

Planung interferenzbeschränkter zellulärer  
Funknetze mittels Set-Covering und -Packing

von

Normen Rochau



Planung interferenzbeschränkter zellulärer Funknetze mittels  
*Set-Covering* und *-Packing*

Diplomarbeit  
bei Prof. Dr. Martin Grötschel

vorgelegt von Normen Rochau  
am Fachbereich Mathematik der  
Technischen Universität Berlin

Oktober 2004



Die selbständige und eigenhändige Anfertigung versichere ich an Eides  
statt.

Berlin, den 14. Oktober 2004



# Danksagung

Ich bedanke mich bei Andreas Eisenblätter.



# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>1</b>
<b>2</b>	<b>Grundbegriffe der Funknetzplanung</b>	<b>5</b>
2.1	Ein Funknetz . . . . .	5
2.2	Allgemeiner Planungsansatz . . . . .	10
<b>3</b>	<b>Mathematische Grundlagen</b>	<b>15</b>
3.1	Set Covering, Set Packing, Set Partitioning . . . . .	15
3.2	0/1-Programmierung . . . . .	20
3.3	Ein Beispiel: Das MAX-CUT Problem . . . . .	23
<b>4</b>	<b>Planung interferenzarmer GSM-Netze</b>	<b>27</b>
4.1	Optimierungsansatz . . . . .	29
4.1.1	Kapazitätsplanung . . . . .	29
4.1.2	Kopplungsmatrizen und Frequenzzuweisung . . . . .	31
4.1.3	Thesen zur Planung interferenzarmer GSM-Netze . . . . .	34
4.2	Das pixelbasierte Modell . . . . .	36
4.2.1	Erste Formulierung . . . . .	36
4.2.2	Verbesserung der Formulierung . . . . .	39
4.2.3	Linearisierung . . . . .	42
4.2.4	Schnittungleichungen . . . . .	43
4.3	Das Obermengen-Modell . . . . .	44
4.3.1	Linearisierung . . . . .	45
4.3.2	Schnittungleichungen . . . . .	49
4.4	Vereinfachungen des IP-Modells . . . . .	59
4.4.1	Reduktion der Zahl der Obermengen . . . . .	60
4.4.2	Die Rolle des $\delta$ -Schwellwerts . . . . .	62
4.5	Heuristiken . . . . .	65
4.5.1	Tiltheuristik . . . . .	65
4.5.2	Lokale Suche . . . . .	66
4.6	Ergebnisse . . . . .	72
4.7	Zusammenfassung . . . . .	77

<b>5</b>	<b>Planung interferenzarmer UMTS-Netze</b>	<b>79</b>
5.1	Optimierungsansatz . . . . .	80
5.1.1	Eigenschaften der UMTS-Technologie . . . . .	80
5.1.2	Planungsgrößen . . . . .	82
5.1.3	Lastkopplungsmatrizen . . . . .	83
5.1.4	Thesen zur Planung interferenzarmer UMTS-Netze . .	87
5.2	Formulierung als ganzzahliges Programm . . . . .	89
5.3	Ergebnisse . . . . .	92
5.4	Zusammenfassung . . . . .	95
<b>6</b>	<b>Zusammenfassung</b>	<b>99</b>

# Kapitel 1

## Einführung

In dieser Diplomarbeit beschäftigen wir uns mit der Planung der Funkchnittstelle zellulärer Funknetze. Die Herausforderung besteht darin, für die Antennen Standorte und Ausrichtungen so zu bestimmen, dass man mit einer minimalen Zahl von Antennen eine maximale Ausbeute an Netzabdeckung, Kapazität und Signalqualität erreicht. Die Folgen schlechter Planung sind für den Kunden mangelhafte Abdeckung und schlechte Signalqualität, und für den Netzbetreiber erhöhte Kosten, weil mehr Antennen auf mehr Standorten betrieben werden, als bei guter Planung notwendig wäre.

Wir stellen in dieser Arbeit ein technologieübergreifendes Verfahren zur automatischen Planung der Funkchnittstelle vor. Mit diesem Verfahren ermitteln wir auf Basis von Verkehrsdaten und Prädiktionen über die Signalausbreitung Netzkonfigurationen, die bei garantierter Netzabdeckung und geringem Ressourcenverbrauch eine hohe Servicequalität aufweisen. Das Verfahren basiert auf einem allgemeinen mathematischen Modell. Das Lösungsverfahren ist unabhängig von der betrachteten Technologie. In dieser Arbeit führen wir die Optimierung exemplarisch an der GSM- und an der UMTS-Technologie vor.

Das GSM-System steht durch die Einführung der GPRS- und EDGE-Technologien vor neuen Herausforderungen. Die Systeme werden wegen der höheren Datenraten interferenzanfälliger, sodass neue Planungsverfahren gefragt sind. UMTS ist eine neue Technologie. Da etablierte Planungsmethoden nur schwierig auf UMTS übertragbar sind, sind neue Verfahren erforderlich. Wegen der hohen Komplexität der neuen Technologien gehört die Planung von Hand der Vergangenheit an.

