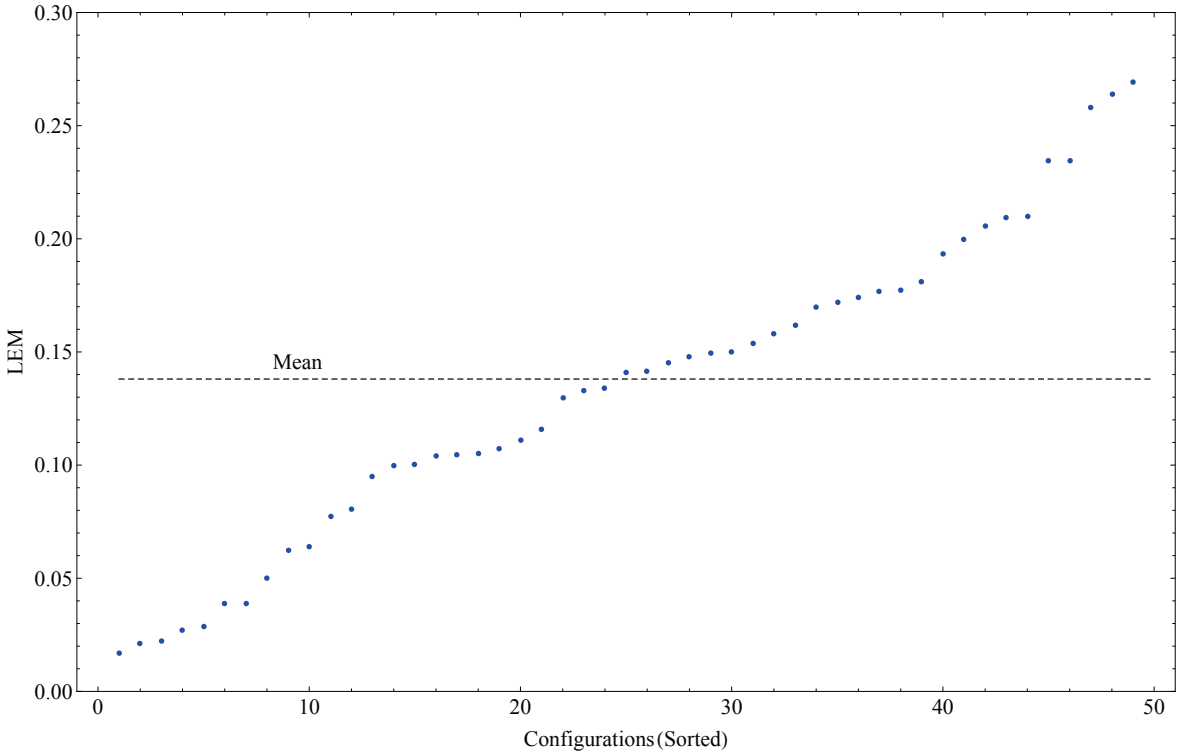


Figure 8.25: *LEM* for analyzed network configurations (sorted) [MCS13b]

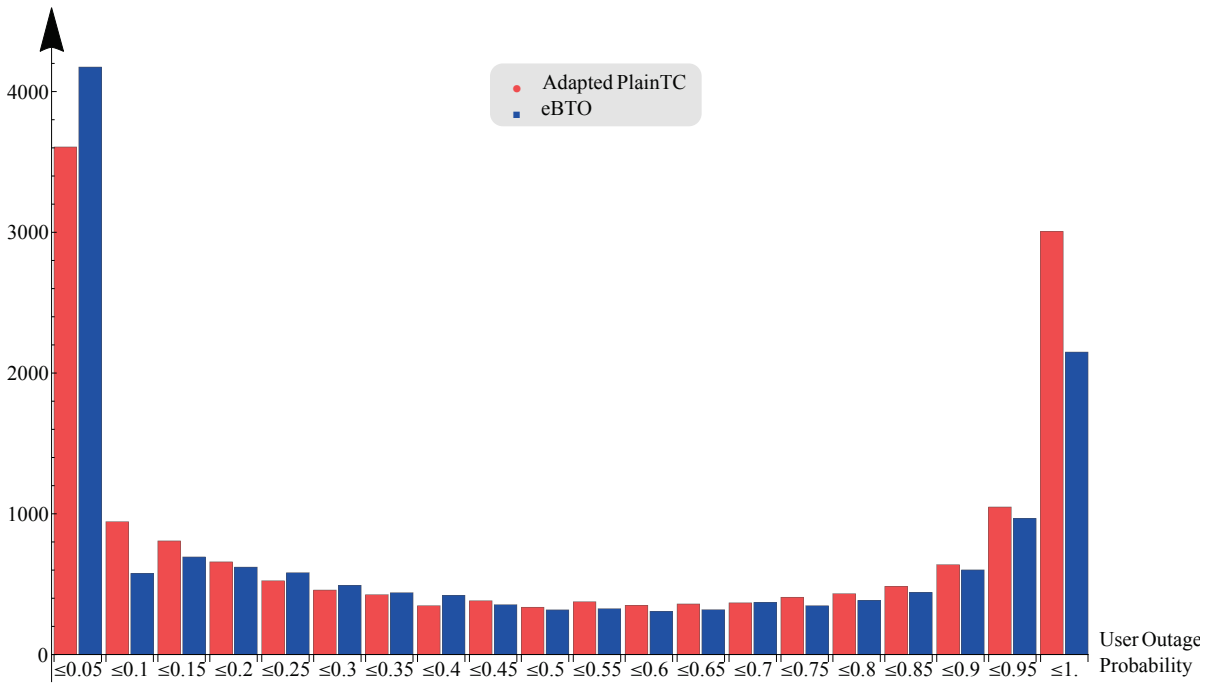


Figure 8.26: Histogram of user outage probability (aggregated over 50 configurations) [MCS13b]

abilities experienced by users across all 50 topology configurations. The histogram bins have a width of 5%, resulting in a total of 20 bins. For most of the bins, the two distributions only differ marginally, which indicates that there is no significant performance difference with respect to user outage. However, the share of users pertaining to the bin with lowest outage probability is significantly higher for the BTO-operated network. Accordingly, the opposite is true for the high outage probability bin, where users of the *PlainTC*-operated network form a significant majority. This overall shift of users to lower outage probabilities in the BTO scenario stems from two different characteristics: First, a certain percentage of users do not request a service until the last quarter of the simulated time, which, due to the high portion of failed backhaul nodes at that time, has a more profound impact on the *PlainTC*-operated than on the BTO-operated network. Second, *PlainTC* is more affected by the cascading effect. Here, the impact of nodes going out of operation due to drained batteries is reinforced if these nodes lie on any path that connects other nodes to a gateway node. In essence, the shutdown of the node results in cutting off of several others from the gateway node. Such a situation becomes more likely with increasing node failure fraction ζ, which is reached earlier in the case of *PlainTC*.

8.4.2 Extended Backhaul Topology Optimization

In the eBTO algorithm, a measure for the algorithm's impact on resource allocation fairness from the user perspective forms an explicit part of the objective function. Therefore, the author additionally evaluates Jain's fairness index as well as the Gini coefficient, which are computed for each of the 50 topology configurations [MCS13c]. This yields two fairness indicators across the entire base of users. However, network domain aspects are analyzed first by evaluating the effect of eBTO on network throughput. Figure 8.27 illustrates a histogram of average data rate experienced by active users (data rates during an ongoing data service were considered). Bin widths have not been dimensioned in a linear manner. Rather, Figure 8.27 depicts increasing bin widths, from 0.1 to 0.5 to 5 Mbit/s. For the eBTO-operated network, a general shift to higher data rates can be observed. Considerably fewer occurrences of very low data rates (≤ 0.2 Mbit/s) and a slight increase of users having a high average data rate (≥ 5 Mbit/s) stand out. This improvement mainly has to be attributed to the fact that backhaul nodes on average have a longer lifetime, thus servicing their respective access points for a longer time and providing higher mean data rates to users. Specifically, the average data rate provided to active users is at 3.23 Mbit/s for the *PlainTC*-operated and 3.76 Mbit/s for the eBTO-operated network, respectively. Hence, by balancing energy consumption more evenly across available nodes, eBTO does not only reduce power consumption but also improves user throughput and outage rates.

Figure 8.28 illustrates Jain's fairness index with respect to average user data rate for the *PlainTC*-operated and the eBTO-operated backhaul network. Moreover, the gap between the operation modes is highlighted by a vertical line. In two-thirds of the configurations, BTO outperforms *PlainTC*, in more than 10 cases by more than 0.2 points on the index's $[0, 1]$-scale. In contrast, for those cases that report a better rating of *PlainTC*, the difference generally is only marginal. Only in a small number of configurations, *PlainTC* receives a significantly higher rating (≥ 0.2 points) than the according eBTO-operated WBN. On average, the eBTO performs about 0.07 points better than *PlainTC*, which is an improvement of more than 18%.

Similar conclusions can be drawn from Figure 8.29. For each of the 50 configurations, it depicts the Gini coefficient with respect to average user data rate in either operation mode, *PlainTC* and eBTO. It is important to

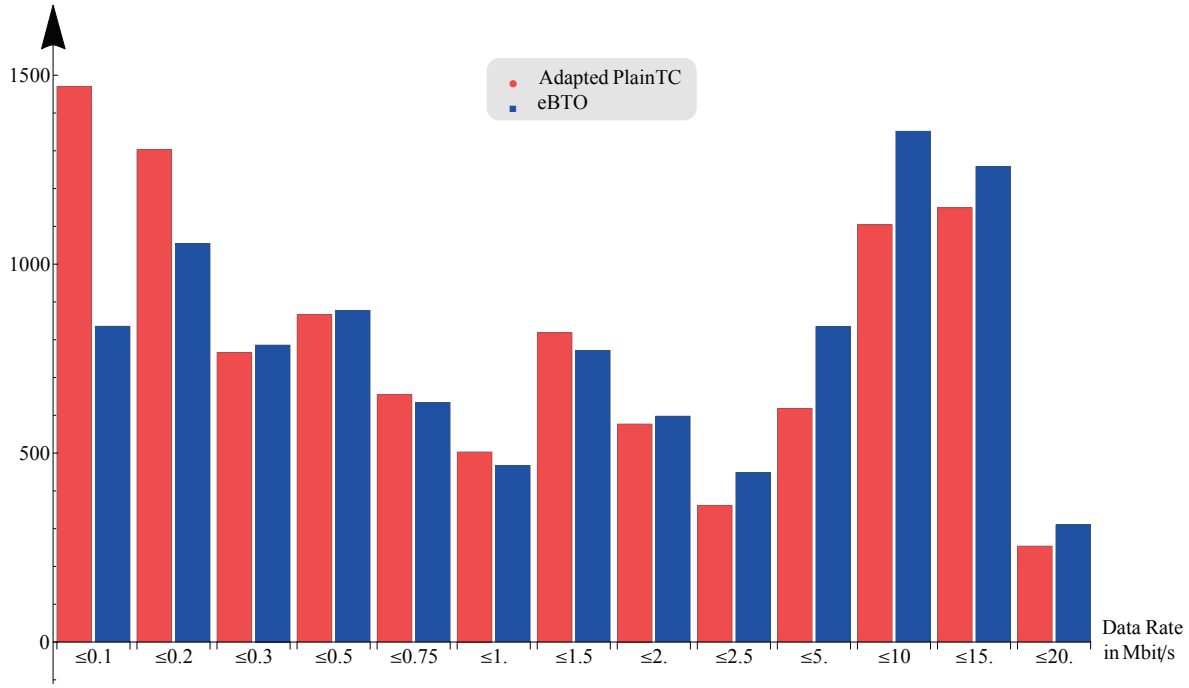


Figure 8.27: Histogram of average user data rate (aggregated over 50 configurations) [MCS13b]

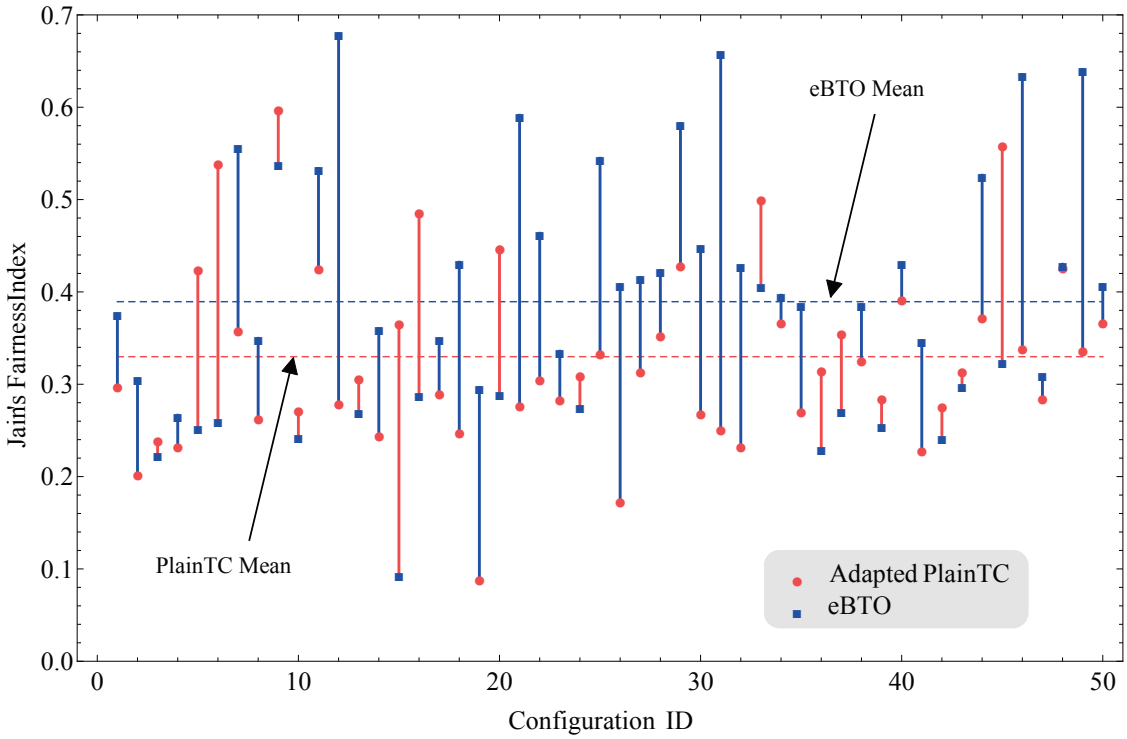


Figure 8.28: Evaluation of Jain's fairness index for *PlainTC*-operated and BTO-operated networks [MCS13c]

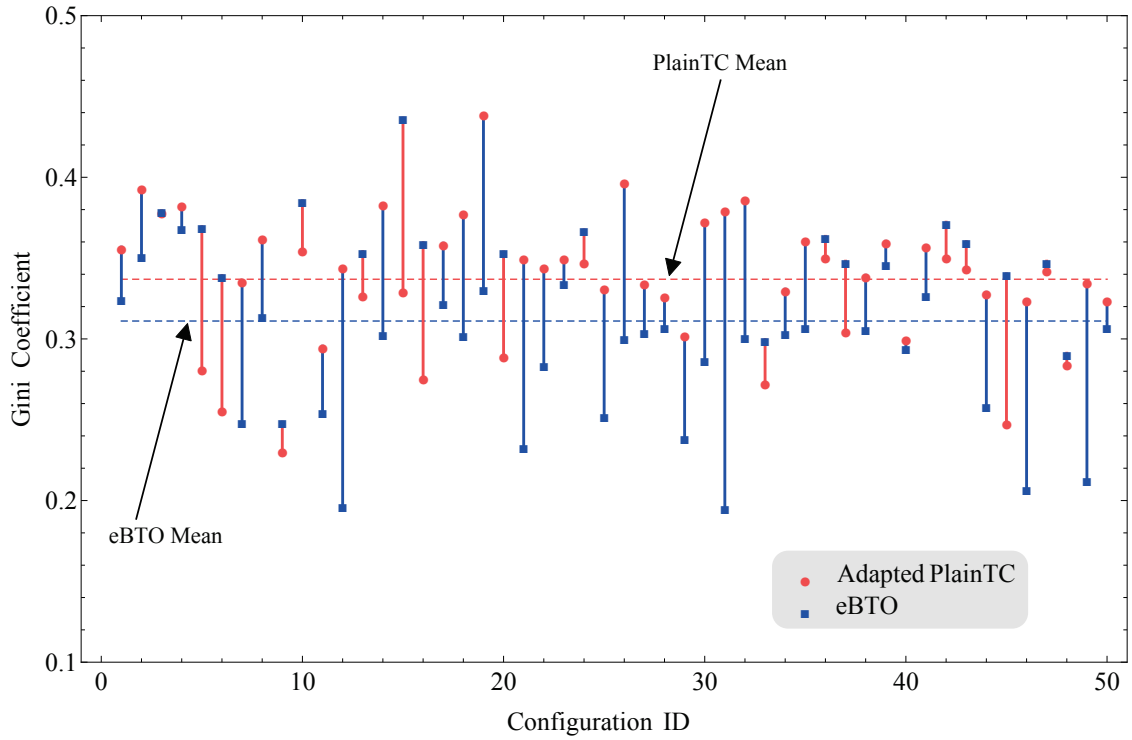


Figure 8.29: Evaluation of Gini coefficients for *PlainTC*-operated and BTO-operated networks [MCS13c]

bear in mind that, in contrast to Jain's fairness index, lower values of the Gini coefficient indicate a fairer distribution of available resources. While overwhelmingly confirming the evaluation obtained by Jain's fairness index, the Gini coefficient exhibits some different results in individual head-to-head comparisons. This particularly holds for some comparisons with tight results (as, e.g., for configurations 47 and 48), where eBTO is considered fairer by Jain's index, and *PlainTC* by the Gini coefficient. This is mainly attributed to the fact that the two metrics use mutually inverse scales. Nevertheless, Figure 8.29 denotes a clear superiority of eBTO over *PlainTC*, in several cases by more than 0.1 points. An according result is achieved only once by a *PlainTC*-operated network (configuration 15). Therefore, on average, eBTO outperforms *PlainTC* by 0.03 points or almost 10%.

8.4.3 Sensitivity Analysis

A sensitivity analysis with respect to parameters related to energy supply and demand (PV panel size, battery capacity, node power consumption) for both scenarios has already been performed in Section 8.2. In the following, the sensitivity of average user throughput is analyzed. This includes questions like "how does user throughput vary with gateway density or terminal

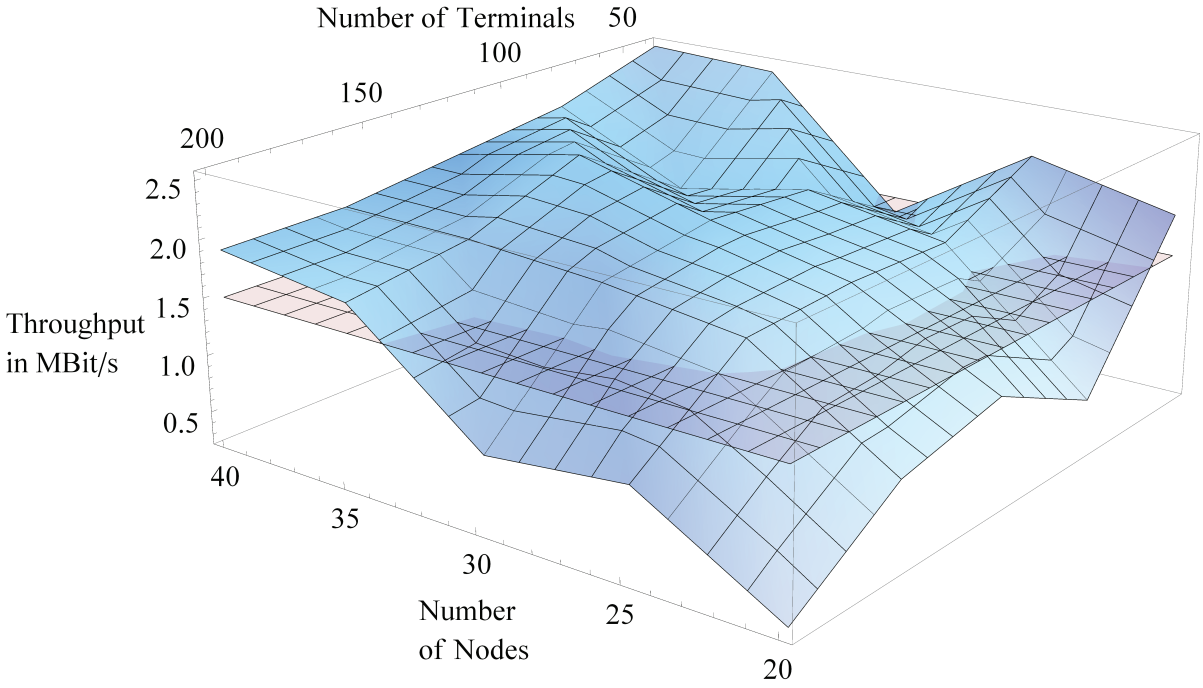
population. For this purpose, the following network deployment parameters have been varied within selected boundaries [MCKS13]:

- number m of backhaul nodes,
- number N_{UT} of user terminals,
- ratio $\frac{m}{N_{UT}}$ of nodes to terminals,
- ratio $\frac{m_{\mathrm{GW}}}{N_{UT}}$ of gateway nodes to terminals,
- ratio $\frac{m_{\mathrm{GW}}}{m}$ of gateway nodes to backhaul nodes.

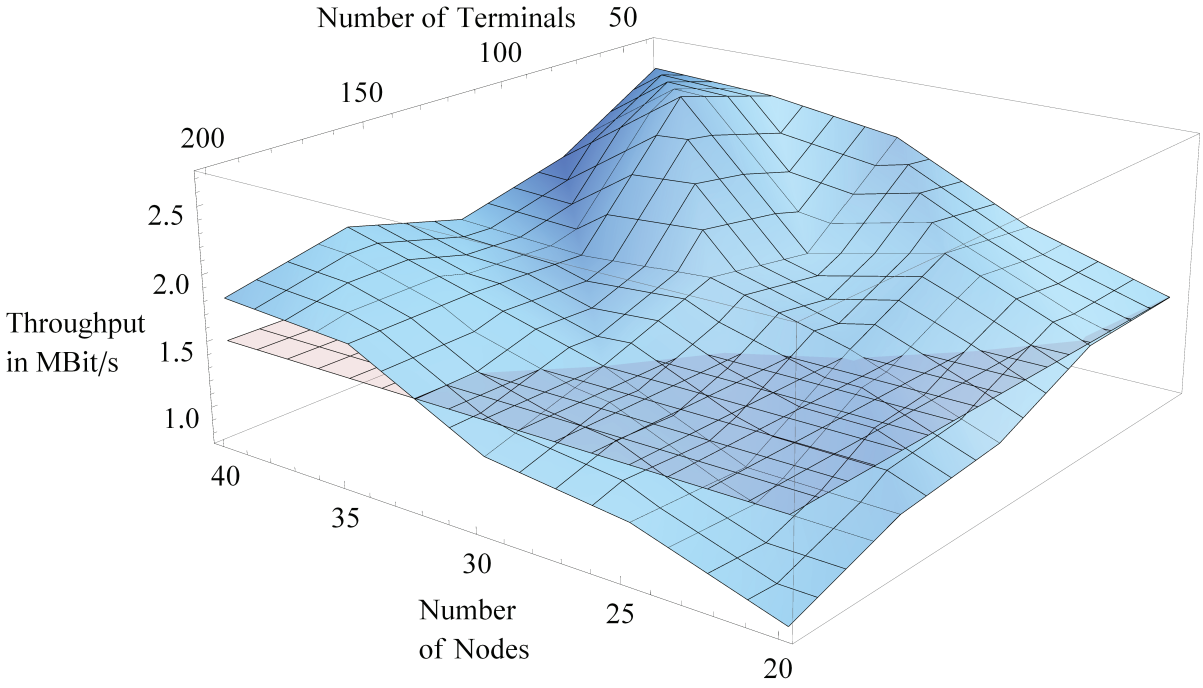
Other input variables, such as user mobility, data traffic, number of backhaul radio interfaces per node, or radio access network configuration have remained fixed [MCKS13].