Unsere Planungsidee setzt bei der Behandlung der Interferenz an. Interferenz führt in Funknetzen zu schlechter Signalqualität und Kapazitätseinbußen. Die Bewertung der Interferenzverhältnisse ist in der Planungsphase sowohl in GSM als auch in UMTS ein äußerst komplexer Vorgang. Daher ist man schlecht beraten, solche Bewertungen während eines Optimierungsverfahrens durchzuführen. Wir stellen in dieser Arbeit Methoden vor, wie

man für die Netzkonfigurationen, die während der Optimierung untersucht werden, Voraussagen über die entstehende Interferenz machen kann. Dieses Interferenzpotenzial wird in sogenannten *Kopplungsmatrizen* ausgedrückt. Die darin enthaltenen *Kopplungswerte* sind vereinfacht gesagt ein Maß für die Überlappung der Sendebereiche der Antennen. Dieses Maß ist komplex genug, um eine differenzierte Abschätzung der Interferenzverhältnisse zu erlauben, und einfach genug, um während eines Optimierungsalgorithmus wiederholt ermittelt zu werden. Die Minimierung des Interferenzpotenzials führt zur Ermittlung von Netzkonfigurationen hoher Qualität.

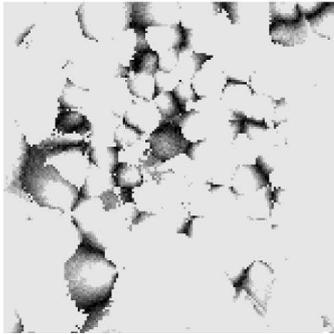
Die erste Formulierung hat die Form eines nichtlinearen 0/1-Programms. Für die Optimierung werden in dieser Arbeit verschiedene Lösungsverfahren vorgestellt. Wir lösen dieses Programm zum einen, indem wir es in ein ganzzahliges Programm (IP) transformieren. Wir stellen hier zwei Arten der Transformation vor. Außerdem entwickeln wir heuristische Methoden. Neben einer Startheuristik, die mit hoher Geschwindigkeit gute Lösungen produziert, stellen wir einen lokalen Suchalgorithmus vor, der diese Lösungen zusätzlich verbessert.

Unser Ansatz ist nicht nur für die Erschließung neuer Gebiete oder Aufstellung neuer Antennen interessant. Er eignet sich auch zur Optimierung bestehender Netze, bei deren Betrieb die oben angesprochenen Probleme auftreten. Wir konzentrieren uns in dieser Arbeit auf die Optimierung des elektrischen Tilts der Antennen. Änderungen des elektrischen Tilts können von Netzbetreibern per Fernsteuerung durchgeführt werden. Mit unseren Methoden haben Netzbetreiber also die Möglichkeit, ihre Netzkonfiguration kostengünstig und in beliebiger Häufigkeit neuen Gegebenheiten anzupassen.

Unsere Ergebnisse belegen, dass unser Planungsansatz zu deutlichen Verbesserungen des Interferenzverhaltens führt. Als Beispiel für die Leistungsfähigkeit unseres Verfahrens betrachte man Abbildung 1.1. Wir sehen links die Ausgangskonfiguration und rechts das optimierte Netz. Die dunkleren Gebiete geben für ein GSM-Netz vereinfacht gesagt die Orte wieder, wo das Empfangssignal keine gute Qualität hat. Je nach Empfänger sind die eingehenden Signale verzerrt oder nicht entschlüsselbar. Wir sehen, dass diese Gebiete durch die Optimierung deutlich verkleinert werden konnten.

## Aufbau der Arbeit

In Kapitel 2 geben wir eine Einführung in die Grundbegriffe der Funknetzplanung. Die Antennen eines zellulären Funknetzes sind von drei Bereichen umgeben, ihrem *best server*-Bereich, der Zelle und dem Störungsbereich. Wir erklären diese Begriffe und erläutern an ihnen die drei Hauptaspekte der Funknetzplanung: Abdeckung, Interferenz und Kapazität. Des Weiteren beschreiben wir, welche negativen Auswirkungen Interferenz auf die Netzabdeckung haben kann und warum Netze mit wenig Interferenz eine höhere



(a)



(b)

---

IST-2000-28088 Momentum, <http://momentum.zib.de>

Abbildung 1.1: Vergleich der Interferenz von Ausgangsnetz (a) und optimiertem Netz (b).

Kapazität haben. Unser Planungsansatz ist an diesen drei Aspekten ausgerichtet: Wir minimieren bei garantierter Abdeckung das Interferenzpotenzial; die nachfolgende Kapazitätsanalyse zeigt, dass durch die Minimierung des Interferenzpotenzials die Gesamtinterferenz reduziert wurde und damit die Kapazität gesteigert.