Figure 8.30(a) depicts average data throughput $\overline{\nu}$ depending on user and node population. The horizontal plane is placed at 1.5 Mbit/s. In the analyzed network configurations, gateway nodes have a share of 10% among all nodes. The figure shows some typical characteristics of a meshed point-to-point structure. For a fixed number of users, an increasing number of nodes improves the average data rate per user. In case of a fixed number of nodes, an increasing number of users significantly reduces the average data rate since more users have to share a fixed backhaul capacity. These general characteristics apply across the entire range of input variables. Overall, observed mean throughput is $\overline{\nu} = 1.67$ Mbit/s, varying from $\overline{\nu} \approx 0.2$ Mbit/s for a ratio of nodes to users of $\frac{m}{N_{UT}} = 0.1$ to $\overline{\nu} \approx 2.5$ Mbit/s for $\frac{m}{N_{UT}} = 1.25$. Moreover, it can be observed that, in order to preserve a minimum average data rate of $\overline{\nu} = 1.5$ Mbit/s per user, $\frac{m}{N_{UT}}$ should not fall below a threshold of approximately 18 to 20%. In Figure 8.30(b), according results for WBNs exhibiting a gateway share of 20% among all nodes are depicted. The general characteristics of average throughput are the same. However, fewer outliers are observed and the intersection line where the 1.5 Mbit/s plane is cut is more regular. These observations are attributed to the fact that a higher percentage of gateway nodes results in a smoother distribution of gateway capacity. In the case of 20% gateway node share, the critical ratio $\frac{m}{N_{UT}}$ required for exceeding an average throughput of 1.5 Mbit/s lies at 15 to 17%. Moreover, average throughput across all parameter constellations slightly increases to $\overline{\nu} = 1.70$ Mbit/s.

Figure 8.31(a), which shows average throughput over node-to-user ratio $\frac{m}{N_{\mathrm{UT}}}$ ($\frac{m_{\mathrm{GW}}}{m} = 0.1$), indicates the incremental benefit of additional backhaul nodes in terms of average user throughput. Simulation results are fitted using a logarithmic function, as indicated by the straight blue line.

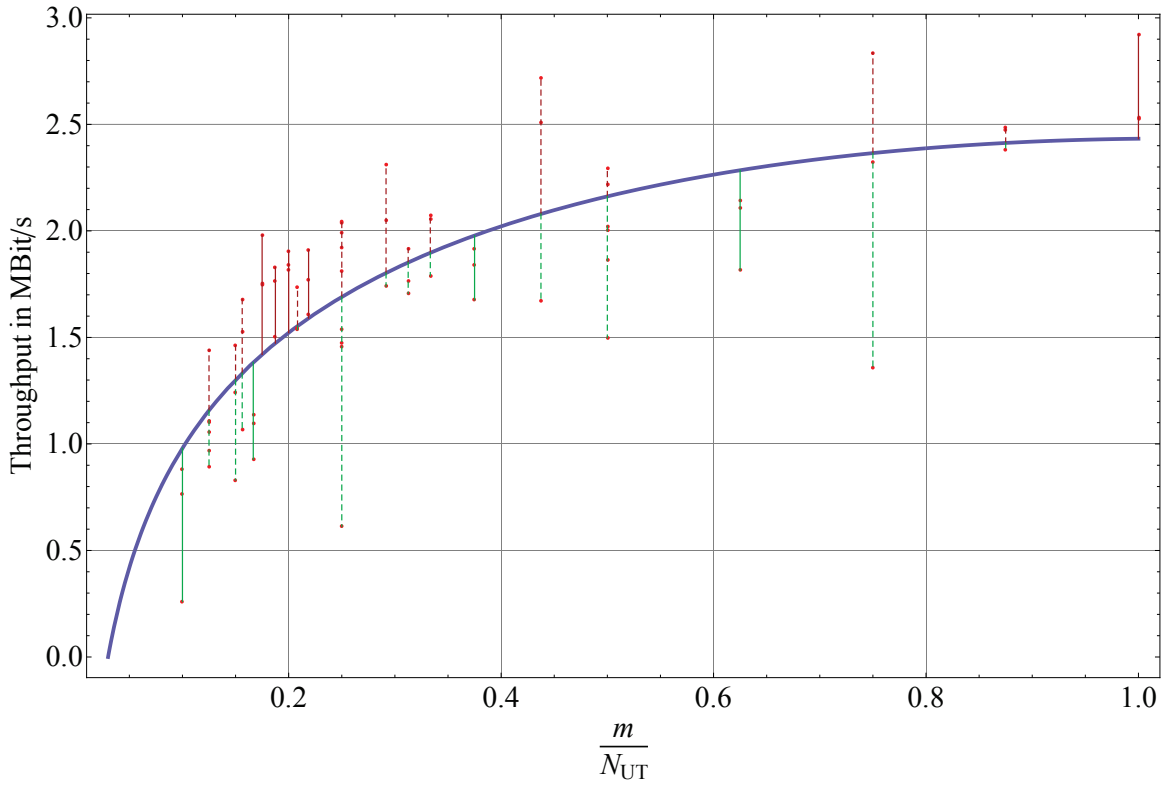


(a) 10% gateway nodes

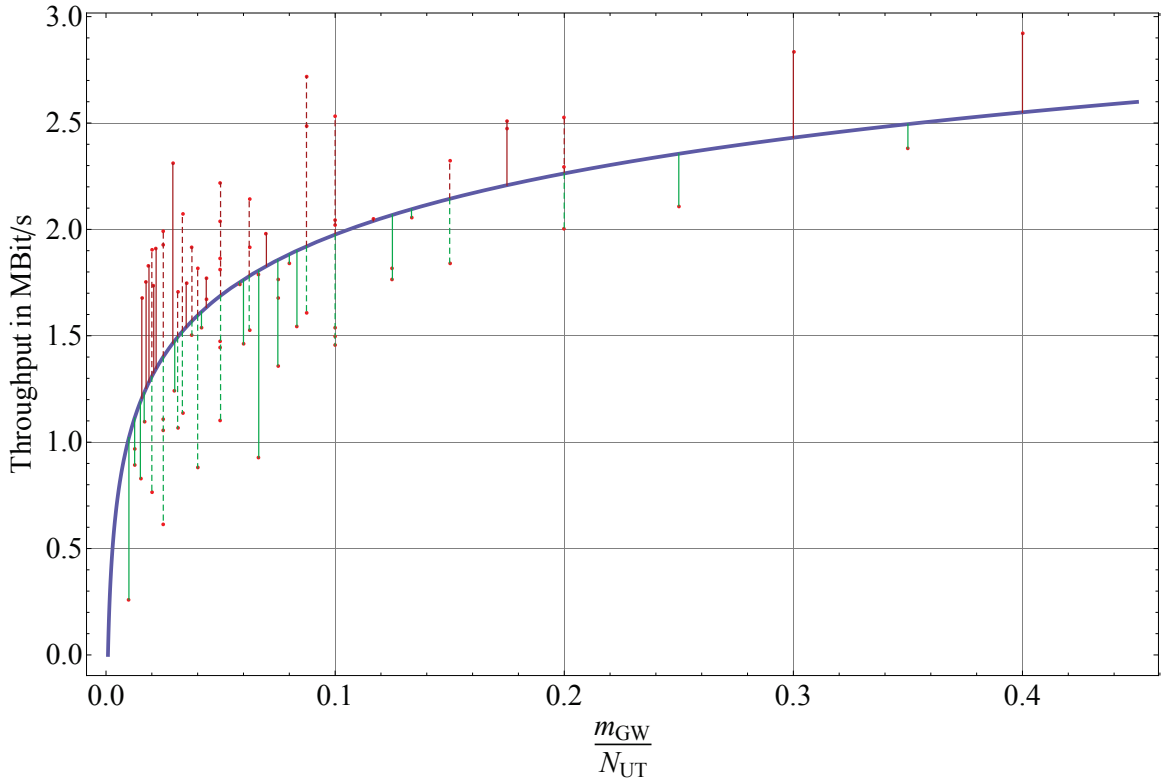


(b) 20% gateway nodes

Figure 8.30: Sensitivity of user throughput for a fix share of gateways



(a) Depending on backhaul nodes per user



(b) Depending on gateway nodes per user

Figure 8.31: Sensitivity of user throughput depending on nodes per user

The graph goes into saturation when reaching a percentage of $\frac{m}{N_{UT}} \approx 0.6$. A further increase in the number of nodes only yields marginal returns. Moreover, as observed in Figure 8.30(a), the plot confirms the minimum ratio of $\frac{m}{N_{UT}} \approx 0.19$ for exceeding an average throughput of $\overline{\nu} = 1.5\,\mathrm{MBit/s}$. Similarly, Figure 8.31(b) depicts the benefit of increasing the density of gateway nodes. An initial increase in the number of gateways per user dramatically improves average user throughput. However, when reaching a value of $\frac{m_{\mathrm{GW}}}{N_{UT}} \approx 0.3$, saturation settles in. Moreover, even in the interval $0.1 \leq \frac{m_{\mathrm{GW}}}{N_{UT}} \leq 0.2$, incremental gains already reduce significantly. This confirms the low improvement in overall mean data rate as observed in Figure 8.30.

In summary, the analysis shows that an increase of the absolute number of both backhaul and gateway nodes improves total throughput. However, the analysis of normalized quantities, such as number of (gateway) nodes per user or share of gateway nodes among all nodes, reveals that only marginal returns can be expected as soon as certain thresholds have been passed.

8.5 Summary

Chapter 8 has performed a multi-domain numerical evaluation of the concepts and algorithms developed in this work. For this purpose, a simulation tool following the principles outlined in the evaluation methodology presented in Section 3.3 has been developed. It appropriately models the operation of an energy-autarkic WBN by incorporating state-of-the-art concepts from wireless networking, energy supply and consumption, as well as context-awareness. Using this simulation tool, the author has analyzed the requirements, opportunities, and constraints of autarkic, PV-based energy supply, in particular PV panel sizes and battery capacities necessary for a stable and continuous operation in the given deployment scenarios. Further, Topology Control (TC) algorithms for WBN deployment and operation have been evaluated with respect to their impact on KPIs from different domains. It has been demonstrated that the deployment algorithms can effectively control parameters such as edge connectivity of a network graph and realize high network capacity ($s-t$-flows) despite varying density of gateway nodes. Further, the developed BTO algorithms have shown to significantly reduce energy consumption on node and system level, while QoS remains at virtually equivalent levels. As a further benefit, BTO has proven to delay battery discharging, therefore guaranteeing extended node

lifetime. As a matter of fact, fairness of resource allocation among users does not suffer from these measures, either. Finally, a sensitivity analysis comparing network capacity and user throughput for varying ratios of backhaul nodes, gateways, and users has yielded reasonable operating points for these quantities.

Chapter 9

Conclusion

As of today, suburban, rural, and topographically challenging areas in industrialized as well as developing countries have been frequently suffering from the limited availability of broadband Internet access. In light of the rapidly growing popularity of mobile Internet and online services in general, this does not only affect end users. For companies, particularly Small- and Medium-size Enterprises (SMEs), such conditions constitute considerable competitive disadvantages if important internal processes or distribution channels rely on broadband network connectivity. In extreme cases, further implications include the loss of attractive and sustainable jobs, degradation of purchasing power, migration of young and well-educated people to urban areas, drop of tax revenues of communities, decline of public and private infrastructure and buildings, as well as a shortfall in cultural activities. In this context, coordinated, point-to-point Wireless Backhaul Networks (WBNs) are a promising, cost-efficient solution for provisioning broadband Internet connectivity to sparsely populated areas. This work has addressed and resolved some of the major challenges related to the deployment and operation of this class of wireless networks.

9.1 Summary

Mobile network operators currently find themselves exposed to rapid technological developments and dynamic market shifts. With respect to the aforementioned problems, they face the challenge of profitably providing broadband connectivity in rural regions. Lower overall revenues due to a smaller subscriber base coincide with high investments for new network equipment, thus increasing the inhibition threshold for investing in latest technologies. Two examples include operators' hesitation to deploy optical

fiber in a comprehensive manner and the German government's efforts to commit operators to supply LTE-A radio access networks in selected rural areas.

In order to overcome these problems, the author has developed the notion of coordinated WBNs as a promising approach to cost-efficiently provide broadband network connectivity. Using point-to-point radio technologies, WBNs frequently rely on standards such as IEEE 802.11 or microwave implementations. However, while realizing enormous cost savings due to the utilization of cheaper hardware, this class of networks exhibits significant performance gaps when compared to carrier-grade networks. Quality of Service (QoS) and reliability levels do not yet match the performance of operator networks, particularly in case of autarkic energy supply (e.g., from PhotoVoltaic (PV) panels). Therefore, numerous challenges need to be resolved before a wide-spread market launch becomes realistic. In this work, the author has addressed and formalized a selection of problems in the area of energy-autarkic wireless backhauling and made the following contributions.

Formal Problem Statement and Evaluation Methodology

The author has established the category of energy-autarkic, coordinated WBNs as a point-to-point multi-radio, multi-channel evolution of Wireless Mesh Networks (WMNs) where a significant share of backhaul nodes are not connected to the power grid but rather possess alternative power supply equipment, such as PV modules or wind turbines. Subsequently, limitations and challenges of deployment and operation of WBNs have been summarized. Besides being subject to minimum performance requirements, such networks simultaneously face serious constraints with respect to their energy balance. In other words, deployment and operation cope with interdependencies in the physical, network, and context domains. In order to formally depict according trade-offs and optimally balance multiple operational objectives, the author has defined these domains and constructed an extensible formal multi-domain problem statement. A novel evaluation methodology developed by the author serves as a guideline to efficiently evaluate the impact of operation algorithms on network performance. The methodology has been implemented in a system-level simulation tool where various scenarios can be reproduced. Within this tool, concepts from graph theory are used to design an extended model for WBNs that is able to incorporate requirements for carrier-grade performance, energy balance con-

straints, and available context information (e.g., on user movements). This ensures an overall system model close to reality.

Design of a Versatile Context Management Platform
The collection, processing, and distribution of heterogeneous contextual data allowing for a context-enabled operation of energy-autarkic WBNs is realized by a context management platform. Within the scope of this thesis, functional concepts and protocols for a scalable and extensible context management architecture have been designed and implemented. According work of the author has significantly contributed to the development of the context management platform developed in the European research project *C-Cast* [CC10]. Moreover, the author has further refined the underlying producer-consumer model for a seamless embedding of the platform in various areas, among them radio access networks [MKSS10], Internet of Things [SKMS11], indoor and outdoor localization [MRS+10], as well as security and privacy [MKSS11b].

Topology Control (TC) Enabling Carrier-grade Networking and Management of Energy Balance
Based on the general problem statement, several algorithms addressing TC in WBNs have been developed. For initial network deployment, two novel algorithms, Maximum Path Redundancy (MPR) and Cycle-based TC, are presented. In graph theoretical terms, they address the minimum cost k-edge-connected spanning subgraph problem and the Quadratic Assignment Problem (QAP). While the former focuses on maximizing edge connectivity (redundancy levels) and capacity of the network graph, the latter aims at an efficient usage of radio interfaces by reasonably trading off graph connectedness against hop count between nodes. With respect to network operation, the author has designed two Backhaul Topology Optimization (BTO) algorithms implementing extensible multi-objective optimization problems. Context-aware BTO maximizes the weighted objectives of high $(s-t)$-flow (i.e., network capacity) and extended node lifetime in the light of several scenario-specific constraints for energy supply, irradiation conditions, network topology, as well as user service requests and mobility. The extended Backhaul Topology Optimization (eBTO) algorithm additionally considers fairness of resource allocation among users.

Analytical and Quantitative Evaluation

With respect to the analytical evaluation of the developed algorithms, the author has computed upper and lower bounds for power consumption and derived according analytical expressions for minimum and maximum lifetime of nodes with autarkic energy supply. In this context of reliable network operation in case of self-sufficient energy supply, required minimum battery capacities and solar panel areas have been determined depending on the variables geographical latitude, typical irradiation levels and weather conditions, as well as PV module characteristics. Moreover, symbolic approximation levels for the Maximum Path Redundancy (MPR) and for the cycle-based TC algorithm have been derived. Comparing the respective values of the defined objective function, they are able to approach the optimal solution by at least $\frac{2}{d_{\max}}$, where $d_{\max}$ is the maximum node degree (generally, $d_{\max} = 3$). Analytical expressions have further been used as a benchmark for the quantitative evaluation performed by means of the aforementioned simulation tool. For example, the evaluation has confirmed analytical results for PV panel sizes and battery capacities necessary for a stable and continuous operation in varying deployment scenarios. Further, TC algorithms for WBN deployment and operation have been evaluated with respect to their impact on KPIs from different domains. It has been demonstrated that the deployment algorithms effectively control parameters, such as edge connectivity of a network graph and realize high network capacity ($(s-t)$-flow) despite varying density of gateway nodes. Further, the developed BTO algorithms have shown to significantly reduce energy consumption on node and system level, while Quality of Service (QoS) remains at virtually equivalent levels. As a further benefit, BTO has proven to delay battery discharging, therefore extending node lifetime on average by more than 21%. As a matter of fact, fairness of resource allocation has also been improved by 10 to 18%, depending on utilized metrics. Finally, a sensitivity analysis that compares network capacity and user throughput for varying ratios of backhaul nodes, gateways, and users has yielded reasonable operating points for these quantities.

Overall, the suitability of energy-autarkic WBNs for selected deployment scenarios in topographically challenging regions with unreliable (nonexisting) supply from power grid has been shown. This also comprises a major contribution towards the deployment of carrier-grade point-to-point radio networks for providing broadband Internet in rural areas. Further, the author's work, particularly the analyses on energy efficiency, lifetime

of backhaul nodes with autarkic energy supply, and network redundancy levels, contributed to the research project SolarMesh [Sol14], funded by the German Ministry for Eduction and Research. In this context, the analytical and numerical evaluations of this work have supported the development of a wireless backhaul network infrastructure for provisioning carrier-grade services to rural areas in Germany and the realization of a large-scale testbed including nodes with autarkic energy supply in the Rhein-Sieg region near Bonn.

9.2 Outlook

The aforementioned conceptual, architectural, algorithmic, and quantitative contributions of the author have yielded a number of important findings with respect to an optimized deployment and operation of energy-autarkic point-to-point WBNs. However, several of the limitations of this network class have not been in the scope of this work and remain to be solved.

- *Carrier-grade network performance*: In terms of network-domain performance analysis, this work has focused on overall network capacity, throughput per user, network graph connectedness (i.e., path redundancy levels), and fairness of resource allocation among users. Here, coordinated WBNs have shown satisfactory to very good performance levels. However, the QoS aspect of latency needs further research and evaluation efforts. In particular, the general applicability and performance of different routing and switching protocols need to be analyzed. According investigations in similar contexts show that routing on network layer adds too much latency to packet delivery, so that switching protocols, such as, Multi Protocol Label Switching (MPLS) or Multi Protocol Label Switching - Traffic Engineering (MPLS-TE), are the preferred option for packet forwarding in a backhaul network.

- *Autarkic energy supply*: Due to the considered deployment scenarios, the analysis performed in this work is restricted to power supply from PV modules. Since it considers the full range of possible irradiation levels at the selected scenario locations, from dark winter skies to sunny days in summer, it has also determined equipment dimensions guaranteeing a 100% continuous operation. For such an ultra-reliable recharging of batteries, this has in parts yielded large, potentially

uneconomic PV panel sizes, thus disclosing the drawbacks of such power supply solutions.

- *General applicability in practice*: Since the concepts and algorithms of this work could be tested in actual testbeds only to a very small extent, they have been restricted to analytical and simulation-based evaluation for the major part. It will be an important task to verify developed algorithms and achievable performance gains in real-world WBN deployments. Although performance gains have been demonstrated and proven on a theoretical level and generally apply, operational, situation-specific validation usually raises practical questions and therefore is mandatory for final success.

In line with these limitations, further research activities are conceivable. Numerous switching/routing protocols for wireless mesh networks exist that specifically address the issue of high and fluctuating latency in multi-hop networks. The fundamental question is whether they (or potential enhancements) can reliably achieve carrier-grade latency and jitter levels. Further, OpenFlow switching [Ope12], which meanwhile has been introduced by some manufacturers, might be a viable solution for WBNs since it increases the flexibility to execute topological changes and effectively introduces Software-Defined Networking (SDN) to point-to-point radio networks. With respect to the implications of self-sufficient power supply, it would be interesting to perform evaluations on requirements for wind turbines. Particularly, if they can economically facilitate a reliable operation. For example, this could be an interesting option for coastline deployments, such as the Helsinki archipelago, the Frisian Islands, or the Stockholm skerries. Finally, while this thesis has laid theoretical foundations, further analyses require sufficiently large testbeds to perform extensive tests of the developed algorithms. A successful validation performed with real, tangible radio equipment constitutes the last step in establishing energy-autarkic coordinated wireless backhaul networks as an accepted solution for wide area networking in rural or topographically challenging areas.

Own Publications

[CMK+10] Carrella, S.; Mannweiler, C.; Klein, A.; Schneider, J.; Schotten, H. D.: A Concept for Context-aware Multihoming with Heterogeneous Radio Access Technologies. *Proceedings of the 5th International Wireless Internet Conference (WICON)*, 2010.

[CMS13] Chakraborty, P.; Mannweiler, C.; Schotten, H. D.: Evolutionary Approach for Multi-objective Optimization of Wireless Mesh Networks. *Proceedings of the 9th IEEE International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2013, pp. 36–40.

[KLM+11] Klein, A.; Lottermann, C.; Mannweiler, C.; Schneider, J.; Schotten, H. D.: A Novel Approach for Combined Joint Call Admission Control and Dynamic Bandwidth Adaptation in Heterogeneous Wireless Networks. *Proceedings of the 7th EURO-NF Conference on Next Generation Internet (NGI 2011)*, 2011.

[KMSS10a] Klein, A.; Mannweiler, C.; Schneider, J.; Schotten, H. D.: A Concept for Context-Enhanced Heterogeneous Access Management. *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, 2010.

[KMSS10b] Klein, A.; Mannweiler, C.; Schneider, J.; Schotten, H. D.: Access Schemes for Mobile Cloud Computing. *Proceedings of the 11th International Conference on Mobile Data Management (MDM)*, 2010.

[KMSS11] Klein, A.; Mannweiler, C.; Schneider, J.; Schotten, H. D.: An Advanced System Concept for Cognitive Spectrum Management and Utilization. *Proceedings of the 10th Workshop "Electrical and Electronic Engineering for Communication" (EEEfCOM)*, 2011.

[MAS+10] Mannweiler, C.; Amann, B.; Schneider, J.; Klein, A.; Schotten, H. D.: A Distributed Context Service Architecture for Heterogeneous Radio Environments. *Proceedings of the 15th "ITG Fachtagung Mobilkommunikation"*, 2010.

[MCKS13] Mannweiler, C.; Chakraborty, P.; Klein, A.; Schotten, H. D.: SolarMesh - Deployment Aspects for Wireless Mesh Networks in Developing Countries. Jonas, Karl; Rai, Idris A.; Tchuente, Maurice (eds.): *e-Infrastructure and e-Services for Developing Countries*. vol. 119 in series *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*. pp. 105–114. Springer, Berlin, Heidelberg„ 2013.

[MCS13a] Mannweiler, C.; Chakraborty, P.; Schotten, H. D.: A Novel Lifetime Extension Measure for Optimized Energy-Autonomous Wireless Backhaul Networks. *Proceedings of the 10th International Symposium in Wireless Communication Systems (ISWCS)*, 2013, pp. 1–5.

[MCS13b] Mannweiler, C.; Chakraborty, P.; Schotten, H. D.: Benefits of Topology Reconfiguration in Autonomous Wireless Backhaul Networks. *Proceedings of the 18th "ITG Fachtagung für Mobilkommunikation"*, 2013.

[MCS13c] Mannweiler, C.; Chakraborty, P.; Schotten, H. D.: Multi-Objective Adjacency Matrix Optimization for Coordinated Wireless Backhaul Networks. *Proceedings of the 24th IEEE International Symposium on Indoor, Mobile and Radio Communications (PIMRC)*, 2013.

[MCS13d] Mannweiler, C.; Chakraborty, P.; Schotten, H. D.: Pareto-Optimal Topologies for Lifetime Extension of Coordinated Wireless Point-to-Point Networks. *Proceedings of the IEEE 78th Vehicular Technology Conference (VTC 2013-Fall)*, 2013.

[MCS14] Mannweiler, C.; Chakraborty, P.; Schotten, H. D.: On the Reliability of Solar-Powered Point-to-Point Radio Backhaul Networks. *Proceedings of the 25th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2014.

[MKSS09a] Mannweiler, C.; Klein, A.; Schneider, J.; Schotten, H. D.: Exploiting User and Network Context for Intelligent Radio Network Access. *Proceedings of the International Conference on Ultra Modern Telecommunications (ICUMT)*, 2009, pp. 1–6.

[MKSS09b] Mannweiler, C.; Klein, A.; Schneider, J.; Schotten, H. D.: Integration von Kontextmodellen für intelligenten Funknetzzugang. *Proceedings of the 14th "ITG Fachtagung Mobilkommunikation"*, 2009.

[MKSS10] Mannweiler, C.; Klein, A.; Schneider, J.; Schotten, H. D.: Context Awareness for Heterogeneous Access Management. *Advances in Radio Science*, vol. 8, 2010, pp. 257–262.

[MKSS11a] Mannweiler, C.; Klein, A.; Schneider, J.; Schotten, H. D.: Context-based User Grouping for Multi-Casting in Heterogeneous Radio Networks. *Advances in Radio Science*, vol. 9, 2011.

[MKSS11b] Mannweiler, C.; Klein, A.; Schneider, J.; Schotten, H. D.: Privacy and Access Right Classification for Mobile Personal Media Environments. *Proceedings of the 5th Essen Workshop on Network Security (EWNS)*, 2011.

[MLK+12] Mannweiler, C.; Lottermann, C.; Klein, A.; Schneider, J.; Schotten, H. D.: Cyber-physical Networking for Wireless Mesh Infrastructures. *Advances in Radio Science*, vol. 10, no. 7, 2012, pp. 113–118.

[MLKS12] Mannweiler, C.; Lottermann, C.; Klein, A.; Schotten, H. D.: SolarMesh - Energy-efficient, Autonomous Wireless Networks for Developing Countries.

Popescu-Zeletin, R.; Jonas, K.; Rai, I. A.; Glitho, R.; Villafiorita, A. (eds.): *e-Infrastructure and e-Services for Developing Countries*. vol. 92 in series *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*. pp. 106–115. Springer, Berlin, Heidelberg, 2012.

[MMS10] Moltchanov, B.; Mannweiler, C.; Simoes, J.: Context Awareness Enabling New Business Models in Smart Spaces. *Proceedings of the 3rd Conference on Smart Spaces and 10th International Conference on Next Generation Wired and Wireless Networking (ruSMART/NEW2AN)*, Berlin, Heidelberg, 2010, pp. 13–25.

[MRS+10] Mannweiler, C.; Raulefs, R.; Schneider, J.; Denis, B.; Klein, A.; Uguen, B.; Laaraiedh, M.; Schotten, H. D.: A Robust Management Platform for Multi-Sensor Location Data Interpretation. *Proceedings of the Future Network & Mobile Summit 2010*, 2010.

[MS14a] Mannweiler, C.; Schotten, H.: Evaluating the Energy Balance of Solar-Powered Coordinated Wireless Backhaul Networks. *Proceedings of the IEEE 80th Vehicular Technology Conference (VTC 2014-Fall)*, 2014.

[MS14b] Mannweiler, C.; Schotten, H.: The Suitability of Energy-Autonomous Point-to-Point Radio Networks for Wireless Backhauling. *Proceedings of the 19th "ITG Fachtagung Mobilkommunikation"*, 2014.

[MSK+13] Mannweiler, C.; Schneider, J.; Klein, A.; Chakraborty, P.; Schotten, H. D.: A Distributed Broker System Enabling Coordinated Access Schemes in Autonomous Wireless Networks. *Proceedings of the 9th IEEE International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2013, pp. 59–64.

[MSKS11] Mannweiler, C.; Schneider, J.; Klein, A.; Schotten, H. D.: From Context to Context Awareness: Model-based User Classification for Efficient Multicasting. *Proceedings of the 15th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2011)*, 2011.

[MSMS10] Mannweiler, C.; Simoes, J.; Moltchanov, B.; Schotten, H. D.: Context-aware Smart Environments Enabling New Business Models and Services. *Proceedings of the ITU Kaleidoscope - Beyond the Internet? - Innovations for future networks and services*, 2010.

[SJMS13] Schneider, J.; Ji, L.; Mannweiler, C.; Schotten, H. D.: Context-based Cognitive Radio for LTE-Advanced Networks. *Proceedings of the 18th "ITG Fachtagung für Mobilkommunikation"*, 2013.

[SKMS11] Schneider, J.; Klein, A.; Mannweiler, C.; Schotten, H. D.: An Efficient Architecture for the Integration of Sensor and Actuator Networks into the Future Internet. *Advances in Radio Science*, vol. 9, 2011.

[SKMS12] Schneider, J.; Klein, A.; Mannweiler, C.; Schotten, H. D.: A Context Management System for a Cost-efficient Smart Home Platform. *Advances in Radio Science*, vol. 10, 2012, pp. 135–139.

[SMKa12] Schneider, J.; Mannweiler, C.; Klein, A.; and Schotten, H. D.: A Coordination Protocol for Distributed Context Management Systems. *Proceedings of the 12th Würzburg Workshop on IP: Joint ITG, ITC, and Euro-NF Workshop Visions of Future Generation Networks (EuroView2012)*, 2012.

[SMKS09] Schneider, J.; Mannweiler, C.; Klein, A.; Schotten, H. D.: Erfassung von Umgebungskontext und Kontextmanagement. *Proceedings of the 14th "ITG Fachtagung Mobilkommunikation"*, 2009.

[SSWM+10] Stanoevska-Slabeva, K.; Wozniak, T.; Mannweiler, C.; Hoffend, I.; Schotten, H. D.: The Emerging Ecosystem for Context Information and the Role of Telecom Operators. *Proceedings of the 14th International Conference on Intelligence in Next Generation Networks (ICIN)*, 2010.

[SWM+11] Schneider, J.; Weinreich, J.; Mannweiler, C.; Klein, A.; Schotten, H. D.: Kostengünstige Implementierung einer Bluetooth-Indoor-Lokalisierungs-Plattform. *Proceedings of the 16th "ITG Fachtagung Mobilkommunikation"*, 2011.

[SWP+10] Staehle, B.; Wamser, F.; Pries, R.; Staehle, D.; Mannweiler, C.; Klein, A.; Schneider, J.; Schotten, H. D.: Application- and Context-aware Radio Resource Management for Future Wireless Networks. *Proceedings of the 10th Würzburg Workshop on IP: Joint ITG and Euro-NF Workshop "Visions of Future Generation Networks" (EuroView)*, 2010.

Bibliography

[3GP10] 3GPP: *Technical Specification Group Radio Access Network; Universal Terrestrial Radio Access (UTRA); Uplink transmit diversity for High Speed Packet Access (HSPA) (Release 10)*. 2010.

[AAH+97] Abowd, G. D.; Atkeson, C. G.; Hong, J.; Long, S.; Kooper, R.; Pinkerton, M.: Cyberguide: A Mobile Context-aware Tour Guide. *ACM Wireless Networks*, vol. 3, no. 5, 1997, pp. 421–433.

[AB62] Anderson, T. W.; Bahadur, R. R.: Classification into Two Multivariate Normal Distributions with Different Covariance Matrices. *Annals of Mathematical Statistics*, vol. 33, no. 2, 1962, pp. 420–431.

[AB01] Anstreicher, K. M.; Brixius, N. W.: A New Bound for the Quadratic Assignment Problem Based on Convex Quadratic Programming. *Mathematical Programming*, vol. 89, no. 3, 2001, pp. 341–357.

[ABG+10] Auer, G.; Blume, O.; Giannini, V.; Godor, I.; Imran, M. A.; Jading, Y.; Katranaras, E.; Olsson, M.; Sabella, D.; Skillermark, P.; Wajda, W.: *Energy Efficiency Analysis of Reference Systems, Areas of Improvements and Target Breakdown: Deliverable 2.3*. 2010.

[ABPW08] Andresen, M.; Bräsel, H.; Plauschin, M.; Werner, F.: Using Simulated Annealing for Open Shop Scheduling with Sum Criteria. Tan, C. M. (eds.): *Simulated Annealing*. pp. 49–76. InTech, 2008.

[ADT00] Andrieu, C.; Doucet, A.; Touzni, A.: Adaptive MAP Multi-User Detection for Fading CDMA Channels. *Proceedings of the 10th IEEE Workshop on Statistical Signal and Array Processing*, 2000, pp. 6–9.

[AGG+11] Auer, G.; Giannini, V.; Godor, I.; Skillermark, P.; Olsson, M.; Imran, M. A.; Sabella, D.; Gonzalez, M. J.; Desset, C.; Blume, O.: Cellular Energy Efficiency Evaluation Framework. *Proceedings of the IEEE 73rd Vehicular Technology Conference (VTC 2011-Spring),*, 2011, pp. 1–6.

[AHRS04] Arkin, E. M.; Hassin, R.; Rubinstein, S.; Sviridenko, M.: Approximations for Maximum Transportation with Permutable Supply Vector and Other Capacitated Star Packing Problems. *Algorithmica*, vol. 39, no. 2, 2004, pp. 175–187.

[AHT13] Ahmid, A.; Höglund, R.; Tsakiroglou, C.: *Expanding the Serengeti Broadband Network (SBN): IP Address Renumbering*. 2013.

[AM90] Aleksander, I.; Morton, H.: *An Introduction to Neural Computing*. Chapman and Hall, 1990.

[ANK10] Alimian, A.; Nordman, B.; Kharitonov, D.: *Network and Telecom Equipment - Energy and Performance Assessment: Energy Consumption Rating Initiative Std., Rev. Draft 3.0.1*. 2010.

[ANNM04] Aoul, Y. H.; Nafaa, A.; Negru, D.; Mehaoua, A.: FAFC: Fast Adaptive Fuzzy AQM Controller for TCP/IP Networks. *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, vol. 3, 2004, pp. 1319 – 1323 Vol.3.

[APH07] Anagnostopoulos, C. B.; Pasias, P.; Hadjiefthymiades, S.: A Framework for Imprecise Context Reasoning. *Proceedings of the IEEE International Conference on Pervasive Services (ICPS)*, 2007, pp. 181–184.

[ARP08] Alizamir, S.; Rebennack, S.; Pardalos, P. M.: Improving the Neighborhood Selection Strategy in Simulated Annealing Using the Optimal Stopping Problem. Tan, C. M. (eds.): *Simulated Annealing*. pp. 363–382. InTech, 2008.

[Aru12] Aruba Networks: *Islamic University in Madinah Deploys Wireless Networks for Students and Staff*. 2012.

[ATI09] ATIS Exploratory Group on Green: *ATIS Report on Environmental Sustainability*. 2009.

[Ava10] Avallone, S.: An Energy-efficient Channel Assignment and Routing Algorithm for Multi-Radio Wireless Mesh Networks. *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, 2010, pp. 1–5.

[BBC+08] Banchs, A.; Bayer, N.; Chieng, D.; La Oliva, A. de; Gloss, B.; Kretschmer, M.; Murphy, S.; Natkaniec, M.; Zdarsky, F.: CARMEN: Delivering Carrier Grade Services over Wireless Mesh Networks. *Proceedings of the 19th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2008, pp. 1–6.

[BBX97] Brown, P. J.; Bovey, J. D.; Xian, C.: Context-aware Applications: From the Laboratory to the Marketplace. *Personal Communications, IEEE*, vol. 4, no. 5, 1997, pp. 58–64.

[BD66] Bellman, R. E.; Dreyfus, S. E.: *Applied Dynamic Programming*. vol. 7962. Princeton University Press, 1966.

[BDR07] Baldauf, M.; Dustdar, S.; Rosenberg, F.: A Survey on Context-aware Systems. *International Journal of Ad-Hoc and Ubiquitous Computing*, vol. 2, no. 4, 2007, pp. 263–277.

[Bel54] Bellman, R. E.: The Theory of Dynamic Programming. *Bulletin of the American Mathematical Society*, vol. 60, no. 6, 1954, pp. 503–516.

[Bez81] Bezdek, J. C.: *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York and USA: Plenum Press, 1981.

[BG04] Brickley, D.; Guha, R. V.: *RDF Vocabulary Description Language 1.0: RDF Schema*. 10.02.2004.

[BGF04] Boukezzoula, R.; Galichet, S.; Foulloy, L.: Observer-Based Fuzzy Adaptive Control for a Class of Nonlinear Systems: Real-Time Implementation for a Robot Wrist. *IEEE Transactions on Control Systems Technology*, vol. 12, no. 3, 2004, pp. 340–351.

[BGN00] Byrd, R. H.; Gilbert, J. C.; Nocedal, J.: A Trust Region Method Based on Interior Point Techniques for Nonlinear Programming. *Mathematical Programming*, vol. 89, no. 1, 2000, pp. 149–185.

[BGY12] Bin Sediq, A.; Gohary, R. H.; Yanikomeroglu, H.: Optimal Tradeoff Between Efficiency and Jain's Fairness Index in Resource Allocation. *IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2012, pp. 577–583.

[BHLM07] Baumann, R.; Heimlicher, S.; Lenders, V.; May, M.: HEAT: Scalable Routing in Wireless Mesh Networks Using Temperature Fields. *Proceedings of the 8th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2007, pp. 1–9.

[BJ87] Bucy, R. S.; Joseph, P. D.: *Filtering for Stochastic Processes With Applications to Guidance*. Chelsea Publishing Series. Chelsea Publishing Company, 1987.

[BK65] Bellman, R. E.; Kalaba, R. E.: *Dynamic Programming and Modern Control Theory*. Academic Press New York, 1965.

[BLRS03] Blough, D. M.; Leoncini, M.; Resta, G.; Santi, P.: The K-Neigh Protocol for Symmetric Topology Control in Ad-hoc Networks. *Proceedings of the 4th ACM International Symposium on Mobile Ad-hoc Networking & Computing*, MobiHoc '03, New York, NY and USA, 2003, pp. 141–152.

[BMJ+98] Broch, J.; Maltz, D. A.; Johnson, D. B.; Hu, Y.-C; Jetcheva, J.: A Performance Comparison of Multi-Hop Wireless Ad-hoc Network Routing Protocols. *Proceedings of the 4th ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, MobiCom '98, New York, NY and USA, 1998, pp. 85–97.

[BMVSSO04] Baturone, I.; Moreno-Velo, F. J.; Sanchez-Solano, S.; Ollero, A.: Automatic Design of Fuzzy Controllers for Car-Like Autonomous Robots. *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 4, 2004, pp. 447–465.

[Bod01] Bodenhofer, U.: *Genetic Algorithms: Theory and Applications*. 2001.

[Bou13] Boucher, C.: *The Fundamental Theorem of Linear Programming: Wolfram Demonstrations Project*. 2013.

[BPSM+08] Bray, T.; Paoli, J.; Sperberg-McQueen, C. M.; Maler, E.; Yergeau, F.: *Extensible Markup Language (XML) 1.0 (Fifth Edition)*. 2008.

[Bra13] Brahmi, N.: METIS: Mobile Communications for 2020 and Beyond. *Proceedings of the 18th "ITG Fachtagung für Mobilkommunikation"*, 2013, pp. 8–10.

[Bri06] Bringsjord, S.: *Artificial Intelligence - Neurocomputational Approaches: Stanford Encyclopedia of Philosophy*. 2006.

[Bro83] Brown, R.G.: *Introduction to Random Signal Analysis and Kalman Filtering*. Wiley, 1983.

[Bro96] Brown, M.: Supporting User Mobility. *Proceedings of IFIP World Conference on Mobile Communications*, 1996, pp. 69–77.

[BRR+13] Bjorkman, B.; Rembert, B.; Rossenhoevel, C.; Purdy, C.; Bercovich, D.; Semaan, G.; Fishburn, M.; Tilley, P.; Ranganathan, R.: *Multiple Classes of Service in Carrier Ethernet Mobile Backhaul Networks: MEF Best Practices Document*. 2013.

[BSS06] Bazaraa, M. S.; Sherali, H. D.; Shetty, C. M.: *Nonlinear Programming: Theory and Algorithms*. 3 edition. Hoboken and N.J: Wiley-Interscience, 2006.

[BT95] Boggs, P. T.; Tolle, J. W.: Sequential Quadratic Programming. *Acta numerica*, vol. 4, no. 1, 1995, pp. 1–51.

[BZMK09] Baker, N.; Zafar, M.; Moltchanov, B.; Knappmeyer, M.: Context-aware Systems and Implications for Future Internet. *Future Internet Assembly*. pp. 335–344. IOS Press, 2009.

[CABM05] Couto, D. S. J. de; Aguayo, D.; Bicket, J.; Morris, R.: A High-Throughput Path Metric for Multi-Hop Wireless Routing. *Wireless Networks - Special issue: Selected papers from ACM MobiCom 2003*, vol. 11, no. 4, 2005, pp. 419–434.

[CBD02] Camp, T.; Boleng, J.; Davies, V.: A Survey of Mobility Models for Ad-hoc Network Research. *Wireless Communications & Mobile Computing (WCMC): Special Issue On Mobile Ad-hoc Networking: Research, Trends and Applications*, vol. 2, no. 5, 2002, pp. 483–502.

[CC08] C-Cast: *Requirements and Concepts for Context Casting Service Eenablers and Context Management: C-Cast ICT-2007-216462 WP 3 Deliverable D6*. 01.11.2008.

[CC10] C-Cast: *Provide an End-to-End Context-aware Communication Framework: Project of the European Union Framework Programme 7*. 2010.

[CCG98] Cali, F.; Conti, M.; Gregori, E.: IEEE 802.11 Wireless LAN: Capacity Analysis and Protocol Enhancement. *Proceedings of the 17th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, vol. 1, 1998, pp. 142 –149 vol.1.

[CCG00] Calì, F.; Conti, M.; Gregori, E.: Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit. *IEEE/ACM Trans. Netw*, vol. 8, no. 6, 2000, pp. 785–799.

[CCLS01] Cao, J.; Cleveland, W. S.; Lin, D.; Sun, D. X.: On the Nonstationarity of Internet Traffic. *Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, SIGMETRICS '01, New York, NY and USA, 2001, pp. 102–112.

[CDM+00] Cheverst, K.; Davies, N.; Mitchell, K.; Friday, A.; Efstratiou, C.: Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences. *Proceedings of the Conference on Human Factors in Computing Systems (SIGCHI)*, CHI '00, New York and USA, 2000, pp. 17–24.

[Cel98] Cela, E.: *The Quadratic Assignment Problem: Theory and Algorithms*. vol. 1 in series *Combinatorial optimization*. Dordrecht and Boston: Kluwer Academic Publishers, 1998.

[CFJ03a] Chen, H.; Finin, T.; Joshi, A.: An Intelligent Broker for Context-aware Systems. *Proceedings of the 5th International Conference on Ubiquitous Computing (UbiComp)*, 2003, pp. 197–198.

[CFJ03b] Chen, H.; Finin, T.; Joshi, A.: An Ontology for Context-aware Pervasive Computing Environments. *The Knowledge Engineering Review*, vol. 18, no. 3, 2003, pp. 197–207.

[Che04] Chen, H. L.: *An Intelligent Broker Architecture for Pervasive Context-aware Systems*. Dissertation, University of Maryland, Baltimore County, 2004.

[CHG+01] Cordon, O.; Herrera, F.; Gomide, F.; Hoffmann, F.; Magdalena, L.: Ten Years of Genetic Fuzzy Systems: Current Framework and New Trends. *Proceedings of the 20th North American Fuzzy Information Processing Society (NAFIPS) International Conference*, vol. 3, 2001, pp. 1241–1246.

[Chu08] Chundu, R.: Mobile Broadband Backhaul: Addressing the Challenge. *Ericsson Review*, vol. 85, no. 3, 2008, pp. 1–4.

[Cis12] Cisco Systems, Inc.: *Visual Networking Index: Forecast and Methodology, 2011–2016: White Paper*. 2012.

[Cis13] Cisco Systems, Inc.: *Visual Networking Index: Forecast and Methodology, 2012-2017: White Paper*. 2013.

[CJ03] Clausen, T.; Jacquet, P.: *Optimized Link State Routing Protocol (OLSR)*. RFC 3626 (Experimental), 2003.

[CJBM02] Chen, B.; Jamieson, K.; Balakrishnan, H.; Morris, R.: Span: An Energy-efficient Coordination Algorithm for Topology Maintenance in Ad-hoc Wireless Networks. *Wireless Networks*, vol. 8, no. 5, 2002, pp. 481–494.

[CKY10] Chen, T.; Kim, H.; Yang, Y.: Energy Efficiency Metrics for Green Wireless Communications. *Proceedings of the International Conference on Wireless Communications and Signal Processing (WCSP)*, 2010, pp. 1–6.

[CL99] Chen, S.; Luk, B. L.: Adaptive Simulated Annealing for Optimization in Signal Processing Applications. *EURASIP Signal Processing Journal*, vol. 79, no. 1, 1999, pp. 117–128.

[CM99] Claffy, K. C.; McCreary, S.: *Internet Measurement and Data Analysis: Passive and Active Measurement*. 1999.

[CTB+95] Cooperstock, J. R.; Tanikoshi, K.; Beirne, G.; Narine, T.; Buxton, W. A. S.: Evolution of a Reactive Environment. *Procceedings of the Conference on Human Factors in Computing Systems (CHI)*, CHI '95, New York, NY and USA, 1995, pp. 170–177.

[CV95] Cortes, C.; Vapnik, V.: Support-Vector Networks. *Mach. Learn*, vol. 20, no. 3, 1995, pp. 273–297.