In Kapitel 3 geben wir kurze Einführungen in die hier berührten Gebiete der Mathematik. Grundkenntnisse der ganzzahligen Optimierung werden vorausgesetzt, ebenso wie Grundkenntnisse der Komplexitätstheorie. Zunächst geben wir eine Einführung in die Optimierungsprobleme *set covering*, *set packing* und *set partitioning*. Nachdem wir nichtlineare 0/1-Programme vorgestellt haben, zeigen wir, wie man sie linearisiert. Für führen diesen Vorgang am Beispiel des MAX-CUT Problems durch.

In Kapitel 4 stellen wir unseren Ansatz zur Planung von GSM-Netzen vor. Wir stellen hier zunächst das Frequenzzuweisungsproblem vor. Erst nach erfolgter Frequenzzuweisung kann man eine Aussage machen über die zu erwartenden Interferenzwerte im Netz. Die Einführung von Kopplungsmatrizen macht die wiederholte Durchführung der Frequenzzuweisung überflüssig. Ermittelt man die Summe ihrer Elemente, so kann man schon *vor* der Frequenzzuweisung Aussagen darüber treffen, welche Interferenzwerte zu erwarten sind. Diese wertvolle Eigenschaft verwenden wir in unseren Optimierungsalgorithmen. Unser Ziel lautet, eine Netzkonfiguration zu ermitteln, für die die Summe der Einträge der Kopplungsmatrix minimal ist.

Zunächst geben wir eine nichtlineare 0/1-Formulierung für dieses Problem. Wir entwickeln zwei Methoden, das Problem in ganzzahlige Programme zu transformieren, eine pixelbasierte und eine obermengenbasierte. Das obermengenbasierte Modell ist kompakter und lässt sich für kleine Instan-

zen lösen. Allerdings treten Laufzeitprobleme bei größeren Instanzen auf. Wir entwickeln verschiedene Klassen von Schnittungleichungen, um die Problemformulierung zu verbessern, erreichen aber nur wenig. Deutliche Fortschritte erzielen wir durch die Entwicklung heuristischer Verfahren. Das von uns vorgestellte lokale Suchverfahren findet auf allen untersuchten Instanzen Lösungen, die zu deutlichen Verbesserungen der Interferenzwerte führen. Wir stellen unsere Ergebnisse am Ende von Kapitel 4 vor.

In Kapitel 5 befassen wir uns mit der Planung von UMTS-Netzen. Hier werden die Kopplungsmatrizen so definiert, dass Verkehrswerte mit eingegeben. Mit den Kopplungsmatrizen kann man hier Leistungswerte für Antennen bestimmen, anhand derer man die Netzauslastung beurteilt, die sich für den gegebenen Verkehr ergibt. Die Systemauslastung erlaubt Rückschlüsse über die Qualität der gefundenen Konfiguration. Ist die Auslastung trotz geringen Verkehrsaufkommens hoch, so ist das ein Zeichen für eine schlechte Konfiguration: Die Antennen sind hier nämlich gezwungen, einen großen Teil ihrer Sendeleistungen dafür zu verwenden, die Störungen, die durch die Signale anderer Sender verursacht werden, zu neutralisieren. Die mathematischen Modelle der Kapitel 4 und 5 sind dieselben. Daher entwickeln wir in Kapitel 5 keine neuen Methoden zur Verbesserung der IP-Formulierung. Wir konzentrieren uns vielmehr auf die Ergebnisse, die mit der lokalen Suche ermittelt wurden. Sie belegen, dass unser Planungsansatz auch in UMTS zu verbesserten Netzkonfigurationen führt.

Das letzte Kapitel ist der Zusammenfassung der Ergebnisse dieser Arbeit gewidmet.

Grundlage unserer Berechnungen waren Daten, die vom MOMENTUM-Projekt (siehe [17]) zusammengestellt wurden. Die verwendeten Verkehrsdaten sowie die Pfadverlust-Prädiktionen wurden mit MOMENTUM-Daten erstellt. Bei der Erstellung einiger Abbildung verwendeten wir MOMENTUM-Visualisierungswerkzeuge. Diese Abbildungen sind gesondert gekennzeichnet.

## Kapitel 2

# Grundbegriffe der Funknetzplanung

Dieses Kapitel besteht aus zwei Teilen. Im ersten geben wir eine Einführung in die von uns verwendeten Mittel zur Planung zellulärer Funknetze, im zweiten erläutern wir unseren Planungsansatz in allgemeiner Form.

In Abschnitt 2.1 erklären wir zunächst den Aufbau eines Funknetzes und was mit den Begriffen *best server*-Bereich, *Zelle* und *Störungsbereich* gemeint ist. Danach beschreiben wir, wie wir diese Bereiche mit Pfadverlust-Prädiktionen bestimmen. Wir erklären hier auch die drei Planungsaspekte Abdeckung, Kapazität und Interferenz, und wie sie zusammenhängen.

In Abschnitt 2.2 erläutern wir den in dieser Arbeit verfolgten Planungsansatz in allgemeiner, technologieunabhängiger Form. Unsere Hauptidee besteht darin, eine Methode zur Prognose der Interferenzverhältnisse einer Netzkonfiguration zu entwickeln. Die eigentliche Ermittlung der Interferenz ist in der Regel kompliziert und ist daher in einem Optimierungsverfahren schlecht verwendbar. Sie erfolgt bei uns in nachgeordneten, komplexen Berechnungen. Das Interferenzpotenzial einer Netzkonfiguration wird durch *Kopplungsmatrizen* bestimmt und durch unsere Optimierungsverfahren minimiert. Die Definition der Kopplungsmatrizen ist technologieabhängig. Ihr Vorteil besteht darin, dass die in ihnen codierte Kopplung deutlich leichter ermittelt werden kann als die tatsächliche Interferenz.

Wir übertragen unseren allgemeinen Planungsansatz in dieser Arbeit auf die Technologien GSM und UMTS. Die Umsetzungen unterscheiden sich deutlich, und werden daher getrennt in den Kapiteln 4 und 5 behandelt.

### 2.1 Ein Funknetz

Bei der Planung der Funkschnittstelle von Funknetzen befassen wir uns mit der Aufgabe, Standorte für Antennen zu bestimmen und ihre Ausrichtung festzulegen. Eine Antenne bedient dabei die Kunden (*mobile stations*) in

ihrer unmittelbaren Umgebung. Die Menge der Orte, an denen eine Antenne von allen Antennen des Netzes als stärkste empfangen werden kann, heißt *best server-Bereich*. Die Antenne heißt an solchen Orten *best server*. Wir nehmen in unseren Modellen an, dass eine *mobile station* immer vom *best server* bedient wird. In der Realität trifft dies in den meisten Fällen zu.

Ein Empfänger benötigt zur Entschlüsselung der eingehenden Signale eine gewisse Signalstärke. Der Bereich, in dem das Signal einer Antenne so gut empfangen werden kann, dass eine Entschlüsselung möglich ist, heißt *Zelle*. Die Zellen enthalten die *best server*-Bereiche und sind mitunter deutlich größer als diese. Wir verwenden die Begriffe Zelle und Antenne in dieser Arbeit weitgehend synonym, wobei aus dem Kontext hervorgeht, ob der Sender selbst oder die von ihm bediente Fläche gemeint ist.

Ein dritter Bereich, der durch Standort und Ausrichtung einer Antenne definiert wird, ist der *Störungsbereich*. In diesem Bereich sind die Signale der Antenne noch stark genug, um andere Signale zu stören, aber nicht mehr stark genug, als dass eine Kommunikation möglich wäre. In Abbildung 2.1 sind die drei Bereiche einer Antenne skizziert.

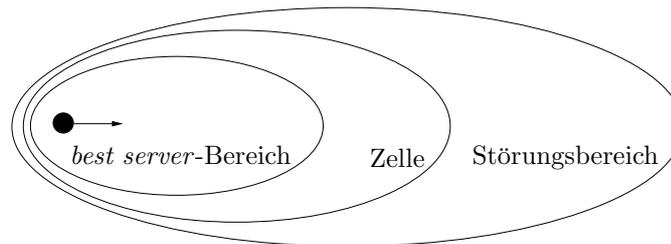
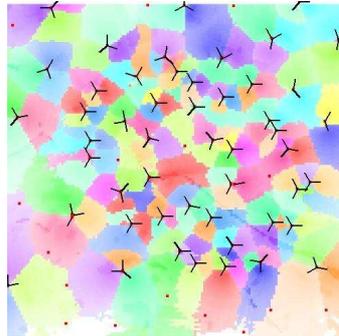


Abbildung 2.1: *Best-server*-Bereich, Zelle und Störungsbereich einer Antenne.

Ein Planungsgebiet wird durch die *best server*-Bereiche der Antennen in Teilgebiete aufgeteilt. Eine Darstellung dieser Aufteilung ist in Abbildung 2.2 zu sehen. Wir nennen eine solche Abbildung eine *best server-Karte*. In dieser Abbildung befinden sich in den meisten Fällen drei Antennen an einem Standort. Ihre horizontale Ausrichtung ist zumeist um einhundertzwanzig Grad verschoben. Jede Antenne ist in der Abbildung von ihrem *best server*-Bereich umgeben.