[CV12] Ceriani, L.; Verme, P.: The Origins of the Gini Index: Extracts from Variabilità e Mutabilità (1912) by Corrado Gini. *The Journal of Economic Inequality*, vol. 10, no. 3, 2012, pp. 421–443.

[CZB+10] Correia, L. M.; Zeller, D.; Blume, O.; Ferling, D.; Jading, Y.; Go-Anddor, I.; Auer, G.; van der Perre, L.: Challenges and Enabling Technologies for Energy Aware Mobile Radio Networks. *IEEE Communications Magazine*, vol. 48, no. 11, 2010, pp. 66–72.

[DA99] Dey, A. K.; Abowd, G. D.: *The Context Toolkit: Aiding the Development of Context-aware Applications*. 1999.

[DA00] Dey, A. K.; Abowd, G. D.: Towards a Better Understanding of Context and Context Awareness. *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, 2000, pp. 304–307.

[DAPW98] Dey, A. K.; Abowd, G.; Pinkerton, M.; Wood, A.: CyberDesk: A Framework for Providing Self-Integrating Context-aware Services. *Knowledge-Based Systems*, vol. 11, no. 1, 1998, pp. 47–54.

[Deu12a] Deutscher Wetterdienst (DWD): *Globalstrahlung: Die Energie der Sonne*. 2012.

[Deu12b] Deutscher Wetterdienst (DWD): *Globalstrahlung in der Bundesrepublik Deutschland*. 2012.

[Die01] Diebold, F. X.: *Elements of Forecasting*. 2 edition. Mason, Ohio: Thomson/South-Western, 2001.

[Die06] Diestel, R.: *Graphentheorie*. 3rd edition. Berlin [u.a.]: Springer, 2006.

[Dim13] Dimroth, F.: *World Record Solar Cell with 44.7% Efficiency*. 2013.

[DKJ+03] Djuric, P.M.; Kotecha, J.H.; Jianqui Z.; Yufei H.; Ghirmai, T.; Bugallo, M.F.; Miguez, J.: Particle Filtering. *IEEE Signal Processing Magazine*, vol. 20, no. 5, 2003, pp. 19–38.

[DMCB98] Davies, N.; Mitchell, K.; Cheverst, K.; Blair, G.: *Developing A Context Sensitive Tourist Guide*. 1998.

[DN09] Dey, A. K.; Newberger, A.: Support for Context-aware Intelligibility and Control. *Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI)*, 2009, pp. 859–868.

[DPZ04] Draves, Richard; Padhye, Jitendra; Zill, Brian: Routing in Multi-Radio, Multi-hop Wireless Mesh Networks. *Proceedings of the 10th International Conference on Mobile Computing and Networking*, MobiCom '04, New York, NY and USA, 2004, pp. 114–128.

[Dre10] Dreyfus, S.: Modern Computational Applications of Dynamic Programming. *Journal of Industrial and Systems Engineering*, vol. 4, no. 3, 2010, pp. 152–155.

[DT97] Dantzig, G. B.; Thapa, M. N.: *Linear Programming 1: Introduction*. Secaucus and USA: Springer-Verlag New York, Inc, 1997.

[EG04] ETSI; 3GPP: *Universal Mobile Telecommunications System (UMTS) - Spacial Channel Model for Multiple Input Multiple Output (MIMO) Simulations*. 27.01.2004.

[Egl90] Eglese, R.W.: Simulated Annealing: A Tool for Operational Research. *European Journal of Operational Research*, vol. 46, no. 3, 1990, pp. 271–281.

[EHC+93] Elrod, S.; Hall, G.; Costanza, R.; Dixon, M.; Des RiviÃ¨res, J.: Responsive Office Environments. *Commun. ACM*, vol. 36, no. 7, 1993, pp. 84–85.

[EIKP08] Eiter, T.; Ianni, G.; Krennwallner, T.; Polleres, A.: Rules and Ontologies for the Semantic Web. Baroglio, C.; Bonatti, P.; Maluszynski, J.; Marchiori, M.; Polleres, A.; Schaffert, S. (eds.): *Reasoning Web*. vol. 5224 in series *Lecture Notes in Computer Science*. pp. 1–53. Springer Berlin / Heidelberg, 2008.

[Ein05] Einstein, A.: *Investigations on the Theory of the Brownian Movement*. 1905.

[EK72] Edmonds, J.; Karp, R. M.: Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems. *Journal of the ACM (JACM)*, vol. 19, no. 2, 1972, pp. 248–264.

[Eri11] Ericsson AB: *More than 50 Billion Connected Devices: Taking Connected Devices to Mass Market and Profitability: White Paper*. 2011.

[Eri12] Ericsson AB: *It All Comes Back to Backhaul: White Paper*. Feb. 2012.

[ES96] Elbaum, R.; Sidi, M.: Topological Design of Local Area Networks Using Genetic Algorithms. *IEEE/ACM Transactions on Networking (TON)*, vol. 4, no. 5, 1996, pp. 766–778.

[EV07] Ergen, M.; Varaiya, P.: Decomposition of Energy Consumption in IEEE 802.11. *Proceedings of the IEEE International Conference on Communications (ICC)*, 2007, pp. 403–408.

[Eva11] Evans, D.: *The Internet of Things: How the Next Evolution of the Internet Is Changing Everything: White Paper*. 2011.

[FC11] Ferreira, L. S.; Correia, L. M.: Energy-efficient Radio Resource Management in Self-Organised Multi-Radio Wireless Mesh Networks. *Proceedings of the 22nd IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2011, pp. 232–236.

[Fen06] Feng, G.: A Survey on Analysis and Design of Model-Based Fuzzy Control Systems. *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 5, 2006, pp. 676–697.

[FF55] Ford, L. R.; Fulkerson, D. R.: *A Simple Algorithm for Finding Maximal Network Flows and an Application to the Hitchcock Problem*. Rand Corporation, 1955.

[FGG97] Friedman, N.; Geiger, D.; Goldszmidt, M.: Bayesian Network Classifiers. *Machine Learning*, vol. 29, no. 2-3, 1997, pp. 131–163.

[FKS97] Fickas, S.; Kortuem, G.; Segall, Z.: Software Organization for Dynamic and Adaptable Wearable Systems. *Digest of Papers, 1st International Symposium on Wearable Computers*, 1997, pp. 56–.

[For92] Fortune, S.: Voronoi Diagrams and Delaunay Triangulations. *Computing in Euclidean geometry*, vol. 1, 1992, pp. 193–233.

[FPN+05] Floréen, P.; Przybilski, M.; Nurmi, P.; Koolwaaij, J.; Tarlano, A.; Wagner, M.; Luther, M.; Bataille, F.; Boussard, M.; Mrohs, B.; Lau, S.: Towards a Context Management Framework for MobiLife. *Proceedings of the IST Mobile & Wireless Communications Summit*, 2005.

[Fri95] Fritzke, B.: A Growing Neural Gas Network Learns Topologies. Tesauro, G.; Touretzky, D. S.; Leen, T. K. (eds.): *Advances in Neural Information Processing Systems: 7*. pp. 625–632. Cambridge and USA: MIT Press, 1995.

[Fri97] Fritzke, B.: *Some Competitive Learning Methods: Java Paper*. 1997.

[Fri98] Fritzke, B.: *Vektorbasierte neuronale Netze*. Dissertation, Friedrich-Alexander-Universitaet Erlangen-Nuernberg, Erlangen, 1998.

[GCL08] Gen, M.; Cheng, R.; Lin, L.: *Network Models and Optimization: Multiobjective Genetic Algorithm Approach*. London: Springer London, 2008.

[Gel74] Gelb, A.: *Applied Optimal Estimation*. MIT Press, 1974.

[Geo14] GeoModel Solar: *SolarGIS*. 2010-2014.

[GGB+02] Gustafsson, F.; Gunnarsson, F.; Bergman, N.; Forssell, U.; Jansson, J.; Karlsson, R.; Nordlund, P.-J.: Particle Filters for Positioning, Navigation, and Tracking. *IEEE Transactions on Signal Processing*, vol. 50, no. 2, 2002, pp. 425–437.

[GJ03] Garey, M. R.; Johnson, D. S.: *Computers and Intractability: A Guide to the Theory of NP-Completeness*. 24th print edition, A series of books in the mathematical sciences. New York: W.H. Freeman and Co., 2003.

[GK00] Gupta, P.; Kumar, P. R.: The Capacity of Wireless Networks. *IEEE Transactions on Information Theory*, vol. 46, no. 2, 2000, pp. 388–404.

[GLAS99] Garcia-Luna-Aceves, J. J.; Spohn, M.: Source-Tree Routing in Wireless Networks. *Proceedings of the Seventh International Conference on Network Protocols*, 1999, pp. 273–282.

[GPZ04] Gu, T.; Pung, H. K.; Zhang, D. Q.: A Middleware for Building Context-aware Mobile Services. *Proceedings of the 59th IEEE Vehicular Technology Conference (VTC-Spring)*, vol. 5, 2004, pp. 2656 – 2660 Vol.5.

[GSS93] Gordon, N. J.; Salmond, D. J.; Smith, A. F. M.: Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation. *IEE Proceedings F: Radar and Signal Processing=*, vol. 140, no. 2, 1993, pp. 107–113.

[Hay05] Hay, R.: *WLAN Throughput and Coverage*. 08.12.2005.

[HB95] Haeberlin, H.; Beutler, C.: Normalized Representation of Energy and Power for Analysis of Performance and On-line Error Detection in PV-Systems. *Proceedings of the European Photovoltaic Solar Energy Conference*, 1995, pp. 1–4.

[HB14] Honsberg, C.; Bowden, S.: *A Collection of Resources for the Photovoltaic Educator*. 2014.

[HGPC99] Hong, X.; Gerla, M.; Pei, G.; Chiang, C.-C.: A Group Mobility Model for Ad-hoc Wireless Networks. *Proceedings of the 2nd ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, MSWiM '99, New York, NY and USA, 1999, pp. 53–60.

[HJJ03] Henderson, D.; Jacobson, S. H.; Johnson, A. W.: The Theory and Practice of Simulated Annealing. Glover, F.; Kochenberger, G. A. (eds.): *Handbook of Metaheuristics*. pp. 287–319. International Series in Operations Research & Management Science. Kluwer Academic Publishers, 2003.

[HM54] Hammersley, J. M.; Morton, K. W.: Poor Man's Monte Carlo. *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 16, no. 1, 1954, pp. p 23–38.

[HNBR97] Hull, R.; Neaves, P.; Bedford-Roberts, J.: Towards Situated Computing. *Digest of Papers, 1st International Symposium on Wearable Computers*, 1997, pp. 146–153.

[HPS02] Haas, Z. J.; Pearlman, M. R.; Samar, P.: *The Zone Routing Protocol (ZRP) for Ad-Hoc Networks*. 2002.

[HR06] Hassin, Refael; Rubinstein, Shlomi: An Improved Approximation Algorithm for the Metric Maximum Clustering Problem with Given Cluster Sizes. *Information processing letters*, vol. 98, no. 3, 2006, pp. 92–95.

[HS13] Hollands, K. G. T.; Suehrcke, H.: A Three-State Model for the Probability Distribution of Instantaneous Solar Radiation, with Applications. *Solar Energy*, vol. 96, 2013, pp. 103–112.

[HSP+03] Hofer, T.; Schwinger, W.; Pichler, M.; Leonhartsberger, G.; Altmann, J.; Retschitzegger, W.: Context Awareness on Mobile Devices - the Hydrogen Approach. *Proceedings of the 36th International Conference on System Sciences (HICSS)*, HICSS '03, Washington (DC) and USA, 2003, pp. 292.1–.

[HT05] Huang, C.-F.; Tseng, Y.-C.: The Coverage Problem in a Wireless Sensor Network. *Mobile Networks and Applications*, vol. 10, no. 4, 2005, pp. 519–528.

[HT10] Holma, H.; Toskala, A.: *WCDMA for UMTS: HSPA Evolution and LTE*. 5th ed edition. Chichester, West Sussex: Wiley, 2010.

[Hu93] Hu, L.: Topology Control for Multihop Packet Radio Networks. *IEEE Transactions on Communications*, vol. 41, no. 10, 1993, pp. 1474–1481.

[IEE09] IEEE: *IEEE Standard for Information Technology - Telecommunications and Information Exchange between systems - Local and metropolitan area networks - Specific requirements: Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) - Specifications Amendment 5: Enhancements for Higher Throughput*. 2009.

[Int97] International Telecommnication Union: *Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000*. 1997.

[Irm08] Irmer, R.: *NGMN Radio Access Performance Evaluation Methodology*. 2008.

[Jac93] Jacobs, O. L. R.: *Introduction to Control Theory*. Oxford science publications. Oxford University Press, 1993.

[JCH84] Jain, R.; Chiu, D.-M.; Hawe, W. R.: *A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer Systems*. Eastern Research Laboratory, Digital Equipment Corporation, 1984.

[JGJS99] Jordan, M. I.; Ghahramani, Z.; Jaakkola, T. S.; Saul, L. K.: An Introduction to Variational Methods for Graphical Models. *Mach. Learn.*, vol. 37, no. 2, 1999, pp. 183–233.

[JLH+99] Johansson, P.; Larsson, T.; Hedman, N.; Mielczarek, B.; Degermark, M.: Scenario-Based Performance Analysis of Routing Protocols for Mobile Ad-hoc Networks. *Proceedings of the 5th ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, MobiCom '99, New York, NY and USA, 1999, pp. 195–206.

[JLL+07] Jo, D.; Lee, J.; Lee, S.; Ha, T.; Kwon, T.; Choi, Y.: Signal Dragging: Effects of Terminal Movement on War-Driving in CDMA/WCDMA Networks. *Proceedings of the International Conference on Location and Context Awareness (LoCA)*, LoCA'07, Berlin and Heidelberg, 2007, pp. 211–227.

[JM96] Johnson, D. B.; Maltz, D. A.: Dynamic Source Routing in Ad-hoc Wireless Networks. Imielinski, T.; Korth, H. (eds.): *Mobile Computing*. pp. 153–181. Kluwer Academic Publishers, 1996.

[JNL99] Joa-Ng, M.; Lu, I.-T.: A Peer-to-Peer Zone-Based Two-Level Link State Routing for Mobile Ad-Hoc Networks. *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, 1999, pp. 1415–1425.

[Joi12] Joint Research Centre, Institute for Energy, Renewable Energy Unit: *Photovoltaic Geographical Information System (PVGIS): Geographical Assessment of Solar Resource and Performance of Photovoltaic Technology*. 10.02.2012.

[JS03] Jun, J.; Sichitiu, M. L.: The Nominal Capacity of Wireless Mesh Networks. *IEEE Transactions on Wireless Communications*, vol. 10, no. 5, 2003, pp. 8–14.

[JSK+13] Jiang, C.; Shi, Y.; Kompella, S.; Hou, Y. T.; Midkiff, S. F.: Bicriteria Optimization in Multi-Hop Wireless Networks: Characterizing the Throughput-Energy Envelope. *IEEE Transactions on Mobile Computing*, vol. 12, no. 9, 2013, pp. 1866–1878.

[JU97] Julier, S. J.; Uhlmann, J. K.: A New Extension of the Kalman Filter to Nonlinear Aystems. *11th International Symposium on Aerospace/Defense Sensing (AeroSense), Simulations and Controls*, 1997, pp. 182–193.

[Kal60] Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, 1960, pp. 35–45.

[KB61] Kalman, R. E.; Bucy, R. S.: New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering*, vol. 83, no. 3, 1961, pp. 95–108.

[KD03a] Kotecha, J. H.; Djuric, P. M.: Gaussian Particle Filtering. *IEEE Transactions on Signal Processing*, vol. 51, no. 10, 2003, pp. 2592–2601.

[KD03b] Kotecha, J. H.; Djuric, P. M.: Gaussian Sum Particle Filtering. *IEEE Transactions on Signal Processing*, vol. 51, no. 10, 2003, pp. 2602–2612.

[KDE07] Kröni, V.; Durisch, W.; Ebers, S.: *Leistungsfähigkeit einiger netzgekoppelter Inverter und Solarmodule*. 11.12.2007.

[KGV83] Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P.: Optimization by Simulated Annealing. *Science*, vol. 220, no. 4598, 1983, pp. 671–680.

[KKKP00] Kirousis, L. M.; Kranakis, E.; Krizanc, D.; Pelc, A.: Power Consumption in Packet Radio Networks. *Theor. Comput. Sci.*, vol. 243, no. 1-2, 2000, pp. 289–305.

[KL03] Keerthi, S. S.; Lin, C.-J.: Asymptotic Behaviors of Support Vector Machines with Gaussian Kernel: Neural Computation. *Neural Computation*, vol. 15, no. 7, 2003, pp. 1667–1689.

[KM03] Korpipää, P.; Mäntyjärvi, J.: An Ontology for Mobile Device Sensor-Based Context Awareness. *Proceedings of the 4th International and Interdisciplinary Conference on Modeling and Using Context*, CONTEXT'03, Berlin and Heidelberg, 2003, pp. 451–458.

[KMH+07] Kyösti, P.; Meinilä, J.; Hentilä, L.; Zhao, X.; Jämsä, T.; Schneider, C.; Narandzić, M.; Milojević, M.; Hong, A.; Ylitalo, J.; Holappa, V.-M.; Alatossava, M.; Bultitude, R.; Jong, Y. de; Rautiainen, T.: *WINNER II Channel Models*. 2007.

[KMK+03] Korpipää, P.; Mantyjarvi, J.; Kela, J.; Keranen, H.; Malm, E. J.: Managing Context Information in Mobile Devices. *IEEE Pervasive Computing*, vol. 2, no. 3, 2003, pp. 42–51.

[KN92] Kaplan, R. S.; Norton, D. P.: The Balanced Scorecard: Measures that Drive Performance. *Harvard Business Review*, vol. 70, no. 1, 1992, pp. 71–79.

[KP99] Kozan, E.; Preston, P.: Genetic Algorithms to Schedule Container Transfers at Multimodal Terminals. *International Transactions in Operational Research*, vol. 6, no. 3, 1999, pp. 311–329.

[KSB98] Kortuem, G.; Segall, Z.; Bauer, M.: Context-aware, Adaptive Wearable Computers as Remote Interfaces to Intelligent Environments. *Digest of Papers, 2nd International Symposium on Wearable Computers*, 1998, pp. 58–65.

[Laq03] Laquai, B.: *Abschätzung des möglichen Energieertrags einer Photovoltaikanlage*. 23.02.2003.

[LBC+01] Li, J.; Blake, C.; Couto, D. S. J. de; Lee, H. I.; Morris, R.: Capacity of Ad-hoc Wireless Networks. *Proceedings of the 7th International Conference on Mobile Computing and Networking (MobiCom)*, Rome and Italy, 2001, pp. 61–69.

[LC00] Lo, C.-C.; Chang, W.-H.: A Multi-bjective Hybrid Genetic Algorithm for the Capacitated Multi-point Network Design Problem. *Trans. Sys. Man Cyber. Part B*, vol. 30, no. 3, 2000, pp. 461–470.

[LCW97] Lam, D.; Cox, D.C; Widom, J.: Teletraffic Modeling for Personal Communications Services. *IEEE Communications Magazine*, vol. 35, no. 2, 1997, pp. 79–87.

[Lee08] Lee, E. A.: Cyber Physical Systems: Design Challenges. *Proceedings of the 11th IEEE International Symposium on Object-Oriented Real-Time Distributed Computing (ISORC)*, 2008, pp. 363–369.

[LH03] Liang, B.; Haas, Z. J.: Predictive Distance-Based Mobility Management for Multidimensional PCS Networks. *IEEE/ACM Trans. Netw*, vol. 11, no. 5, 2003, pp. 718–732.

[LHB+05] Li, L.; Halpern, J. Y.; Bahl, P.; Wang, Y.-M.; Wattenhofer, R.: A Cone-based Distributed Topology Control Algorithm for Wireless Multi-hop Networks. *IEEE/ACM Transactions on Networking (TON)*, vol. 13, no. 1, 2005, pp. 147–159.

[LKWG11] Lange, C.; Kosiankowski, D.; Weidmann, R.; Gladisch, A.: Energy Consumption of Telecommunication Networks and Related Improvement Options. *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, no. 2, 2011, pp. 285–295.

[LLT07] Lee, L. H.; Lee, C. U.; Tan, Y. P.: *A Multi-objective Genetic Algorithm for Robust Flight Scheduling Using Simulation*. 2007.

[LYGL05] Lei, J.; Yates, R.; Greenstein, L.; Liu, H.: Wireless Link SNR Mapping onto an Indoor Testbed. *Proceedings of the First International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (Tridentcom)*, 2005, pp. 130–135.

[May82] Maybeck, P. S.: *Stochastic Models, Estimation and Control*. Mathematics in Science and Engineering. Academic Press, 1982.

[MB94] McCarthy, J.; Buvac, S.: *Formalizing Context (Expanded Notes)*. 1994.

[MD97] Mahmood M. Z.; Dassanayake, P.: User Mobility Modeling and Characterization of Mobility Patterns. *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 7, 1997, pp. 1239–1252.

[Met11] Metro Ethernet Forum (MEF): *What is Carrier Ethernet?* 2011.

[Mic94] Michalewicz, Z.: *Genetic Algorithms + Data Structures = Evolution Programs*. 2 edition. Berlin, New York: Springer-Verlag, 1994.

[MKPS05] Megerian, S.; Koushanfar, F.; Potkonjak, M.; Srivastava, M. B.: Worst and Best-Case Coverage in Sensor Networks. *IEEE Transactions on Mobile Computing*, vol. 4, no. 1, 2005, pp. 84–92.

[MLK01] Murata, M.; Laurent, S. St.; Kohn, D.: *XML Media Types.* RFC 3023 (Proposed Standard), 2001.

[MM06] Moguerza, J. M.; Muñoz, A.: Support Vector Machines with Applications. *Statistical Science*, vol. 21, no. 3, 2006, pp. 322–336.

[MN06] Martin, M.; Nurmi, P.: A Generic Large-Scale Simulator for Ubiquitous Computing. *Proceedings of the 3rd International Conference on Mobile and Ubiquitous Systems: Networking & Services (MobiQuitous)*, San Jose, California and USA, 2006, pp. 1–3.

[MNNA09] Mudali, P.; Nyandeni, T. C.; Ntlatlapa, N.; Adiguns, M. O.: Design and Implementation of a Topology Control Scheme for Wireless Mesh Networks. *Proceedings of the 9th IEEE AFRICON*, 2009, pp. 1–6.

[Mob06] MobiLife: *Bring Advances in Mobile Applications and Services within the Reach of Users: Project of the European Union Framework Programme FP 6.* 2006.

[Mol11] Molisch, A. F.: *Wireless Communications.* 2 edition. Weinheim: Wiley, 2011.

[MP43] McCulloch, W. S.; Pitts, W.: A Logical Calculus of the Ideas Immanent in Nervous Activity. *The bulletin of mathematical biophysics*, vol. 5, no. 4, 1943, pp. 115–133.

[MS91] Martinez, T. M.; Schulten, K. J.: A neural gas network learns topologies. Kohonen, T. (eds.): *Artificial Neural Networks.* pp. 397–402. Amsterdam and Netherlands: Elsevier, 1991.

[Mul10] Mulligan, R.: Coverage in Wireless Sensor Networks: A Survey. *Network Protocols and Algorithms*, vol. 2, no. 2, 2010, pp. 27–53.

[Mur98] Murphy, K.: *A Brief Introduction to Graphical Models and Bayesian Networks.* 1998.

[Myl05] Myllymäki, P.: *Advantages of Bayesian Networks in Data Mining and Knowledge Discovery.* 2005.

[NS09] Nagarajan, V.; Sviridenko, M.: On the Maximum Quadratic Assignment Problem. *Mathematics of Operations Research*, vol. 34, no. 4, 2009, pp. 859–868.

[NT93] Nering, E. D.; Tucker, A. W.: *Linear Programs and Related Problems.* Computer Science and Scientific Computing Series. ACADEMIC PressINC, 1993.

[Obj07] Object Management Group (OMG): *Unified Modeling Language (UML).* 2007.

[OE07] Ozgovde, A.; Ersoy, C.: WCOT: A Realistic Lifetime Metric for the Performance Evaluation of Wireless Sensor Networks. *Proceedings of the 18th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2007, pp. 1–5.

[Ope12] Open Networking Foundation: *Software-Defined Networking: The New Norm for Networks: White Paper*. 13/04/2012.

[PB94] Perkins, C. E.; Bhagwat, P.: Highly Dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers. *SIGCOMM Comput. Commun. Rev.*, vol. 24, no. 4, 1994, pp. 234–244.

[PB02] Polyrakis, A.; Boutaba, R.: The Meta-Policy Information Base. *IEEE Network*, vol. 16, no. 2, 2002, pp. 40–48.

[PB10] Patras, P.; Bernardos, C. J.: *Final Architecture Deliverable: CARMEN Deliverable D1.6*. 2010.

[PBRD03] Perkins, C.; Belding-Royer, E.; Das, S.: *Ad hoc On-Demand Distance Vector (AODV) Routing*. RFC 3561 (Experimental), 2003.

[Pea85] Pearl, J.: Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning. *Proceedings of the 7th Conference of the Cognitive Science Society (CogSci)*, 1985, pp. 329–334.

[Pea88] Pearl, J.: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA and USA: Morgan Kaufmann Publishers Inc, 1988.

[PMP+10] Pedersen, T.; Mensing, C.; Papakonstantinou, K.; Oktem, T.-M.; Slock, D.; Tenoux, T.; García, M.; Angel GarcÃa, M.; Youssef, J.; Denis, B.; Laaraiedh, M.; Uguen, B.: *Hybrid Localization Techniques: WHERE Project Deliverable D2.3*. 07.05.2010.

[PR85] Postel, J.; Reynolds, J.: *File Transfer Protocol*. RFC 959 (Standard), 1985. Updated by RFCs 2228, 2640, 2773, 3659.

[PY98] Passino, K. M.; Yurkovich, S.: *Fuzzy Control*. Addison-Wesley, 1998.

[Qua04] Qualcomm Atheros, Inc: *Power Consumption and Energy Efficiency Comparisons of WLAN Products: Technical Paper*. April 2004.

[Qua06] Quaschning, V.: *Regenerative Energiesysteme: Technologie - Berechnung - Simulation*. München: Hanser, 2006.

[RAH98] Rekimoto, J.; Ayatsuka, Y.; Hayashi, K.: Augment-able Reality: Situated Communication through Physical and Digital Spaces. *Digest of Papers, 2nd International Symposium on Wearable Computers*, 1998, pp. 68–75.

[Ram08] Ramos, C. L.: A Monitoring Lifetime Metric for Sensor Networks. *Proceedings of the IEEE Military Communications Conference (MILCOM)*, 2008, pp. 1–7.

[RE12] Rocha, R. C. A.; Endler, M.: *Context Management for Distributed and Dynamic Context-aware Computing*. SpringerBriefs in Computer Science. London: Springer London, 2012.

[Rit06] Ritzenhoff, P.: *Erstellung eines Modells zur Simulation der Solarstrahlung auf beliebig orientierte Flächen und deren Trennung in Diffus- und Direktanteil*. 25.02.2006.

[RLH06] Rekhter, Y.; Li, T.; Hares, S.: *A Border Gateway Protocol 4 (BGP-4)*. RFC 4271 (Draft Standard), 2006.

[RM99] Rodoplu, V.; Meng, T. H.: Minimum Energy Mobile Wireless Networks. *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, 1999, pp. 1333–1344.

[RN02] Russell, S. J.; Norvig, P.: *Artificial Intelligence: A Modern Approach*. 2 edition. Prentice Hall, 2002.

[Rod08] Rodriguez, A.: *RESTful Web Services: The Basics*. 2008.

[RPM98] Ryan, N. S.; Pascoe, J.; Morse, D. R.: Enhanced Reality Fieldwork: The Context-aware Archaeological Assistant. *Computer Applications in Archaeology*, 1998.

[RRH00] Ramanathan, R.; Rosales-Hain, R.: Topology Control of Multihop Wireless Networks Using Transmit Power Adjustment. *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, vol. 2, 2000, pp. 404–413.

[Sal13] Salway, D.: *The "Middle Mile" in Broadband Networks - Definition: Broadband Guide*. 2013.

[San01] Sanchez, M.: *Mobility Models*. 08.11.2001.

[San05] Santi, P.: Topology Control in Wireless Ad-hoc and Sensor Networks. *ACM Computing Surveys*, vol. 37, no. 2, 2005, pp. 164–194.

[Sca85] Scales, L. E.: *Introduction to Non-linear Optimization*. Springer-Verlag New York, Inc, 1985.

[Sch83] Schittkowski, K.: On the Convergence of a Sequential Quadratic Programming Method with an Augmented Lagrangian Line Search Function. *Optimization*, vol. 14, no. 2, 1983, pp. 197–216.

[SDA98] Salber, D.; Dey, A. K.; Abowd, G. D.: *Ubiquitous Computing: Defining an HCI Research - Agenda for an Emerging Interaction Paradigm*. 1998.

[SEF+98] Sumi, Y.; Etani, T.; Fels, S.; Simonet, N.; Kobayashi, K.; Mase, K.: C-MAP: Building a Context-aware Mobile Assistant for Exhibition Tours. Ishida, T. (eds.): *Community Computing and Support Systems*, London, UK, 1998, pp. 137–154.

[SGL11] Sofra, N.; Gkelias, A.; Leung, K. K.: Route Construction for Long Lifetime in VANETs. *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, 2011, pp. 3450–3461.

[SLP04] Strang, T.; Linnhoff-Popien, C.: A Context Modeling Survey. *Proceedings of the 6th International Conference on Ubiquitous Computing (UbiComp)*, 2004.

[SM01] Sánchez, M.; Manzoni, P.: ANEJOS: A Java Based Simulator for Ad-hoc Networks. *Future Generation Computer Systems 17 (2001) 573–583*, vol. 17, no. 5, 2001, pp. 573–583.

[Sob02] Sobrinho, J. L.: Algebra and Algorithms for QoS Path Computation and Hop-by-Hop Routing in the Internet. *IEEE/ACM Transactions on Networking (TON)*, vol. 10, no. 4, 2002, pp. 541–550.

[Sol14] SolarMesh: *Energieeffizientes, Energieffizientes, autonomes, großflächiges Sprach- und Datenfunknetz mit flacher IP-Architektur: Research Project funded by the German Federal Ministry for Education and Research.* 2011-2014.

[Sor60] Sorenson, H.W.: *Kalman Filtering: Theory and Application.* IEEE Press selected reprint series. IEEE Press, 1960.

[SPI08] SPICE: *Service Platform for Innovative Communication Environment: Project of the European Union Framework Programme FP 6.* 2008.

[SS04] Shen, N.; Smit, H.: *Calculating Interior Gateway Protocol (IGP) Routes Over Traffic Engineering Tunnels.* RFC 3906 (Informational), 2004.

[SSB99] Sivakumar, R.; Sinha, P.; Bharghavan, V.: CEDAR: A Core-Extraction Distributed Ad-hoc Routing Algorithm. *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, 1999, pp. 1454–1465.

[SSJ11] Simsek, B.; Sittig, E.; Jonas, K.: *Connecting the Unconnected: Internet Access for Five Billion People.* 2011.

[ST94] Schilit, B. N.; Theimer, M. M.: *Disseminating Active Map Information to Mobile Hosts.* 1994.

[Ste56] Steinhaus, H.: Sur la division des corps matériels en parties. *Bull. Acad. Polon. Sci*, vol. 1, no. 12, 1956, pp. 801–804.

[Sug85] Sugeno, M.: An Introductory Survey of Fuzzy Control. *Information Sciences*, vol. 36, no. 1–2, 1985, pp. 59–83.

[SV02] Salsano, S.; Veltri, L.: QoS Control by Means of COPS to Support SIP-Based Applications. *IEEE Network*, vol. 16, no. 2, 2002, pp. 27–33.

[SZJ+08] Srinivasan, R.; Zhuang, J.; Jalloul, L.; Novak, R.; Park, J.: *IEEE 802.16m Evaluation Methodology Document (EMD).* 2008.

[Tal09] Talbot, T.: *Verizon NEBS(tm) Compliance: TEEER Metric Quantification.* 2009.

[Tan08] Tan, C. M. (eds.): *Simulated Annealing.* InTech, 2008.

[TC01] Tseng, C.-S.; Chen, B.-S.: H-Infinity Fuzzy Estimation for a Class of Nonlinear Discrete-Time Dynamic Systems. *IEEE Transactions on Signal Processing*, vol. 49, no. 11, 2001, pp. 2605–2619.

[TFP+02] Trimintzios, P.; Flegkas, P.; Pavlou, G.; Georgiadis, L.; Griffin, D.: Policy-Based Network Dimensioning for IP Differentiated Services Networks. *Proceedings of the IEEE Workshop on IP Operations and Management*, 2002, pp. 171–176.

[The13] The Open Mesh Project: *B.A.T.M.A.N. advanced.* 2013.

[Tra13] Trango Systems: *Wireless Backhaul Application Diagram.* 2013.

[TV05] Tse, D.; Viswanath, P.: *Fundamentals of Wireless Communication.* Cambridge: Cambridge University Press, 2005.

[vA01] van Laerhoven, K.; Aidoo, K.: Teaching Context to Applications. *Personal Ubiquitous Computing*, vol. 5, no. 1, 2001, pp. 46–49.

[vGP95] van Batenburg, F. H. D.; Gultyaev, A. P.; Pleij, C. W. A.: An APL-Programmed Genetic Algorithm for the Prediction of RNA Secondary Structure. *Journal of Theoretical Biology*, vol. 174, no. 3, 1995, pp. 269–280.

[VH98] Vanoli, K.; Heeke, S.: *Unterichtsmaterial zu Wärme von der Sonne.* 1998.

[vHH+09] van Bechhofer, S.; Harmelen, F.; Hendler, J.; Horrocks, I.; McGuinness, D. L.; Patel-Schneider, P. F.; Stein, L. A.: *OWL: Web Ontology Language Reference.* 13.11.2009.

[W3C09] W3C OWL Working Group: *OWL 2 Web Ontology Language Document Overview.* 27.10.2009.

[WGS01] Wu, J.; Gao, M.; Stojmenovic, I.: On Calculating Power-aware Connected Dominating Dets for Efficient Routing in Ad-hoc Wireless Networks. *Proceedings of the International Conference on Parallel Processing (ICPP)*, 2001, pp. 346–354.

[WHFG92] Want, R.; Hopper, A.; Falcão, V.; Gibbons, J.: The Active Badge Location System. *ACM Transactions on Information Systems*, vol. 10, no. 1, 1992, pp. 91–102.

[Win01] Winograd, T.: Architectures for Context. *Human-Computer Interaction*, vol. 16, no. 2, 2001, pp. 401–419.

[WJH97] Ward, A.; Jones, A.; Hopper, A.: A New Location Technique for the Active Office. *IEEE Personal Communications*, vol. 4, no. 5, 1997, pp. 42–47.

[WL99] Wu, J.; Li, H.: On Calculating Connected Dominating Set for Efficient Routing in Ad-hoc Wireless Networks. *Proceedings of the 3rd International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications (DIAL-M)*, DIALM '99, New York, NY and USA, 1999, pp. 7–14.

[WLBW01] Wattenhofer, R.; Li, L.; Bahl, P.; Wang, Y.-M.: Distributed Topology Control for Power-efficient Operation in Multihop Wireless Ad-hoc Networks. *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, vol. 3, 2001, pp. 1388–1397 vol.3.

[Wol14] Wolfram Research, Inc.: *Wolfram Mathematica: Definitive System for Modern Technical Computing*. 2014.

[WZGP04] Wang, X. H.; Zhang, D. Q.; Gu, T.; Pung, H. K.: Ontology Based Context Modeling and Reasoning using OWL. *Proceedings of the 2nd IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2004, pp. 18–22.

[XHE01] Xu, Y.; Heidemann, J.; Estrin, D.: Geography-informed Energy Conservation for Ad-hoc Routing. *Proceedings of the 7th International Conference on Mobile Computing and Networking (MobiCom)*, Rome and Italy, 2001, pp. 70–84.

[XK04] Xue, F.; Kumar, P. R.: The Number of Neighbors Needed for Connectivity of Wireless Networks. *Wireless Networking*, vol. 10, no. 2, 2004, pp. 169–181.

[YWK06] Yang, Y.; Wang, J.; Kravets, R.: Load-balanced Routing for Mesh Networks. *SIGMOBILE Mobile Computing and Communications Review*, vol. 10, no. 4, 2006, pp. 3–5.

[Zad65] Zadeh, L. A.: Fuzzy Sets. *Information Control*, vol. 8, no. 3, 1965, pp. 338–353.

[Zim07] Zimmermann, A.: *Context Management and Personalisation: a Tool Suite for Context-aware and User-aware Computing*. Dissertation, RWTH Aachen, 2007.

[ZP96] Zhang, N. L.; Poole, D.: Exploiting Causal Independence in Bayesian Network Inference. *Journal of Artificial Intelligence Research*, vol. 5, no. 1, 1996, pp. 301–328.

[ZZM+06] Zhdanova, A. V.; Zoric, J.; Marengo, M.; van Kranenburg, H.; Snoeck, N.; Sutterer, M.; Räck, C.; Droegehorn, O.; Arbanowski, S.: Context Acquisition, Representation and Employment in Mobile Service Platforms. *Proceedings of the IST Mobile and Wireless Communications Summit*, 2006.

List of Acronyms

3GPP	3rd Generation Partnership Project
AAA	Authentication, Authorization, and Accounting
ACK	ACKnowledgment
AI	Artificial Intelligence
AMR	Adaptive Multi-Rate
ANN	Artificial Neural Network
AoA	Angle of Arrival
AODV	Ad-hoc On-Demand Distance Vector
AP	Access Point
API	application programming interface
ARMA	Auto-Regressive Moving Average
ARPU	Average Revenue Per User
ARQ	Automatic Repeat reQuest
AS	Autonomous System
ASA	Authorized Shared Access
B.A.T.M.A.N.	Better Approach To Mobile Ad-hoc Networking
BPSK	Binary Phase-Shift Keying
BER	Bit Error Rate
BGP	Border Gateway Protocol
BN	Bayesian Network
BSC	Balanced Scorecard
BTO	Backhaul Topology Optimization
BTS	Base Transceiver Station
CAGR	Compound Annual Growth Rate
CAPEX	capital expenditures
CARMEN	CARrier grade wireless MEsh Network
CBTC	Cone-Based Topology Control
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CEDAR	Core-Extraction Distributed Ad-hoc Routing

CPS	Cyber-Physical System
CSC	Channel Switching Cost
CSMA/CD	Carrier Sense Multiple Access/Collision Detection
CSMA/CA	Carrier Sense Multiple Access/Collision Avoidance
DHCP	Dynamic Host Configuration Protocol
DNA	DeoxyriboNucleic Acid
DP	Dynamic Programming
DSDV	Destination Sequence Distance Vector
DSL	Digital Subscriber Line
DSR	Dynamic Source Routing
DTX	Discontinuous Transmission
DVB-T	Digital Video Broadcast - Terrestrial
DWD	German Weather Service (Deutscher Wetterdienst)
eBTO	extended Backhaul Topology Optimization
ECR	Energy Consumption Rating
ECRW	Weighted Energy Consumption Rating
EGP	Exterior Gateway Protocol
EIGRP	Enhanced Interior Gateway Routing Protocol
EKF	Extended Kalman Filter
ETSI	European Telecommunications Standards Institute
EU	European Union
EPI	Energy Proportionality Index
ETT	Expected Transmission Time
ETX	Expected Transmission count
FDM	Frequency Division Multiplex
FDMA	Frequency Division Multiple Access
FER	Frame Error Rate
FP7	Framework Programme 7
FSL	Free Space Loss
FSM	Finite State Machine
FTP	File Transfer Protocol
GA	Genetic Algorithm
GAF	Geographic Adaptive Fidelity
GDP	Gross Domestic Product
GPS	Global Positioning System
GSM	Global System for Mobile Communications
GUI	Graphical User Interface
GW	GateWay

HC	Hop Count
HSPA	High-Speed Packet Access
HTTP	HyperText Transfer Protocol
ICT	Information and Communication Technology
ID	IDentification
IEEE	Institute of Electrical and Electronics Engineers
IIR	Infinite Impulse Response
IP	Internet Protocol
IRU	Interference-aware Resource Usage
ISD	Inter-Site Distance
ISM	Industrial, Scientific, and Medical
IT	Information Technology
KF	Kalman Filter
KPI	Key Performance Indicator
LEM	Lifetime Extension Measure
LILT	Local Information Link-State Topology
LINT	Local Information No Topology
LOS	Line Of Sight
LP	Linear Program
LQSR	Link Quality Source Routing
LSR	Label-Switched Router
LTE	Long Term Evolution
LTE-A	LTE Advanced
MA	Moving Average
MAC	Medium Access Control
MAN	Metropolitan Area Network
MANET	Mobile Ad-hoc NETwork
MBMS	Multimedia Broadcast Multicast Service
MCS	Modulation and Coding Scheme
MIC	Metric of Interference and Channel switching
MIMO	Multiple Input and Multiple Output
MPLS	Multi Protocol Label Switching
MPLS-TE	Multi Protocol Label Switching - Traffic Engineering
MPR	Maximum Path Redundancy
MSDU	MAC Service Data Unit
MST	Minimum Spanning Tree
MTU	Maximum Transmission Unit
NAP	Neighbor Addition Protocol

NIC	Network Interface Card
NLOS	Non-Line Of Sight
NLP	Non-Linear Program
NP	Non-deterministic Polynomial time
NRP	Neighbor Reduction Protocol
OFDMA	Orthogonal Frequency Division Multiple Access
OFDM	Orthogonal Frequency Division Multiplex
OLSR	Optimized Link State Routing
OPEX	OPerational EXpenditures
OSGi	*originally* Open Services Gateway initiative
OSI	Open Systems Interconnection
OWL	Web Ontology Language
PA	Power Amplifier
PER	Packet Error Rate
PMP	Point-to-MultiPoint
PSTN	Public Switched Telephone Network
PV	PhotoVoltaic
QAM	Quadrature Amplitude Modulation
QAP	Quadratic Assignment Problem
QoE	Quality of Experience
QoS	Quality of Service
QP	Quadratic Program
RAN	Radio Access Network
RDF	Resource Description Framework
REST	REpresentational State Transfer
RF	Radio Frequency
RFC	Request For Comments
r.h.s.	right hand side
RRM	Radio Resource Management
RSSI	Received Signal Strength Indicator
Rx	Receiver
SA	Simulated Annealing
SDN	Software-Defined Networking
SINR	Signal-to-Interference-plus-Noise Ratio
SIS	Sequential Importance Sampling
SISO	Single Input and Single Output
SLA	Service Level Agreement
SME	Small or Medium-sized Enterprise

SNR	Signal-to-Noise Ratio
SOAP	*originally* Simple Object Access Protocol
SOM	Self-Organizing Maps
SON	Self-Organizing Networks
SQP	Sequential Quadratic Programming
SVM	Support Vector Machine
TC	Topology Control
TCP	Transmission Control Protocol
TEER	Telecommunications Energy Efficiency Ratio
TEEER	Telecommunications Equipment Energy Efficiency Rating
ToA	Time of Arrival
Tx	Transmitter
UKF	Unscented Kalman Filter
UMTS	Universal Mobile Telecommunications System
UT	User Terminal
VDSL	Very High Speed Digital Subscriber Line
VoIP	Voice over IP
WAN	Wide Area Network
WBN	Wireless Backhaul Network
WCDMA	Wideband Code Division Multiple Access
WCETT	Weighted Cumulative Expected Transmission Time
WLAN	Wireless Local Area Network
WMN	Wireless Mesh Network
WRP	Wireless Routing Protocol
WSN	Wireless Sensor Network
XML	eXtensible Mark-up Language
ZHLS	Zone-based Hierarchical Link State
ZRP	Zone Routing Protocol

List of Symbols

$\mathcal{E}$	set of edges of graph G
$\mathbb{N}$	set of natural numbers including 0
$\mathcal{N}(\mu, \sigma)$	normal distribution with mean μ and standard deviation σ
$\mathcal{O}(\cdot)$	order of a function (*big O notation*)
$\mathcal{P}_{i,j}$	set of paths from v_i to v_j
$\mathbb{R}$	space of real numbers
$\mathbb{R}^n$	n-dimensional vector space of real numbers, $n \geq 2$
$\mathcal{S}$	set of nodes v_i with $d(v_i) < d_{\max}$
$\mathcal{S}^-$	set of nodes v_i with $d(v_i) > d_{\max}$
$\mathcal{T}$	set of time instances
$\mathcal{V}$	set of vertices of graph G
$\mathcal{V}_{\mathrm{GW}}$	set of gateway nodes in G
$\mathcal{X}_i$	set of radio interfaces at node v_i
α	weighting coefficient
$\boldsymbol{\beta}$	m-dimensional vector of battery levels (in %) at nodes v_i
χ	edge connectivity of graph G
Δ_π	change in transmission power
$\delta_{i,j}$	Euclidean distance between vertices v_i and v_j
$\Phi(\tau)$	auto-correlation function
φ	angle between two edges in G
γ	approximation level of a heuristic algorithm
η	efficiency (e.g., of a photovoltaic module)
λ	memory degree of a stationary Gaussian process
ν	data rate (throughput)
Π	permutation function
$\boldsymbol{\pi}$	m-dimensional vector of instantaneous power consumption at nodes v_i
π_i	instantaneous power consumption, power assignment

	function of node v_i, $\pi_i : X_i \to [0, \pi_{\max}]$
ς	realization of a random variable with standard normal distribution
Θ	Gauss-Markov process
θ	angle of user movement, $\theta = \arctan\left(\frac{v_y}{v_x}\right)$
ϑ	angle of aperture
$\rho_{i,j}$	capacity utilization of link between nodes v_i and v_j
$\mathbf{\Psi}(G) = (\psi_{i,j})$	incidence matrix of G
ζ	fraction of nonoperative nodes due to drained batteries
A	area
$\mathbf{A}(G) = (a_{i,j})$	adjacency matrix of graph G
$\mathbf{C}(G) = (c_{i,j})$	capacity matrix of graph G
$c(\theta, \varphi)$	correction factor for solar panel tilt θ and orientation φ
$c(\mathbf{A})$	maximum flow ($s-t$-flow) of a graph with adjacency matrix $\mathbf{A}$
$\mathbf{D}(G)$	distance matrix of graph G
$d_G(v)$, $d(v)$	degree of vertex v
E	energy
e_{i_1,i_2}	edge between vertices v_{i_1} and v_{i_2}
e_j	jth edge in $\mathcal{E}$
G	undirected graph, $G = (\mathcal{V}, \mathcal{E})$
g	percentage of gateway nodes in a WBN
H	yearly global irradiation, $[H] = 1\,\frac{\mathrm{kWh}}{\mathrm{m^2\,a}}$
m	number of vertices in graph G
$N_{x,i}$	number of radio interfaces at node v_i
n	number of edges in graph G
P	set of coordinates
$\mathbf{p}(v)$	coordinates of vertex v
PR	performance ratio of a photovoltaic module
Q	electric charge, capacity of a battery
$\mathbf{R} = (r_{i,j})$	{0-1}-reachability matrix of graph G, $r_{i,j} = 1$ if v_j is within radio coverage of v_i
$s(t)$	speed of an object at time t, $s(t) = \|\mathbf{v}(t)\|$
t	time instance
U	utility or objective function
U^A	objective function value for the solution of the approximation algorithm
U^*	objective function value for the optimal solution

$\mathbf{v}(t) = (v_x(t), v_y(t))$	(two-dimensional) velocity of an object at time t
v_i	ith vertex in V
$\mathbf{W}_{i,j} = (w_{i,j})$	edge weight matrix of graph G
Y	yearly energy yield of photovoltaic module, $[Y] = 1\frac{\text{kWh}}{\text{a}}$

Appendix A

Fundamental Concepts of Graph Theory

Table A.1 contains descriptions and definitions of graph theoretical concepts relevant in the scope of this thesis [Die06].