Zur Ermittlung einer solchen Karte benötigt man *Pfadverlust-Prädiktionen*. Sie basieren auf Wellenausbreitungsmodellen (siehe [20]). Wellenausbreitungsmodelle beschreiben, wie ein Signal bei der Ausbreitung gedämpft wird. Ein Funksignal verliert grundsätzlich mit zunehmender Entfernung an Stärke. Bei der Wellenausbreitung spielen aber auch andere Faktoren eine Rolle, wie die Reflexion von Wellen an Gebäuden oder die Signaldämpfung durch Vegetation. Pfadverlust-Prädiktionen geben an, wie gut das Signal einer Antenne in den Punkten im Planungsgebiet empfangen werden kann.



IST-2000-28088 Momentum, <http://momentum.zib.de>

Abbildung 2.2: Eine *best server*-Karte mit gefärbten Zellen.

Das Verhältnis von abgestrahlter Leistung  $S$  und Empfangsleistung  $E$  wird *Pfadverlust* genannt. Die Pfadverlust-Prädiktionen benutzt man, um die Abdeckungseigenschaften eines Netzes – und damit seine Qualität – schon in der Planungsphase zu beurteilen. Der Pfadverlust wird häufig auf einer logarithmischen Skala in Dezibel (dB) gemessen. Das Verhältnis der Zahlen  $S$  und  $E \neq 0$  wird in der Einheit

$$\frac{S}{E} = 10 \cdot \log_{10}\left(\frac{S}{E}\right)\text{dB}$$

ausgedrückt. Je größer das  $S/E$ -Verhältnis ist, desto größer ist der Pfadverlust. Ein typischer Pfadverlust eines Funksignals auf dem Weg vom Sender zum Empfänger sind 80 dB. Das bedeutet, dass die Empfangsleistung um acht Zehner-Potenzen geringer ist als die Sendeleistung. Die absoluten Sende- und Empfangsleistungen werden wegen der große Unterschiede ebenfalls auf einer logarithmischen Skala angegeben. Die Einheit dieser Skala ist dBm und wird auf 1 mW bezogen:

$$\frac{a}{b} \text{ mW} = 10 \cdot \log_{10}\left(\frac{a}{b}\right)\text{dBm}$$

Eine Sendeleistung von 1000 mW entspricht also 30 dBm. Wird sie einem Pfadverlust von 80 dB unterworfen, so ist die Stärke des Empfangssignals  $-50$  dBm, was 0,00001 mW entspricht. In [19] sind  $-78$  dBm  $\approx 1,5 \cdot 10^{-8}$  mW als typischer Wert angegeben für die Empfangsleistung einer *mobile station* in GSM.

Die Pfadverlust-Prädiktion einer Antenne ist in Abbildung 2.3 dargestellt. Die blauen Bereiche gehören zur Zelle, die roten zum Störungsbereich.

Ein Aspekt der Funknetzplanung ist die *Netzabdeckung*. Ein Punkt gilt als *abgedeckt*, wenn die Signale, die ihn erreichen, eine gewisse Mindestempfangsstärke überschreiten, wenn mit anderen Worten der Punkt in mindestens einer Zelle enthalten ist. In dem Planungsansatz, den wir in dieser



---

IST-2000-28088 Momentum, <http://momentum.zib.de>

Abbildung 2.3: Abdeckungsgebiet einer Antenne.

Arbeit vorstellen, fordern wir, dass bestimmte Punkte des Planungsgebiets abgedeckt sind. Dadurch wird die Menge der möglichen Netzkonfigurationen eingeschränkt. Ein solcher Ansatz heißt *set covering*-Ansatz. Alternative Planungsansätze sehen beispielsweise vor, dass nur ein bestimmter Prozentsatz der Punkte überdeckt ist. Solche weichen Bedingungen machen Sinn, wenn die Netzkonfiguration dafür andere Vorteile aufweist.

Abbildung 2.4 zeigt schematisch, wie das Planungsgebiet mit Zellen überdeckt wird. Wie man sieht, überschneiden sich viele Zellen. Die meisten

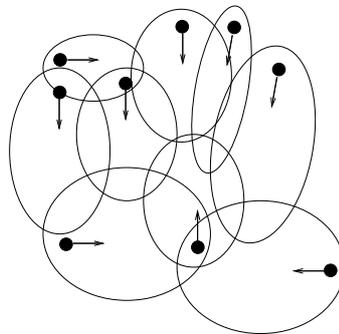
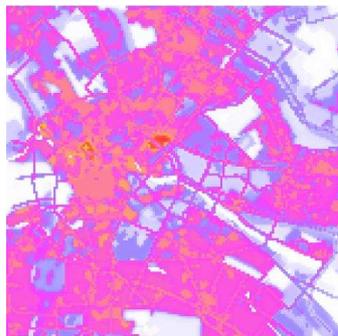


Abbildung 2.4: Überlappende Zellen.

Punkte befinden sich in den Zellen und Störungsbereichen einer Vielzahl von Antennen. Hier können Probleme entstehen, denn ein Empfänger erhält hier Signale von mehreren Antennen. Daher sollten bei der Abdeckungsplanung komplexere Kriterien verwendet werden, als beispielsweise die Minimierung der Anzahl der Antennen. In [14] wurde versucht, durch die *Form* der Zellen Einfluss auf die Überlappungen zu nehmen, in [13] wurden Interferenz- und Verkehrsaspekte mit berücksichtigt.

Empfänger müssen also in der Lage sein, aus der Menge der eingehenden Signale das für sie bestimmte Signal herauszufiltern. Es wurden viele Me-



---

IST-2000-28088 Momentum, <http://momentum.zib.de>

Abbildung 2.5: Eine Verkehrskarte.

thoden entwickelt, die Funksignale so zu übermitteln, dass sie beim Empfang getrennt werden können. In verschiedenen Funktechnologien werden unterschiedliche Methoden verwendet. Daher unterscheiden sich die Planungsaufgaben zum Teil deutlich. Alle Methoden haben gemeinsam, dass sie an Grenzen stoßen, wenn sich zu viele Signale überlagern, wenn sie also zuviel *Interferenz* ausgesetzt sind. Daher ist die im Netz zu erwartende Interferenz ein weiterer Planungsaspekt. Ob und wie stark sich die Signale stören, hängt von den verschiedenen Empfangsstärken ab. Die Überlappung zweier Zellen oder einer Zelle und eines Störungsbereiches bedeutet nicht automatisch, dass die Signale der beiden Antennen sich in jedem Punkt des Überlappungsbereichs stören. Die Interferenzplanung bei garantierter Abdeckung ist das Hauptanliegen dieser Arbeit.

Ein dritter Aspekt, der bei der Funknetzplanung eine Rolle spielt, ist die *Kapazität*. Zu jedem Zeitpunkt befindet sich in einer Zelle eine bestimmte Anzahl von Kunden. Die Antenne sollte in der Lage sein, diese Kunden zu bedienen, das heißt, sie sollte eine ausreichende Kapazität aufweisen. Eine alleinige Planung der Abdeckung und Interferenz macht wenig Sinn, wenn mit so wenigen Antennen geplant wurde, dass das resultierende Netz nicht in der Lage ist, den Bedarf zu decken. Negative Effekte einer mangelhaften Kapazitätsplanung können sein, dass Kunden nicht die Datenraten zur Verfügung gestellt bekommen, die sie anfordern, dass sie keine Verbindung zum Netz herstellen können, oder dass bestehende Verbindungen unterbrochen werden.

Wir benutzen für die Planung Verkehrsprädiktionen. Sie messen jedem Punkt im Planungsgebiet einen Verkehrswert bei. In ländlichen Gebieten gibt es weniger Verkehr als in Ballungsräumen. Daher sind *best server*-Bereiche dort für gewöhnlich bedeutend größer als in der Stadt. Ein Beispiel für eine Verkehrskarte ist in Abbildung 2.5 zu sehen. In ihr ist das Verkehrsaufkommen farblich dargestellt. Je dunkler ein Bereich gefärbt ist, desto größer ist das zu erwartende Verkehrsaufkommen.

Der Aspekt der Interferenz ist sowohl mit der Abdeckung als auch mit der Kapazität verbunden. In der UMTS-Technologie beispielsweise senden Antennen ein Pilotsignal aus. Mit seiner Hilfe nehmen *mobile stations* Kontakt zum Netz auf. Ist das Netz schlecht konfiguriert, so stören sich die Pilotsignale so stark, dass die *mobile stations* die Pilotsignale nicht mehr trennen können. Gebiete, wo solch starke Störungen vorliegen, sind faktisch nicht abgedeckt, obwohl die Signalstärke für eine Kommunikation ausreichend wäre.

Eine ähnliche Überlegung können wir für die Kapazität eines Netzes durchführen. Ist die Interferenz so groß, dass Datenraten gesenkt und Kunden abgewiesen werden müssen, so ist die Kapazität geringer als die eines Netzes, das mit der gleichen Anzahl von Antennen aufgrund einer besseren Konfiguration weniger Interferenz aufweist.

## 2.2 Allgemeiner Planungsansatz

Möchte man die Planungsaufgabe, die wir in Abschnitt 2.1 beschrieben haben in einem Satz zusammenfassen, so kann dieser folgendermaßen lauten: *Ermittle Standorte für die Basisstationen und Ausrichtungen für die Antennen, sodass die Netzabdeckung unseren Bedingungen entspricht, die Kapazität dem zu erwartenden Verkehr angemessen ist und die auftretende Interferenz minimal ist.*

In dieser Arbeit entwickeln wir einen Planungsansatz zur Lösung dieser Aufgabe. Er sieht vor, die Interferenz zu minimieren, wobei die Abdeckungsbedingungen die Zahl der möglichen Netzkonfigurationen einschränken. Die Planungsaufgabe verlangt in unterschiedlichen Technologien unterschiedliche Herangehensweisen. Wir führen die Netzplanung exemplarisch an den Technologien GSM und UMTS durch.