Table A.1: Collection of concepts from graph theory

Concept	**Notation**	**Definition**
Graph	$G = (\mathcal{V}, \mathcal{E})$	A graph is a pair $G = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ being a set of vertices and $\mathcal{E} = \{e_1, e_2, \ldots, e_m\}$ a set of edges (pairs of vertices). The edge $e = \{v_1, v_2\}$ is often written $e = v_1v_2$.
Edge weight	$w(e)$	The function $w : \mathcal{E} \to \mathbb{R}$ assigns a weight to each $e \in \mathcal{E}$. Edge weights have context-specific meaning (e.g. distance or cost).
Vertex degree	$d_G(v)$	The degree $d_G(v)$ or $d(v)$ of a vertex v is the number of edges it is connected to. A graph is regular if all vertices have the same degree, and k-regular if all vertices have degree k.
Order and size	$\lvert\mathcal{V}(G)\rvert$ and $\lvert\mathcal{E}(G)\rvert$	The number of vertices in a graph is denoted by $\lvert\mathcal{V}(G)\rvert$ or m, and is sometimes called the *order*; the number of edges is denoted by $\lvert\mathcal{E}(G)\rvert$ or n and is called size of G.
Path and cycle	$\langle v_1, \ldots, v_m\rangle$	A graph on m vertices is called a path (of length $m-1$) if the vertices can be labeled as $v_1, \cdots, v_m$ in such a way that $\mathcal{E} = \{v_iv_{i+1} : \ i = 1, \ldots, m-1\}$. It is sometimes written $\langle v_1, \ldots, v_m\rangle$. It is called a cycle (of length m) if the vertices can be labeled as $v_1, \ldots, v_m$ in such a way that $E = \{v_iv_{i+1} : \ i = 1, \ldots, m-1\} \cup \{v_mv_1\}$. A cycle is odd/even if m is odd/even.
Subgraph (supergraph)	H	H is a subgraph of G if $\mathcal{V}(H) \subseteq \mathcal{V}(G)$, $\mathcal{E}(H) \subseteq \mathcal{E}(G)$ and H is a graph. G is a supergraph of H.

Table A.1: Collection of concepts from graph theory (continued)

Concept	Notation	Definition
k-partite graph	N/A	A graph G is k-partite if it is possible to partition the vertex set $\mathcal{V}$ into k disjoint subsets $\mathcal{V}_1, \ldots, \mathcal{V}_k$ in such a way that no two vertices in a subset are adjacent, i.e. no two vertices in the subset share an edge (i.e., the induced subgraphs $G[\mathcal{V}_1], \ldots, G[\mathcal{V}_k]$ are all empty).
Adjacency matrix	$A(G)$	The adjacency matrix of a graph $G = (\mathcal{V}, \mathcal{E})$ is the $m \times m$ $(0-1)$-matrix $\mathbf{A}(G) = (a_{i_1,i_2})$, $i_1, i_2 = 1, \ldots, m$, whose rows and columns are labeled by vertices, with entry $a_{i_1,i_2} = 1$ if and only if $e_{i_1,i_2} \in \mathcal{E}$.
Incidence matrix	$M(G)$	The incidence matrix is the $m \times n$ $\{0-1\}$-matrix $\Psi(G) = (\psi_{i,j})$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ whose rows are labeled by vertices and whose columns are labeled by edges, with entry $\psi_{i,j} = 1$ if and only if v_i is incident to e_j.
Isomorphism	$G \cong H$	Two graphs G, H are isomorphic if there is a bijection $f : \mathcal{V}(G) \to \mathcal{V}(H)$ which preserves adjacency: $f(v_1)f(v_2) \in \mathcal{E}(H)$ if and only if $v_1v_2 \in \mathcal{E}(G)$. An isomorphism induces a bijection from $\mathcal{E}(G)$ to $\mathcal{E}(H)$. Necessarily, isomorphic graphs must have the same size.
Neighborhood	$N(v)$	For graph $G = (\mathcal{V}, \mathcal{E})$, the neighborhood of vertex v is $N(v) = \{v_i \vert v_i v \in \mathcal{E},\ v_i \in \mathcal{V} \backslash v\}$.
Clique	K_n	A *maximum clique* of graph G is a complete subgraph of G with the maximum number of vertices. A *maximal clique* of G is a complete subgraph of G that is not contained in any larger complete subgraph.
Graph distance	$\delta_G(v_1, v_2)$	The graph distance (hop count) between vertices v_1 and v_2 is the length of a shortest $v_1; v_2$-path in G.
Eccentricity	ε_v	The eccentricity of vertex v is $\varepsilon_v = \max_{v_i \in \mathcal{V}} \delta_g(v_i, v)$
Centrality		
Diameter	D_G	The diameter of G is $D_G = \max_{v_{i_1}, v_{i_2} \in \mathcal{V}} \delta_G(v_{i_1}, v_{i_2})$
Girth	N/A	The minimum length of a cycle in G.
Connected graph	N/A	A graph G is connected if, between each pair of vertices, there is a path.
Directed graph	N/A	A directed graph or digraph is a graph where the edges have a direction associated with them.
Acyclic graph	N/A	A graph G is acyclic if it has no cycle.
Tree graph	N/A	A graph G is a tree if it is connected and acyclic.

Table A.1: Collection of concepts from graph theory (continued)

Concept	**Notation**	**Definition**
Spanning tree	$T(G)$	A spanning tree $T(G)$ of a connected, undirected graph G is a tree composed of all the vertices and some (or perhaps all) of the edges of G.
Multigraph	N/A	A multigraph or pseudograph is a graph which is permitted to have multiple edges, (also called "parallel edges"), i.e., edges that have the same end nodes.

Appendix B

Routing in Wireless Backhaul Networks

B.1 Routing Protocols

In the following, fundamental routing approaches as well as example protocols are depicted.

B.1.1 Reactive Routing Protocols

Reactive routing protocols compute routing paths on demand basis. A route discovery operation invokes a route-determination procedure. The discovery procedure terminates either when a route has been found or when no route is available after examination of all route permutations. In networks with mobile nodes, e.g., MANETs, active routes may be temporarily unavailable or become completely obsolete. In contrast, WMNs and WBNs in particular, consist of nodes with minimal or no mobility at all. In this context, reactive routing protocols, due their reduced overhead, have better scalability than proactive routing protocols; however, they are usually accompanied by increased delays. Examples for reactive routing protocols include DSR protocol, its extension Link Quality Source Routing (LQSR) [JM96], and AODV protocol [PBRD03].

B.1.2 Proactive Routing Protocols

As indicated by the naming, networks using proactive routing protocols continuously evaluate routes to all reachable nodes and attempt to maintain consistent, up-to-date routing information. Since this is done in a distributed manner by every node, a routing path is immediately available at

each node. Moreover, it is strictly required that all nodes maintain a consistent view of the network topology. In case of a topology change, according updates must be propagated throughout the network. The algorithms underlying the protocols originate from routing principles for wired networks. Due to the rather stationary nature (no node mobility, relatively constant link quality) of WBNs, which also compensates for the inherently high overhead of proactive routing protocols, operators usually prefer them over reactive protocols in backhaul routing. Moreover, they blend well with the central control paradigm postulated by operators. Examples for proactive routing protocols include Destination Sequence Distance Vector (DSDV) protocol [PB94], Optimized Link State Routing (OLSR) protocol [CJ03], and HEAT [BHLM07].

B.1.3 Hybrid Routing Protocols

By utilizing principles from reactive and proactive routing, hybrid protocols aim at combining the respective merits and eliminate shortcomings. They are well suited for hierarchically structured WMNs where the network is partitioned into several zones and different approaches are taken for intra-zone and inter-zone routing. Most often, proactive techniques are applied on lower hierarchical levels (routing within a zone), reactive techniques on higher levels (routing across zones). Zone-based Hierarchical Link State (ZHLS) protocol [JNL99], Zone Routing Protocol (ZRP) [HPS02], and the layer 2 protocol B.A.T.M.A.N. advanced [The13] are examples for hybrid routing protocols. Summarizing, Figure B.1 depicts a classification of the most common routing protocols along two dimensions, namely, the discussed approaches to route computation and, as a second common criterion, the kind of hierarchy the protocol enforces on a network. With respect to this second dimension, routing protocols either divide network nodes into sets of nonuniform importance (e.g., core vs. edge nodes) or assume a flat network, i.e., all nodes are of equal significance and basically share the same functionality. Generally, reactive routing protocols tend to be more eligible for smaller, less structured, flat networks, such as, MANETs or WSNs, whereas proactive routing approaches are more compatible to larger, more structured, hierarchical networks where QoS levels and SLAs have to be guaranteed, e.g., carrier-class backhaul and backbone networks. Due to their increased design overhead, hybrid protocols are usually applied in application-specific scenarios with custom-tailored protocol

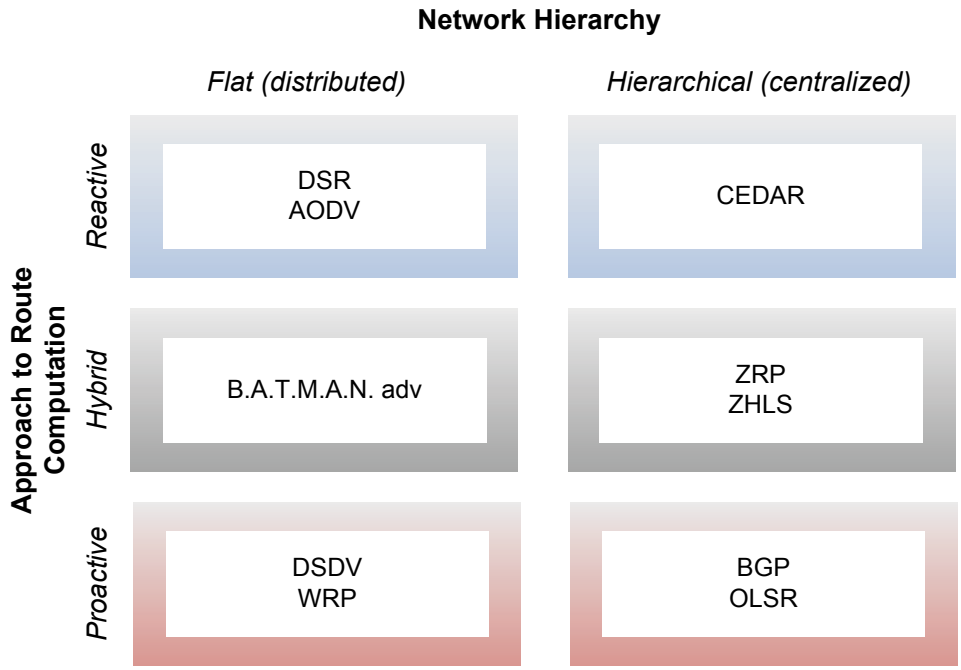


Figure B.1: Classification of routing protocols

characteristics. However, whenever feasible and appropriate, exceptions to these general rules exist, such as the proactive OLSR protocol, which is also very common for routing in MANETs. Moreover, protocols such as Core-Extraction Distributed Ad-hoc Routing (CEDAR) [SSB99] exploit hierarchical network structures to improve reactive routing and Wireless Routing Protocol (WRP) is an example for a proactive protocol explicitly developed for flat networks. Overall, due to the characteristics of the discussed routing approaches, proactive protocols are the preferred choice for coordinated, carrier-grade WBNs.

B.2 Routing Metrics

Routing usually requires metrics that are used to unanimously determine a ranking of available routes. The following paragraphs elaborate on the most common metrics used in today's network routing protocols.

Hop Count

Hop count is a metric that simply counts and minimizes the number of intermediate nodes from source to destination node. Being among the most commonly used routing metrics, it is used in protocols such as, DSR, AODV, or DSDV. It reflects the effects of path lengths on routing perfor-

mance and can find loop-free paths with minimum hop count. As a major drawback, hop count consider neither capacity, nor latency, nor packet error and loss rates of wireless links, thus neglecting important QoS parameters of available paths. Hence, exclusively using a hop count metric may result in suboptimal routing performance.

Expected Transmission Count

The Expected Transmission count (ETX) counts the amount of MAC layer frames required for a successful transmission of a packet for each point-to-point link. The metric for a path of multiple hops through a network is defined as the ETX sum over all links on the path. Consequently, both larger hop count and higher PERs result in a worse ETX metric. Similar to the hop count measure, protocols using ETX can compute optimum paths rather quickly and avoid loops. ETX has been proposed in [CABM05] and has been shown to outperform the hop count metric when used in protocols such as DSDV. As a major drawback, ETX performs less efficient in network setups containing links with different data rates or nodes with multiple radios.

Expected Transmission Time

As an extension to ETX, Expected Transmission Time (ETT) incorporates the capacity of a wireless channel into the routing metric. In [DPZ04], ETT_i of a link i is defined as the duration of a successful transmission of a packet. For a multi-hop path, the respective individual figures are added to compute overall $\mathrm{ETT} = \sum_i \mathrm{ETT}_i$. Given packet size s and data rate ν_i of link i, ETT_i directly relates to ETX_i according to $\mathrm{ETT}_i = \mathrm{ETX}_i \cdot \frac{s}{d_i}$. While exhibiting the same advantages as ETX and additionally including channel capacity, it does consider the possibility of interference problems in single radio environments. This limitation is tackled by Weighted Cumulative Expected Transmission Time (WCETT).

Weighted Cumulative Expected Transmission Time

In multi-radio environments, links can use separate channels to reduce interference, thereby improving routing performance. In order to distinguish between hops that are on different channels, i.e., to account for channel diversity, [DPZ04] introduces the notion of WCETT. This metric is com-

posed of two separate terms. The first one, weighted with $1-\alpha$, $\alpha \in [0,1]$, is ETT, as described in Section B.2. The second term, weighted by α, calculates the bottleneck channel for a given multi-hop path given that each link i on the path uses one of the channels $j = 1, \ldots, k$, i.e., the channel with the maximum ETT is determined:

$$\text{WCETT} = (1-\alpha)\sum_i \text{ETT}_i + \alpha \max_j \sum_{i|\text{link } i \text{ uses channel } j} \text{ETT}_i,\ \forall\, 1 \le j \le k. \tag{B.1}$$

The second term is designed in a way as to increase if more hops of a path transmit on the same channel, i.e., it searches for paths with higher channel diversity thus reducing the number of nodes transmitting on the same frequency. The parameter α can be tuned according to the specific requirements. Since WCETT inherently lacks the characteristic of isotonicity [Sob02], no algorithm exists that can efficiently compute the WCETT-optimal path between two nodes in an arbitrary network.

Metric of Interference and Channel-Switching

The Metric of Interference and Channel switching (MIC), as proposed in [YWK06], improves WCETT by additionally incorporating inter-path interference. For a path p consiting of edges e_j and nodes v_i, the metric is defined as follows:

$$\text{MIC}\,(p) = \frac{1}{m\ \min(\text{ETT})} \sum_{j\,|\,e_j \in p} \text{IRU}_j + \sum_{i\,|\,v_i \in p} \text{CSC}_i, \tag{B.2}$$

where m is the total number of nodes in the network and min(ETT) is the smallest ETT in the network, which can be estimated based on the lowest transmission rate of the wireless network cards. The two contributors to MIC, Interference-aware Resource Usage (IRU) and Channel Switching Cost (CSC), are defined as:

$$\text{IRU}_j = \text{ETT}_j \times N(j), \tag{B.3}$$

$$\text{CSC}_i = \begin{cases} w_1, & \text{if} \quad \text{CH}(\text{prev}(i)) \neq \text{CH}(i) \\ w_2, & \text{if} \quad \text{CH}\,(\text{prev}(i)) = \text{CH}(i) \end{cases}, 0 \le w_1 \le w_2, \tag{B.4}$$

where N_j is the number of neighboring nodes experiencing interference to their links by the transmission on link j by node i. Essentially, IRU_i denotes

the aggregated time that neighboring nodes use for their transmissions on the frequency also occupied by link j, thus capturing the *inter-flow* interference. Therefore, IRU favors a path that results in less interference for neighbors of node i. Further, $\mathrm{CH}(i)$ depicts the channel used by node i for transmitting to the successive node along path p, while $\mathrm{prev}(i)$ represents the channel used for receiving from the previous node. Consequently, CSC is a measure for *intra-flow* interference since paths where consecutive links use the same channel receive a worse evaluation ($\mathrm{CSC}_i = w_2$) than paths that alternate their channel assignments ($\mathrm{CSC}_i = w_1$), essentially favoring paths with more diversified channel assignments. In the context of WBNs, MIC is particularly useful since it is able to adequately reflect the advantages of multi-radio multi-channel wireless point-to-point networks.

Appendix C

Context Management Functions

In the following, as a an extension to Section 4.3.2, a selection of fundamental functions of context management systems is presented.

Context Source Discovery Mechanism

Since context brokers and sources mutually need to be aware of each other, context management systems support protocols that on the one hand allow sources to register with brokers and, on the other hand, allow brokers to regularly check if sources are still available. Common approaches for the former include direct advertisements or broadcasted announcements sent by context sources. These messages include a basic description of the source, i.e., IDentification (ID), position, network address, and provided context information. An XML advertisement example as used in EU Framework Programme 7 (FP7) project *C-CAST* [CC08] is depicted in Listing C.1. The latter is usually realized by regular so-called "alive" messages by sources or according requests by the broker.

Context Source Registry

Context brokers maintain tables for listing registered context sources and their description. This registry works similar to a telephone directory, where any kind of system entity can apply filters to look up requested data.

Context Data Life Cycle Policies

Context data life cycle policies assure a reliable data categorization and according processing across the entire platform. Besides age of data, demand, availability and security requirements, importance, size, and type are relevant criteria for categorization.

Listing C.1: ContextML Advertisement as Implemented in EU FP7 Project *C-CAST*

```
<?xml version="1.0" encoding="utf-8"?>
<contextML>
  <ctxAdvs>
    <ctxAdv>
      <contextProvider id="UKl_TNetworkCxP" v="0.1"/>
      <urlRoot>http://www.eit.uni-kl.de/wicon/TerminalkCxP</urlRoot>
      <scopes>
        <scopeDef n="wifi">
          <url>getCurrentContext</url>
          <entityTypes>imei</entityTypes>
          <inputDef>
            <inputEL name="scope" type="String"/>
            <inputEL name="entity" type="String"/>
          </inputDef>
        </scopeDef>
      </scopes>
    </ctxAdv>
  </ctxAdvs>
</contextML>
```

Context Data Cache and History

A **cache** secures a reliable availability and a fast dissemination of context data. Many systems and services (particularly network transport services) increasingly rely on close-to-real-time information in order to fully exhibit their designed behavior. However, due to limited cache capacity, outdated as well as less time-critical context data is moved to a data repository, usually a **database**. In many cases, this involves additional data processing (e.g., compression) and a change of representation format to store data more efficiently.

Data and Knowledge Representation

Context information needs to be represented and modeled for being machine interpretable and exchangeable using defined interfaces. The goals are to support easy manipulation, easy extension, efficient search and query access and scalability. In many cases, representation is often tailored according to the problem domain, to the specific objectives of the system, or to utilized inference mechanisms. Strang and Linnhoff-Popien [SLP04] identify generic requirements: The modeling approach should (1) be able to cope with high dynamics and distributed processing and composition, (2) allow for partial validation independently of complex interrelationships, (3) enable rich expressiveness and formalism for a shared understanding, (4) indicate richness and quality of information, (5) not assume completeness and unambiguousness, (6) be applicable to existing infrastructures and frameworks.

According to [BDR07], context models can be classified into six different model categories, namely key-value models, markup scheme models (e.g., XML [MLK01, BPSM+08]), graphical models (e.g., Unified Modeling Language, *UML* [Obj07]), object oriented models (e.g., [HSP+03]), logic-based models (e.g., [MB94]), and ontology based models (e.g., Resource Description Framework Schema, RDFS [BG04], OWL [vHH+09], OWL 2 [W3C09], see also Section 4.3.3).

Mechanisms for Privacy, Security, and Trust

In analogy to similar functions in cellular communication systems, AAA are typical means to prevent provisioning data to unauthorized persons or services, hence protecting privacy. A requester must prove both his identity and his authorization in order to access data. Accounting procedures provide transparency by tracking and logging who has accessed which data at what time. If user shall be convinced to provide their data to context management systems, reliable AAA mechanisms establish user trust in the system. Additionally, today's systems incorporate auditing, billing, and charging functions, allowing for shared utilization of common resources.