Wir beschränken uns in dieser Arbeit auf die Optimierung der Konfiguration bestehender Netze. Das mathematische Modell, das wir hier entwickeln, ist jedoch auch auf Szenarien anwendbar, in denen neue Antennenstandorte ermittelt werden sollen, oder wo Gebiete neu erschlossen werden sollen. Die Menge von Antennenstandorten ist in unseren Szenarien festgelegt, ebenso wie die Anzahl der Antennen pro Standort. Uns interessiert, welche Verbesserungen durch Veränderung der Antennenausrichtungen erreicht werden können.

Die Antennen eines Funknetzes haben eine horizontale Ausrichtung, genannt *Azimut*, und eine vertikale Ausrichtung, genannt *Tilt*. Die Ausrichtung der Antenne bestimmt die Hauptstrahlrichtung – die Richtung, in der ihr Signal am stärksten ist. Das besondere der vertikalen Ausrichtung ist, dass sie sowohl auf mechanischem als auch auf elektrischem Wege erzeugt werden kann. Die Hauptstrahlrichtung der Antenne wird im zweiten Fall verändert, ohne dass sich die mechanische Ausrichtung der Antenne ändert. Der Vorteil der elektrischen Neigungsveränderung ist, dass sie ferngesteuert erfolgen

kann, wohingegen Änderungen der horizontalen oder mechanischen vertikalen Ausrichtung manuell durchgeführt werden müssen. Für Netzplaner ist diese Eigenschaft von großem Wert, denn durch elektrische Veränderung der Neigung können bestehende Netze auf einfache Weise neuen Gegebenheiten – beispielsweise neue Antennen im Abdeckungsgebiet – angepasst werden. Dabei ist die Variation des elektrischen Tilts einer Antenne genauso wirksam, wie die Variation des mechanischen Tilts.

Die Neigungswinkel sind bei GSM-Netzen in der Regel geringer als bei UMTS Netzen, da die Interferenzreduktion hauptsächlich durch Frequenzzuweisung erreicht wird und man sich bei der Ausrichtung traditionell auf Abdeckungsaspekte konzentriert hat. Geringere Neigungen führen in der Regel zu besserer Abdeckung. In UMTS-Netzen ist die Veränderung der Ausrichtungen jedoch ein primäres Werkzeug der Interferenzoptimierung. In GSM-Netzen liegen die Neigungswinkel häufig zwischen null und vier Grad, in UMTS Netzen zwischen null und acht Grad. In beiden Fällen sind jedoch Neigungswinkel bis zu zwölf Grad oder mehr möglich.

In unseren Modellen verwenden wir zur Beurteilung einer Konfiguration nicht direkt die Interferenz, sondern eine Größe, die wir als *Interferenzpotenzial* oder *Kopplung* bezeichnen möchten. Wir tun dies, weil die für eine Netzkonfiguration auftretende Interferenz in der Regel schwer zu ermitteln ist. In der GSM-Technologie wird nach Bestimmung der Zellkapazitäten eine Frequenzplanung durchgeführt. Diese ist aufwändig und erfordert komplexe Algorithmen. Erst nach erfolgter Frequenzplanung kann man Aussagen treffen über die im Netz zu erwartende Interferenz. Unsere Optimierungsalgorithmen greifen wegen der Komplexität nicht auf die Ergebnisse einer Frequenzzuweisung zurück. Stattdessen definieren wir Kopplung als Maß für die zu erwartende Interferenz. Starke Kopplung zweier Zellen führt nicht automatisch starker Interferenz. Existieren in einem Netz allerdings viele stark gekoppelte Antennenpaare, so führt die Frequenzzuweisung im Allgemeinen zu schlechten Ergebnissen.

In UMTS-Netzen kann man erst Aussagen über die Interferenz machen, wenn man Informationen hat über die zu erwartenden Sendeleistungen der Antennen. In UMTS-Netzen werden die Sendeleistungen sehr schnell der momentanen Situation angepasst (*fast power control*, siehe [15]). Die Interferenz hängt in erheblichen Maße von diesen Sendeleistungen ab. Wir treffen Aussagen über die Qualität eines UMTS-Netzes, indem wir die durchschnittlichen erwarteten Sendeleistungen der Antennen bestimmen. Ihre Ermittlung ist nach unserem Modell nicht so aufwändig wie die Frequenzzuweisung in GSM. Sie ist aber so aufwändig, dass wir sie in unseren Algorithmen vermeiden wollen. Auch in UMTS definieren wir Kopplungswerte als Maß für das Interferenzpotenzial von Antennenpaaren.