Appendix D

Optimization Techniques

D.1 Linear Optimization

A sample linear program for a two-dimensional (x_1,x_2) solution space is given as follows:

$$\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & -(12x_1 + 7x_2) && \text{(D.1)}\\
\text{s.t.} \quad & -12x_1 + 18x_2 \geq 171 && \text{(D.2)}\\
& 24x_1 + 30x_2 \geq -105 && \text{(D.3)}\\
& -x_1 - \ 5x_2 \geq -50 && \text{(D.4)}\\
& -12x_1 - \ 5x_2 \geq -105 && \text{(D.5)}\\
& 5x_1 - \ 4x_2 \geq 62. && \text{(D.6)}
\end{aligned}$$

Figure D.1 provides the graphical solution to the problem. The linear objective function (D.1) is depicted by the straight line. For finding the optimum, the line has to be shifted in parallel (in case of minimization in the direction of decreasing objective function values) until it reaches a corner or edge of the feasible region. The region of feasible solutions is limited by a two-dimensional polygon (a pentagon in this example) which in turn is derived by the inequality constraints (D.2) through (D.6). A parallel shift of the line to the upper right is equivalent to decrease of the objective function value. The optimum is depicted on the right of Figure D.1, where the straight line reaches a corner of the pentagon, which is, at the same time, the optimum solution $(\hat{x}, \hat{y})$. With a further shift, the line would leave the feasible region [Bou13]. Maximization problems can be transformed into minimization problems using a linear transformation to be applied on the LP. Moreover, since solving non-linear problems can be extremely challenging and resource-intensive, a common approach is to linearize these

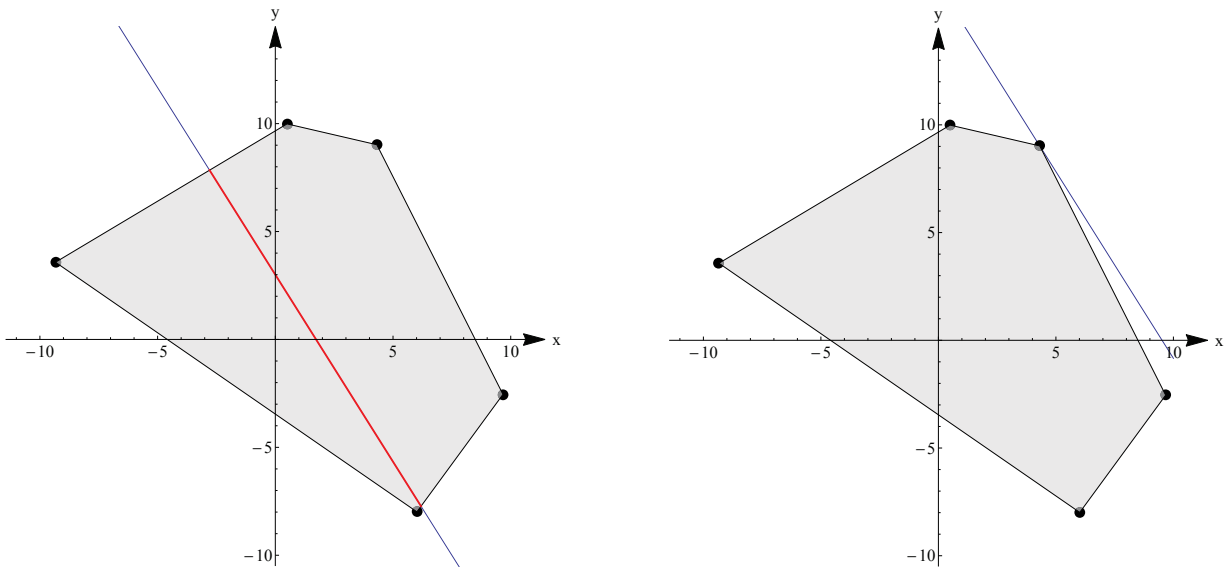


Figure D.1: Illustration of a simple linear program with two variables (x,y) and five inequality constraints

problems using according transformation techniques. However, in many cases, such a linearization is accompanied by a structural change of the problem, i.e., the solution of the LP only serves as an approximation for the optimum solution and, depending on the design of the LP, does not necessarily fall into the feasible region of the original problem [NT93].

D.2 Clustering Techniques

Multivariate clustering is a statistical method for dividing a set of multi-dimensional data objects into subsets or so called clusters. Clustering algorithms can be classified according to the way they assign objects to clusters (hard or fuzzy) or the way they are initialized (fix or flexible amount of clusters). Frequently used techniques include K-Means [Ste56], Fuzzy C-Means [Bez81], Neural Gas [MS91][Fri97], or Growing Neural Gas [Fri95], depending on application area and data characteristics. Generally, any clustering algorithm has two distinct objectives:

- **Homogeneity within a cluster** - objects assigned to the same cluster shall be as similar as possible.
- **Heterogeneity between clusters** - objects assigned to different clusters shall be as dissimilar as possible.

Table D.1: Common distance metrics in clustering algorithms

Name	**Function** ($\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $p \in \mathbb{R}$)
Minkowski distance	$\sqrt[p]{\sum_{i=1}^{n}(x_i - y_i)^p}$
Manhattan distance ($p = 1$)	$\sum_{i=1}^{n} \lvert(x_i - y_i)\rvert$
Euclidean distance ($p = 2$)	$\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$
Chebyshev distance ($p \to \infty$)	$\max_i \lvert x_i - y_i\rvert$
Cosine similarity	$\cos(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x} \cdot \mathbf{y}}{\lVert\mathbf{x}\rVert \lVert\mathbf{y}\rVert} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}}$

The extend of similarity between two objects is calculated based on similarity or distance metrics. Table D.1 lists common distance metrics utilized in clustering. Some authors, e.g., [Fri98], differentiate between clustering and classification, where the latter produces a-priori information on the resulting classes. However, within the scope of this work, where clustering algorithms are applied for structuring a wireless backhaul network and combining individual nodes in groups, the author will assume the alternative approach to use both terms interchangeably.

D.3 Dynamic Programming

As introduced in 5.1.2, the subproblems are modeled as a sequence of state transitions of length N. At each step k of the sequence, the decision vector $\mathbf{u}_k$ shall be chosen so as to maximize the overall objective function

$$B_N = h(\mathbf{x}_0, \mathbf{u}_1) + h(\mathbf{x}_1, \mathbf{u}_2) + \cdots + h(\mathbf{x}_{N-1}, \mathbf{u}_N) + z(\mathbf{x}_N). \tag{D.7}$$

$z(\cdot)$ is a scalar-valued function to evaluate the utility of the final state $\mathbf{x}_N$. It becomes clear that the optimal value b of B_N depends on the initial state $\mathbf{x}_0$ as well as the consecutive decisions $\mathbf{u}_k$ taken at each transition:

$$b(\mathbf{x}_0) = \max_{\{\mathbf{u}_k | k=1,\ldots,N\}} B_N. \tag{D.8}$$

In case of $N = 1$, Equation (D.8) reduces to $b(\mathbf{x}_0) = \max_{\mathbf{u}_1} (b(\mathbf{x}_0, \mathbf{u}_1) + z(\mathbf{x}_1))$. For $N \geq 2$, b can be written as

$$\begin{aligned} b(\mathbf{x}_0) &= \max_{\{\mathbf{u}_1,\ldots,\mathbf{u}_N\}} (h(\mathbf{x}_0, \mathbf{u}_1) + h(\mathbf{x}_1, \mathbf{u}_2) + \cdots + h(\mathbf{x}_{N-1}, \mathbf{u}_N) + z(\mathbf{x}_N)) \\ &= \max_{\mathbf{u}_1} (\max_{\{\mathbf{u}_2,\ldots,\mathbf{u}_N\}} (h(\mathbf{x}_0, \mathbf{u}_1) + \cdots + h(\mathbf{x}_{N-1}, \mathbf{u}_N) + z(\mathbf{x}_N))) \\ &= \max_{\mathbf{u}_1} (h(\mathbf{x}_0, \mathbf{u}_1) + \max_{\{\mathbf{u}_2,\ldots,\mathbf{u}_N\}} (h(\mathbf{x}_1, \mathbf{u}_2) + \cdots + h(\mathbf{x}_{N-1}, \mathbf{u}_N) + z(\mathbf{x}_N))) \\ &= \max_{\mathbf{u}_1} (h(\mathbf{x}_0, \mathbf{u}_1) + b(\mathbf{x}_1 = t(\mathbf{x}_0, \mathbf{u}_1)). \end{aligned} \tag{D.9}$$

Removing the indexes yields the *Bellman Equation* [BK65]

$$b(\mathbf{x}) = \max_{\mathbf{u}} (h(\mathbf{x}, \mathbf{u}) + b(t(\mathbf{x}, \mathbf{u})), \tag{D.10}$$

which is commonly solved using numerical backward induction [BD66].

D.4 Sequential Quadratic Programming

Sequential Quadratic Programming (SQP) techniques serve as a means to approximate optimal solutions for non-linear optimization problems. As a starting point, the NLP (with m equality and p inequality constraints) is given as:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) && \text{(D.11a)} \\ \text{s.t.} \quad & \mathbf{h}(\mathbf{x}) = \mathbf{0} && \text{(D.11b)} \\ & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, && \text{(D.11c)} \end{aligned}$$

where $x \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^m$, and $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^p$ [Sca85]. In order to replace the NLP by an SQP, the first step involves exchanging the objective function (D.11a) with the scalar-valued Langrangian function

$$\mathcal{L}(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + < \mathbf{u}, \mathbf{h}(\mathbf{x}) > + < \mathbf{v}, \mathbf{g}(\mathbf{x}) >, \tag{D.12}$$

where $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^p$ are the optimal multiplier vectors. Since they are not known, $\mathbf{u}^k$ and $\mathbf{v}^k$ serve as their (iterative) approximations, yielding $(\mathbf{x}^k, \mathbf{u}^k, \mathbf{v}^k)$ as the variables of the Lagrangian function for current iteration k. The Langrangian is used as the objective function of the SQP since the resulting iterates $\mathbf{x}_k$ exhibit good convergence properties. Secondly, the second-order Taylor series approximation of Equation (D.12) about $\mathbf{x}^k$ can

be computed as

$$\mathcal{L}(\mathbf{x},\mathbf{u},\mathbf{v}) \approx \mathcal{L}(\mathbf{x}^k,\mathbf{u}^k,\mathbf{v}^k) + < \nabla\mathcal{L}(\mathbf{x}^k,\mathbf{u}^k,\mathbf{v}^k), \mathbf{d_x} > + \frac{1}{2}\,\mathbf{d_x}^T\mathbf{H}(\mathcal{L}(\mathbf{x}^k,\mathbf{u}^k,\mathbf{v}^k))\,\mathbf{d_x}, \tag{D.13a}$$

where $\mathbf{H}(\cdot)$ denotes the Hessian matrix of a scalar function and $\mathbf{d_x} = \mathbf{x} - \mathbf{x}^k$. In a third step, the constraints are replaced with their first order Taylor series approximations about the current $\mathbf{x}_k$. Since the variables change from $\mathbf{x}$ to $\mathbf{d_x}$, the first term of the right hand side (r.h.s.) of Equation (D.13a) can now also be omitted and the quadratic subproblem of (D.11) can be written as

$$\begin{aligned}
\underset{\mathbf{d_x}}{\text{minimize}} \quad & < \nabla\mathcal{L}(\mathbf{x}^k,\mathbf{u}^k,\mathbf{v}^k), \mathbf{d_x} > + \frac{1}{2}\,\mathbf{d_x}^T\mathbf{H}(\mathcal{L}(\mathbf{x}^k,\mathbf{u}^k,\mathbf{v}^k))\mathbf{d_x} && \text{(D.14a)}\\
\text{s.t.} \quad & \mathbf{J}(\mathbf{h}(\mathbf{x}^k))\mathbf{d_x} + \mathbf{h}(\mathbf{x^k}) = \mathbf{0} && \text{(D.14b)}\\
& \mathbf{J}(\mathbf{g}(\mathbf{x}^k))\mathbf{d_x} + \mathbf{g}(\mathbf{x^k}) \leq \mathbf{0}, && \text{(D.14c)}
\end{aligned}$$

where $\mathbf{J}(\cdot)$ denotes the Jacobian matrix of a vector-valued function [BT95]. In the iterative process of SQP, the solution for $\mathbf{d_x}$ is used to compute $\mathbf{x}^{k+1}$. Similarly, the updated multipliers $\mathbf{u}^{k+1}$ and $\mathbf{v}^{k+1}$ can be computed as a weighted average of the previous multipliers and the available optimal multipliers $\mathbf{u}_{QP}$ and $\mathbf{v}_{QP}$ of QP (D.14) [Sch83]:

$$\begin{aligned}
\mathbf{x}^{k+1} &= \mathbf{x}^k + \xi\,\mathbf{d_x}, && \text{(D.15a)}\\
\mathbf{u}^{k+1} &= (1-\xi)\,\mathbf{u}^k + \xi\,\mathbf{u}_{QP}, && \text{(D.15b)}\\
\mathbf{v}^{k+1} &= (1-\xi)\,\mathbf{v}^k + \xi\,\mathbf{v}_{QP}, && \text{(D.15c)}
\end{aligned}$$

where ξ is a step length parameter. Further, derivatives, constraints, and Lagrangian function need to be updated to perform the next iteration. A so-called merit function evaluating the progress from $\mathbf{x}^k$ to $\mathbf{x}^{k+1}$ serves as a convergence metric and determines when to stop the SQP process.

For a successful execution of SQP, several conditions need to hold. One aspect is to assure that the individual quadratic subproblems have a solution, which requires the constraints to have a non-empty feasible set and the quadratic objective function to be bounded below on this set. Further, both first order necessary conditions and strong second order sufficient conditions guaranteeing local minimums in constrained non-linear optimization have to apply. They also assure that the unknown optimal multipliers $\mathbf{u}$

and $\mathbf{v}$ are unique. Implementation of SQP is also significantly simplified if all of the used functions are three times continuously differentiable. Finally, since the Hessian matrix $\mathbf{H}(\mathcal{L}(\cdot))$ of the Lagrangian function $\mathcal{L}$ frequently is difficult to determine, $\mathbf{H}$ is usually approximated. In turn, the quality of this approximation largely determines the rate at which the sequence of $\mathbf{x}^k$ converges [BT95].

D.5 Support Vector Machines

The concept of SVMs (sometimes also referred to as *Support Vector Network*) has been introduced by Cortes et al. [CV95] as a means to solve binary classification problems. The novelty of the concept is to map input vectors to a higher dimension using a non-linear function. This transformation allows for a linear hyperplane ("decision surface") to be constructed so as to separate the transformed input vectors into two distinct classes with maximum margin between the vectors of the two classes. The vectors are said to be linearly separable. Practically, the algorithm is trained by supplying it with some training data where class membership is already known. The training result is an optimized hyperplane that serves as a linear decision function for classifying testing data.

Specifically, in a first step, an SVM applies a non-linear transformation $\Phi : \mathbb{R}^n \mapsto \mathbb{R}^m$ (with $m \geq n$) on the training input vectors $(\mathbf{x}_i, b_i)$ ($i = 1, \ldots, N$ and b_i denoting the class membership):

$$\mathbf{z}_i = \Phi(\mathbf{x}_i). \tag{D.16}$$

The second step comprises the construction of an optimal hyperplane of form

$$< \mathbf{w}_0, \mathbf{z} > + b_0 = 0, \tag{D.17}$$

so that

$$b_i < \mathbf{w}_0, \mathbf{z}_i > + b_0 \geq 1 \quad \forall\, i = 1, \ldots, N, \tag{D.18}$$

i.e., each training point comes to lie on either of the two sides of the hyperplane. $\mathbf{w}_0$ can be written as a linear combination of a subset $\mathbb{V}$ of the transformed training vectors, the so-called *support vectors*:

$$\mathbf{w}_0 = \sum_{i=1}^{N} b_i \alpha_i \mathbf{z}_i, \tag{D.19}$$

where obviously $\alpha_i = 0 \ \forall \ \mathbf{z}_i \notin \mathbb{V}$ and $b_i < \mathbf{w}_0, \mathbf{z}_i > + b_0 = 1 \ \forall \ \mathbf{z}_i \in \mathbb{V}$. The binary decision function deciding on the class membership of $\mathbf{z}$ can then be written as

$$B(\mathbf{z}) = \text{sign}\left(\sum_{i=1}^{N} \alpha_i < \mathbf{z}_i, \mathbf{z} > + b_0\right), \tag{D.20}$$

where $K(\mathbf{x}_i, \mathbf{x}) := < \Phi(\mathbf{x}_i), \Phi(\mathbf{x}) > = < \mathbf{z}_i, \mathbf{z} >$ is defined as the *Kernel Function* of an SVM denoting the general form of a dot product in the *Hilbert* space [AB62]. Summarizing, the fundamental aspect of SVMs is to find the optimal hyperplane as depicted in Equation (D.17). In case the training data is linearly separable without error, this translates into finding the optimum $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)^T$ by solving the quadratic programming problem

$$\max_{\boldsymbol{\alpha}} \left(\|\boldsymbol{\alpha}\|_1 - \frac{1}{2} \boldsymbol{\alpha}^T \mathbf{D} \boldsymbol{\alpha} \right), \tag{D.21}$$

subject to constraints

$$\boldsymbol{\alpha} \geq \mathbf{0}, \tag{D.22}$$

$$< \boldsymbol{\alpha}, \mathbf{b} > = 0, \tag{D.23}$$

where $\|\cdot\|_1$ is the *Taxicab* norm and $\mathbf{D}$ is a $n \times n$ matrix with elements $d_{ij} = b_i b_j < \mathbf{x}_i, \mathbf{x}_j >$. In case the training data is not linearly separable without error (i.e., at least one of the training data points will lie on the wrong side of the hyperplane), a so-called soft hyperplane that minimizes the number of such points as well as the distance by which they deviate from the "right" class can be constructed. Formally, this usually resolves into a dual convex programming problem [CV95].

The characteristic of high flexibility makes SVMs a frequently used method in such areas as pattern recognition, localization, page ranking evaluation, as well as any kind of object classification. The multitude of available kernel functions K allow for a versatile design of the decision surface [KL03]. A further advantage stems from SVMs' capability of finding well-performing decision functions based on a marginal set of training data. Downsides of the algorithm include the necessity of transforming training data into very high dimensions, the possible construction of computationally intensive quadratic programming problems as well as its incompatibility with data that is not linearly separable, even in higher dimensions. Enhancements to the conventional machine include support vector regression, probabilistic kernel approaches, and combination with neural networks [MM06].

D.6 Fuzzy Logic

Fuzzy inference generally comprises four steps as illustrated in Figure 5.1 and depicted in the following: fuzzification, rule matching, inference, and defuzzification. As input, the process takes n exact (sometimes also referred to as *crisp*) variables $(u_1, u_2, \ldots, u_n)$, the output comprises m exact (crisp) variables $(y_1, \ldots, y_m)$, representing the system variables to be controlled or optimized.

In a first step, fuzzification is used to quantify the certainty μ of an abstract situation $s \in \mathbb{S}$ being true in the presence of exact input variables $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, denoted by membership functions $\mu_s : \mathbb{R}^n \to [0, 1]$, where $\mathbb{S}$ is the set of defined situations, commonly referred to as linguistic variables. In the case of a one-dimensional input $\mathbf{u} = u_1$, typical membership functions include (skewed) triangular or Gaussian functions. According graph shapes apply in the multi-dimensional case, as illustrated in Figures D.2 and D.3. It is important to note that membership functions do not necessarily possess the characteristics of probability density functions.

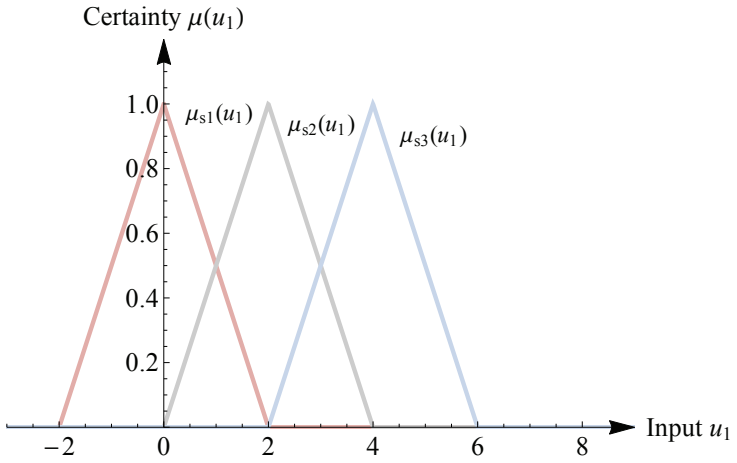


Figure D.2: Triangular fuzzy membership function with one input variable

Secondly, the rule base, containing a set of rules in the format **if** *premise* **then** *consequence*, is considered for identifying active rules, i.e., rules where the premise is met. The premise usually is a Boolean composition of linguistic variables from $\mathbb{S}$, e.g. **if** $s_1 \wedge s_2$ **then** *consequence* c_1. Since fuzzy membership rather than boolean notation is utilized for $s \in \mathbb{S}$, the membership function value $\mu_s(\mathbf{u})$ usually is not either one (s is true) or zero (s is false), but rather $\mu_s(\mathbf{u}) \in]0, 1[$. This implies that a certainty μ should also be introduced for the premise of a rule. Typically, in case of the $\wedge$-composition this resolves into the minimum certainty of all involved

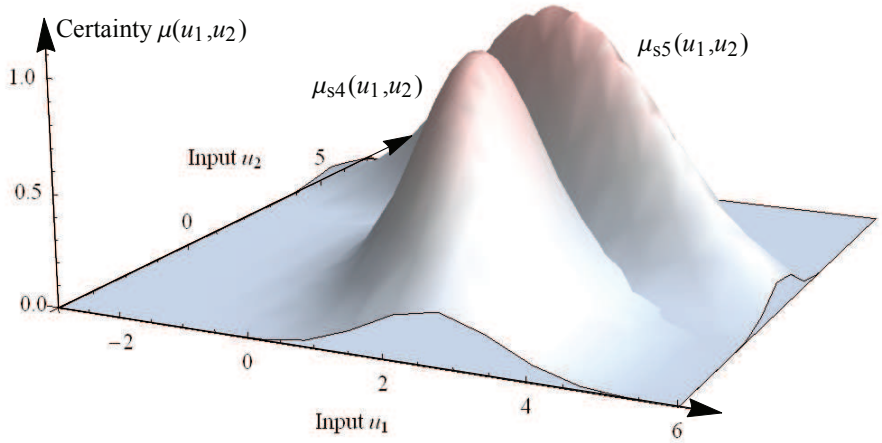


Figure D.3: Gaussian fuzzy membership function with two input variables

linguistic variables, and for the $\vee$-composition into the maximum certainty. For example, if $\mu_{s_1}(\mathbf{u}) = 0.7$ and $\mu_{s_2}(\mathbf{u}) = 0.2$, then $\mu_{s_1 \wedge s_2}(\mathbf{u}) = 0.2$ and $\mu_{s_1 \vee s_2}(\mathbf{u}) = 0.7$. The product sometimes serves as an alternative in case of the $\wedge$-composition, i.e. $\mu'_{s_1 \wedge s_2}(\mathbf{u}) = 0.14$.

In a third step, fuzzy conclusions are drawn based on the set of active rules r_l, $l = 1, \ldots, L$. This inference process mainly comprises the prioritization and combination of the consequences c_l of active rules. Each c_l usually determines how to manipulate a subset of the exact output variables $\mathbf{y}^{(\text{exact})} = (y_1^{(\text{exact})}, \ldots, y_m^{(\text{exact})})$, i.e., it provides a recommended $\mathbf{y}_l$ so that $\mathbf{y}^{(\text{exact})} = \mathbf{y}_l$. Active consequences c_l and their respective recommended outputs $\mathbf{y}_l$ are assigned a weight according to the certainty of their respective premise, yielding tuples $(\mu_{\text{premise}_l}, \mathbf{y}_l)$.

Finally, defuzzification determines exact outputs $y_j^{(\text{exact})}$ from the fuzzy conclusions drawn in the previous step. Different methods are available for this step, among them the *center of gravity*, *center average*, *maximum criterion*, or *maximum average* methods [PY98]. Their applicability not only depends on the context the controller is used in, but also on the cardinality of each y_j. For real-valued y_j, the center average method produce the exact outputs $y_j^{(\text{exact})}$ according to

$$y_j^{(\text{exact})} = \frac{\sum_{l=1}^{L} y_{j,l}\ \mu_{\text{premise}_l}}{\sum_{l=1}^{L} \mu_{\text{premise}_l}}, \tag{D.24}$$

which can eventually be used to set the respective system parameters subject to optimization.

D.7 Kalman Filter

Mathematically, the Kalman filter shall estimate the state $\mathbf{x}_k \in \mathbb{R}^n$ of a discrete-time controlled process at time k (Table D.2 depicts the used notations). State transition $\mathbf{x}_k$ is described by the linear stochastic difference equation

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k. \tag{D.25}$$

The according process observation $\mathbf{z}_k \in \mathbb{R}^m$ is modeled using the observation matrix $\mathbf{H}$:

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k, \tag{D.26}$$

where $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^m$ represent process and measurement noise, respectively. They are assumed to be mutually independent and to follow a time-invariant Normal distribution with zero mean, i.e. $\mathcal{N}_n(\mathbf{0}, \mathbf{\Sigma_v})$ and $\mathcal{N}_m(\mathbf{0}, \mathbf{\Sigma_w})$ with process noise covariance matrix $\mathbf{\Sigma_v} \in \mathbb{R}^{n\times n}$ and measurement noise covariance matrix $\mathbf{\Sigma_w} \in \mathbb{R}^{m\times m}$.

Table D.2: Symbol notation for the Kalman filter

Symbol	Meaning	Symbol	Meaning
k	time step		
$\mathbf{x}_k \in \mathbb{R}^n$	state vector	$\hat{\mathbf{x}}_k^-$	a priori estimate of $\mathbf{x}_k$
$\mathbf{u}_k \in \mathbb{R}^l$	input vector	$\hat{\mathbf{x}}_k$	a posteriori (corrected) estimate of $\mathbf{x}_k$
$\mathbf{A} \in \mathbb{R}^{n\times n}$	state matrix	$\mathbf{e}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k^-$	a priori estimate error
$\mathbf{B} \in \mathbb{R}^{n\times l}$	input matrix	$\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$	a posteriori estimate error
$\mathbf{z}_k \in \mathbb{R}^m$	observation/measurement vector	$\mathbf{p}_k^-$	a priori estimate error covariance
$\mathbf{H} \in \mathbb{R}^{m\times n}$	observation matrix	$\mathbf{p}_k$	a posteriori estimate error covariance
$\mathbf{\Sigma_w} \in \mathbb{R}^{n\times n}$	process noise covariance	$\mathbf{K}$	Kalman gain
$\mathbf{\Sigma_v} \in \mathbb{R}^{m\times m}$	measurement noise covariance	$\mathbf{H}\hat{\mathbf{x}}_k^-$	measurement prediction

Implementing the first of the two basic Kalman filter steps (prediction), the a priori state estimate $\hat{\mathbf{x}}_k^-$ is defined as the expected state at time k, based on the knowledge of the process until time step $k-1$ and *before* taking the measurement $\mathbf{z}_k$, according to Equation (D.25):

$$\hat{\mathbf{x}}_k^- = \mathbf{A}\hat{\mathbf{x}}_{k-1} + \mathbf{B}\mathbf{u}_{k-1}. \tag{D.27}$$

Similarly, in the second step (model update), $\hat{\mathbf{x}}_k$ is defined as the a posteriori estimate of $\mathbf{x}_k$ *after* having measured $\mathbf{z}_k$, i.e. including the additional knowledge gained at time step k. The approach is to approximate $\mathbf{x}_k$ using

a linear combination of the measurement $\mathbf{z}_k$, the a priori estimation $\hat{\mathbf{x}}_k^-$ and the measurement prediction according to Equation (D.26):

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^-). \tag{D.28}$$

The difference $\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^-$ between actual and expected measurement is called *innovation* or *residual*. $\mathbf{K}$ depicts the *Kalman gain* which is to be chosen so as to minimize the a posteriori estimation error

$$\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k. \tag{D.29}$$

The (time variant) covariance of $\mathbf{e}_k$ is depicted as $\mathbf{P}_k$. It can be shown [May82, Bro83, Jac93] that $\mathbf{e}_k$ is minimized by

$$\mathbf{K}_{opt} = \frac{\mathbf{P}_k^- \mathbf{H}^T}{\mathbf{H}\mathbf{P}_k^- \mathbf{H}^T + \mathbf{\Sigma}_\mathbf{v}}. \tag{D.30}$$

$\mathbf{P}_k^-$ depicts the expected estimation error covariance, i.e. before the measurement $\mathbf{z}_k$ has been taken. Intuitively, if the estimation error variance is small (i.e. $\hat{\mathbf{x}}_k^-$ is a reliable approximation of $\mathbf{x}_k$), $\mathbf{K}_{opt}$ should approach zero, putting less weight on the residual. On the other hand, if the measurement noise covariance $\mathbf{\Sigma}_\mathbf{v}$ is small (i.e. the observation $\mathbf{z}_k$ is reliable), more emphasis should be put on the residual. This is reflected in the behavior of Equation (D.30):

$$\lim_{\mathbf{P}_k^- \to \mathbf{0}} \mathbf{K}_{opt} = \mathbf{0}, \tag{D.31}$$

$$\lim_{\mathbf{\Sigma}_\mathbf{v} \to \mathbf{0}} \mathbf{K}_{opt} = \mathbf{H}^{-1}. \tag{D.32}$$

D.8 Particle Filter

In contrast to the Kalman filter, particle filters do not make the assumption of linear, Gaussian-distributed quantities, they rather operate on non-linear state-space and observation models without any pre-defined probability distribution [DKJ+03]:

$$\mathbf{x}_k = g(\mathbf{x}_{k-1}, \mathbf{u}_k) \tag{D.33}$$

and a non-linear observation model

$$\mathbf{z}_k = h(\mathbf{x}_k), \tag{D.34}$$

where $\mathbf{u}$, $\mathbf{x}$, and $\mathbf{z}$ represent system input, system state, and observations at time k, respectively. For the estimation of $\mathbf{x}_k$, the basic idea of particle filters is to recursively approximate the respective probability distributions $p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1})$ (predictive distribution), $p(\mathbf{x}_k \mid \mathbf{z}_{1:k})$ (posterior distribution, $l \geq 1$), and $p(\mathbf{x}_k \mid \mathbf{z}_{1:K})$ (smoothed distribution, $K > k$). Using the *Chapman-Kolmogorov* equation and *Bayes' Rule*, predictive and posterior distributions are computed as follows:

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} \mid \mathbf{z}_{1:k-1}) \, d\mathbf{x}_{k-1}, \tag{D.35}$$

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) = \frac{p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}) p(\mathbf{z}_k \mid \mathbf{x}_k)}{p(\mathbf{z}_k \mid \mathbf{z}_{1:k-1})}. \tag{D.36}$$

The solving of Equations (D.35) and (D.36) usually is computationally intensive, except for the linear Gaussian case which resolves into the Kalman filter. This problem is resolved by approximating the according distributions by drawing discrete samples, so-called particles, according to the Sequential Importance Sampling (SIS) method. Here, M particles $\mathbf{x}_k^{(m)}$ ($m = 1, \ldots, M$) are drawn from the so-called importance distribution $\pi(\mathbf{x})$ and assigned a properly chosen weight $w^{(m)}$. With M particles available, the state estimation $\hat{\mathbf{x}}_k$ is then given by

$$\hat{\mathbf{x}}_k = \sum_{m=1}^{M} w^{(m)} \mathbf{x}_k^{(m)}. \tag{D.37}$$

The posterior distribution is approximated by

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) \approx \sum_{m=1}^{M} w^{(m)} \delta(x_k - x_k^{(m)}), \tag{D.38}$$

where $\delta(\cdot)$ is the Dirac delta function. Using the particles as well as the observations, predictive and smoothed distribution can be approximated in a similar way [PMP+10]. Obviously, the importance distribution π has a major impact on the performance of the filter. Therefore, it should be as close to the posterior distribution as possible. Djuric et al. [DKJ+03] illustrate approaches to appropriately initially select and recursively update both importance distribution as well as particle weights, thus defining the sequential execution of the filter. However, they emphasize that, by employing the described SIS method, the majority of particles degenerates quickly, i.e., their weight becomes very insignificant after only a few se-

quential steps. Therefore, the Sequential Importance Re-sampling (SIR) method has been introduced. The basic extension to SIS is the disposal of particles whose weight falls below a defined threshold. For compensation, particles with heavy weight are split into several particles of reduced weight. The number of particles is usually retained. Further extensions to the conventional filter are the so-called *Gaussian Particle Filter* [KD03a] allowing an improved approximation of the posterior distribution as well as the *Gaussian Sum Particle Filter* [KD03b] featuring approximation based on a linear combination of Gaussian distributions.

D.9 Bayesian Networks

A Bayesian network is an annotated directed acyclic graph $G = (\mathcal{V}, \mathcal{E})$ representing a joint probability distribution over a set of random variables $\mathcal{V}$ forming the vertices of G. The directed edges in $\mathcal{E}$ describe a hierarchy of nodes, indicating the dependency of children nodes on their parent nodes. In order to fully define the network, the conditional probabilities $P(v_i|\pi_i)$ for each realization of the variables $v_i \in \mathcal{V}$ are additionally required. π_i denotes the set of parent variables of v_i.

To derive $P(v_i|\pi_i)$, three important probabilistic rules are exploited in Bayesian networks [Pea85]. First, based on fundamental axioms of probability theory, Bayes has described the quantitative relationship of probabilities $P(v_1)$ and $P(v_2)$ of two events v_1 and v_2 as well as the respective conditional probabilities $P(v_1|v_2)$ and $P(v_2|v_1)$, known as *Bayes' Theorem*:

$$P(v_2|v_1) = P(v_2)\frac{P(v_1|v_2)}{P(v_1)}. \tag{D.39}$$

Second, *Bayes' Rule* can be directly derived from Equation (D.39). Given an interfering event v_I, the a-priori (before v_I happened) and posterior (after v_I) ratio of probabilities of two events v_1 and v_2 relate according to

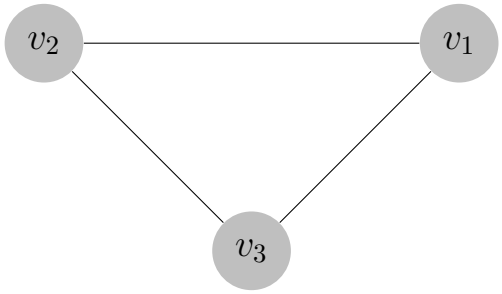
$$\frac{P(v_1|v_I)}{P(v_2|v_I)} = \frac{P(v_I|v_1)}{P(v_I|v_2)} \cdot \frac{P(v_1)}{P(v_2)}. \tag{D.40}$$

Third, in case $P(v_1|v_2)$ is unknown but only the joint probability $P(v_1 \cap v_2)$ is available, Equation (D.39) can be rewritten as

$$P(v_2|v_1) = \frac{P(v_1 \cap v_2)}{P(v_1)}. \tag{D.41}$$

	v_2	
v_1	T	F
F	0.4	0.6
T	0.01	0.99

$P(v_2|v_1)$



v_1	
T	F
0.2	0.8

$P(v_1)$

		v_3	
v_2	v_1	T	F
F	F	0.4	0.6
F	T	0.01	0.99
T	F	0.01	0.99
T	T	0.01	0.99

$P\left(v_3|(v_1, v_2)\right)$

Figure D.4: Bayesian network with three variables and according probability tables, adapted from [Mur98] (T = true, F = false)

Using Eqns. (D.39), (D.40), and (D.41), the joint probability distribution of the variables $v_i \in \mathcal{V}$ can can now be derived [FGG97]:

$$P(v_1, v_2, \ldots, v_n) = \prod_{i=1}^{n} P(v_i|\pi_i). \tag{D.42}$$

Equation (D.42) is particularly exploited for inference in Bayesian networks. Two types of inference can be distinguished, namely *top-down* (predictive) and *bottom-up* (diagnostic). In the former, based on given realization of parent variables, the probability distributions of depending variables is calculated, i.e., the likelihood of a set of possible outcomes is predicted. The latter uses observations of child variables to deduce the probability that parent variables have assumed certain realizations, i.e., to retrospectively evaluate the likelihood of certain causes for observed effects. Since exact inference in Bayesian networks is an NP-hard problem, efficient inference algorithms dealing with specific class of networks have been developed, among them the message passing algorithm, cycle-cutset conditioning, and variable elimination [Pea88], [ZP96]. Approximate inference is performed by, e.g., Monte-Carlo sampling and variational methods [JGJS99]. In prac-

tical situations, graph topology and probability distributions for a Bayesian network are unknown. In such cases, it can learn the according parameters: expert knowledge is used to design an initial network which is fed with training data for subsequent refinement [Pea88].

Figure D.4 depicts an example of a Bayesian network with three variables v_1, v_2, and v_3 and their dependencies. While v_1 is completely independent, both v_2 and v_3 depend on v_1 and v_3 additionally depends on v_2. Accordingly, the tables show the conditional probabilities for the variables. The combination of them yields the joint probability distribution for the three variables.

D.10 Neural Networks

As an extension to Section 5.2.6, Figure D.5 schematically depicts the structure of an artificial neuron The input value β_i (also referred to as *stimulus*) of a neuron is calculated as a weighted sum of the input signals a_j, $\beta_i = \sum_j w_{j,i} a_j$. Depending on the value of β_i, the activation function f_h or f_O decides on the activation of the neuron, i.e., whether the output signal a_i should be activated. Implying a normalized output value, i.e. $a_i \in [0, 1]$, the activation level can either be binary (zero or one, i.e., the activation function is a step function) or assume several degrees of activation between zero and one. a_i then serves as part of the stimulus of all direct successors.

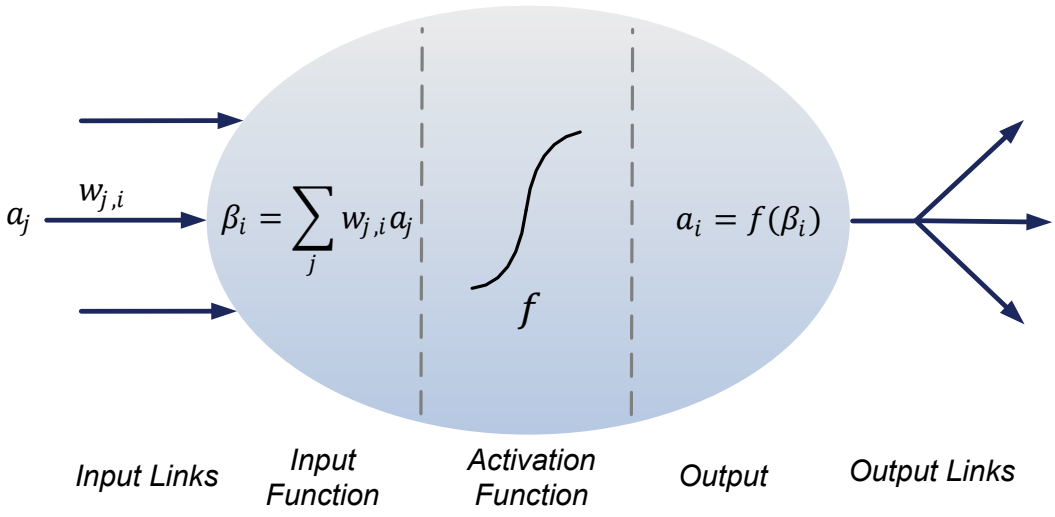


Figure D.5: Design of a neuron in an ANN [Bri06]

In that manner, input stimuli propagate through the network and, depending on the individual activation functions, hidden and output neurons get activated or deactivated. Neural networks are organized in a completely distributed manner, i.e., all computations are performed locally (in the neuron) and no central entity is involved in coordinating network behavior [RN02].

D.11 Genetic Algorithms

Given an initial set of solutions to an optimization problem, genetic algorithms repeatedly execute three operations until a termination condition is met: crossover, mutation, and selection. For a more efficient implementation, a widespread class of genetic algorithms works on binary strings, i.e., a coding function c maps the solutions $\mathbf{x}_i \in \mathbb{R}^m$, $i = 1, \ldots, n$ to binary sequences $\mathbf{s}_i \in \mathbb{B}^k$, with $\mathbb{B} = \{0, 1\}$. In the first step, the crossover method decides on how offspring solutions are derived from parent solutions, i.e., how the genetic material of the parents is mixed to form the new generation. Here, different approaches exist:

- In *one-point crossover*, the binary strings of length n of the parents are split at a randomly chosen position and the two cropped tails are interchanged, thus producing two children.
- *N-point crossover* segments the parent strings into $N + 1$ segments where every second substring is swapped.
- *Segmented crossover* is similar to N-point crossover; however, the number of segments can vary.
- In *uniform crossover*, for each position in the two parent strings, the algorithm randomly decides whether to swap the bits or not.
- For *shuffle crossover*, a permutation is applied to the parent strings before an N-point crossover is performed. The offspring finally undergoes the inverse permutation.

Parents are usually picked based oon their fitness value. The second step comprises the mutation of current solutions, mimicking a deformation of DeoxyriboNucleic Acid (DNA) material due to, e.g., radioactive radiation. Common methods include:

- For *single bit inversion*, the value of a randomly chosen bit of a string is inversed with probability p_M.
- In *bitwise inversion*, each bit of a string is inverted with probability p_M.
- In *random selection*, the entire string is replaced by a randomly chosen one with probability p_M.

A common method for solution selection consists of proportional selection which exploits the fitness function $f(\mathbf{x})$. Here, the selection probability of a solution s_j is proportional to its fitness function value [Bod01]:

$$P(s_j \text{ is selected}) = \frac{\varphi(f(s_j))}{\sum_{i=1}^{n} \varphi(f(s_i))}, \tag{D.43}$$

where a nondecreasing function φ: $\mathbb{R} \to \mathbb{R}^+$ is applied in case not all fitness values are positive. Other selection methods are less restrictive in earlier iterations of the GA to allow for a more inclusive search in the solution space and become more selective in subsequent iterations. Among them, roulette wheel selection, $(\lambda + \sigma)$-selection, truncation selection, and tournament selection belong to the more widely used ones [GCL08]. Moreover, one can also take into account additional considerations, e.g., maintaining a certain level of genetic diversity. Algorithm 3 summarizes the fundamental operation steps of a generic genetic algorithm.

Algorithm 3: Generic genetic algorithm (adapted from [GCL08])

Data: Initial solution set $\mathbb{S}(0) \subseteq \mathbb{R}^m$, coding function $c : \mathbb{R}^m \to \mathbb{B}^k = \{0,1\}^k$, fitness function f
Result: Optimum solution set $\mathbb{S}_{opt} \subseteq \mathbb{R}^m$, $|\mathbb{S}_{opt}| \geq 1$

begin
    t:=0;
    $\mathbb{G}(0) = c(\mathbb{S}(0))$;
    while *termination condition not met* **do**
        create $\mathbb{O}_C(t)$ from $\mathbb{G}(t)$ by crossover routine;
        create $\mathbb{O}_M(t)$ from $\mathbb{G}(t)$ by mutation routine;
        create $\mathbb{G}(t+1)$ from $\mathbb{G}(t)$, $\mathbb{O}_C(t)$and $\mathbb{O}_M(t)$ by selection routine;
        t:=t+1;
    end
    $\mathbb{S}(t) = c^{-1}(\mathbb{G}(t))$;
    $\mathbb{S}_{opt} = \{ s \in \mathbb{S}(t) \mid f(s) = \max_{\mathbf{s}' \in \mathbb{S}(t)} f(s')$;
end

D.12 Simulated Annealing

The simulated annealing algorithm is initiated by a solution $\mathbf{s}_0$ from the set $\mathbb{S}$ of all possible solutions. The objective function $f : \mathbb{S} \rightarrow \mathbb{R}$ to be minimized is used to compare different solutions with each other and to identify the optimum solution $\mathbf{s}_{opt}$ with $f(\mathbf{s}_{opt}) \leq f(\mathbf{s}) \; \forall \; \mathbf{s} \in \mathbb{S}$. Moreover, the neighborhood set $\mathbb{S}_N(\mathbf{s})$ of a solution $\mathbf{s}$ is defined as the set of those solutions that can be reached in a single iteration of the local search algorithm. The design of an appropriate neighborhood function can be a challenging task and very much depends on the problem under consideration. For example, in a packet scheduling problem (i.e., the solution depicts the scheduling sequence of packets), a neighborhood set of a given solution might consist of all those solutions where a certain packet switches its position with the previous or subsequent packet. In many cases, the design of the neighborhood function has proven to have the strongest impact on the overall performance of the algorithm, e.g., [ABPW08]. Starting with $\mathbf{s}_0$, a candidate $\mathbf{s}'$ for the next solution is selected from $\mathbb{S}_N(\mathbf{s}_0)$, either randomly or using a specified rule (e.g., maximum value of f). The decision of whether to accept or reject $\mathbf{s}'$ is taken based on the *Metropolis* acceptance criterion, which is also used for modeling energy state transition of thermodynamic systems [HJJ03]:

$$p(\mathbf{s}' \text{ is accepted as next solution}) = \begin{cases} e^{-\frac{f(\mathbf{s}')-f(\mathbf{s})}{t_k}} & \text{if } f(\mathbf{s}') - f(\mathbf{s}) \geq 0, \\ 1 & \text{if } f(\mathbf{s}') - f(\mathbf{s}) < 0. \end{cases} \tag{D.44}$$

Here, t_k denotes system temperature at time step k. Equation (D.44) incorporates the hill climbing characteristic as the new solution $\mathbf{s}'$ can be accepted even though it results in an increased value for f. However, with decreasing t_k, the probability for this to happen also decreases. In order to stabilize the solution at each temperature t_k, M_k iterations for finding new solutions are allowed before the temperature is cooled down further according to a defined cooling schedule. The steps of finding new solutions repeat until a defined termination criterion is met, e.g., temperature or objective function value of the current solution falling below a defined threshold, a given number of iterations having been executed, or a defined convergence criterion being met. Algorithm 4 summarizes the fundamental steps of the simulated annealing method.

Algorithm 4: Simulated annealing algorithm (adapted from [Egl90])

Data: Initial solution $\mathbf{s} \in \mathbb{S}$, initial temperature $t_0 \geq 0$, cooling schedule for t_k, number M_k of iterations at each temperature t_k, objective function f
Result: Optimum solution $\mathbf{s}_{opt}$
begin
  $k := 0$;
  repeat
    $m := 0$;
    repeat
      generate a solution $\mathbf{s}' \in \mathbb{S}_N(\mathbf{s})$;
      $\Delta := f(\mathbf{s}') - f(\mathbf{s})$;
      if $\Delta < 0$ **then**
        $\mathbf{s} := \mathbf{s}'$;
      else
        $\mathbf{s} := \mathbf{s}'$ with probability $e^{-\frac{\Delta}{t_k}}$;
      end
      $m := m + 1$;
    until $m = M_k$;
    $k := k + 1$;
    determine t_k according to cooling schedule;
  until *termination criterion is met*;
  $\mathbf{s}_{opt} = \mathbf{s}$;
end

Curriculum Vitae

Personal Details

Name	Christian Mannweiler
Place of Birth	Idar-Oberstein

Professional Experience

03/2014 - present	German Research Center for Artificial Intelligence (DFKI GmbH), Kaiserslautern Intelligent Networks research group Senior researcher
09/2008 - 02/2014	University of Kaiserslautern Wireless Communications and Navigation research group Research associate & teaching assistant

Education

09/2008 - 12/2014	University of Kaiserslautern PhD thesis "Context-Enabled Optimization of Energy-Autarkic Networks for Carrier-Grade Wireless Backhauling"
10/2002 - 05/2008	University of Kaiserslautern Diploma degree in Industrial Engineering and Management (*Wirtschaftsingenieurwesen*)