Nachdem unsere Algorithmen Konfigurationen ermittelt haben, nach denen das Interferenzpotenzial minimal ist, führen wir die beschriebenen Maßnahmen durch, um die tatsächlich auftretende Interferenz zu bestimm-

ten. Unsere Analysen werden sowohl in der GSM- als auch in der UMTS-Technologie belegen, dass unser Planungsansatz zu Erfolgen führt. In GSM konnte die Gesamtinterferenz auf allen untersuchten Instanzen reduziert werden, in UMTS auf fast allen.

Wir möchten unser Vorgehen in dieser Arbeit an Abbildung 2.6 deutlich machen.

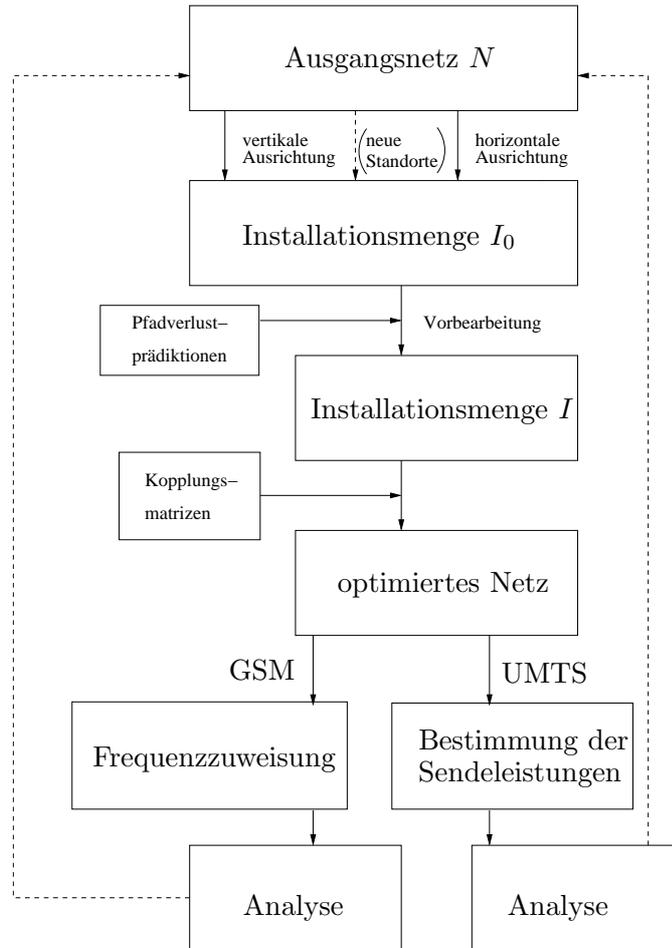


Abbildung 2.6: Unser Vorgehen bei der Optimierung.

Gegeben ist ein Ausgangsnetz  $N$ . (Unter Umständen enthält es keine Antenne, nämlich dann, wenn ein Planungsgebiet neu erschlossen wird.) Für jede Antenne betrachten wir eine Menge von Ausrichtungen. Eine Antenne zusammen mit einer Ausrichtung bezeichnen wir als *Installation*. Betrachten wir für eine Antenne also zwei Ausrichtungen, so gehen beide als Installation in die Menge  $I_0$  ein. Die so erhaltene Installationsmenge  $I_0$  wird mit Hilfe von Pfadverlust-Prädiktionen analysiert. Mitunter stellt sich heraus,

dass einige Installationen aus verschiedenen Gründen nicht infrage kommen. Diese werden in einem Vorbereitungsschritt aus  $I_0$  entfernt, sodass wir die Installationsmenge  $I$  erhalten. Aus dieser wird dann mit Hilfe der Kopplungsmatrizen eine optimierte Konfiguration ausgewählt. Im nächsten Schritt wird für das erhaltene Netz abhängig von der verwendeten Technologie ein relativ aufwändiger Analyseschritt vollzogen. In GSM führen wir eine Frequenzzuweisung durch, in UMTS bestimmen wir die Sendeleistungen. Die Qualität des Netzes wird nicht allein an der auftretenden Interferenz, sondern noch an anderen Indikatoren gemessen.

In vielen Szenarien erhalten wir durch die Netzkonfiguration Rückschlüsse darauf, welche Eigenschaften von Installationen besonders vorteilhaft sind. In einem solchen Fall können wir das optimierte Netz als Ausgangspunkt einer erneuten Optimierung verwenden. Unter Umständen sind wir mit der Netzkonfiguration in einem Teil des Planungsgebietes zufrieden, möchten sie in anderen Teilen aber nachbessern. Hier können wir das optimierte Netz ebenfalls als Ausgangspunkt für einen neuen Durchlauf verwenden.



# Kapitel 3

## Mathematische Grundlagen

In diesem Kapitel erläutern wir einige mathematische Begriffe, die für diese Arbeit von Bedeutung sind. In Abschnitt 3.1 beschreiben wir die drei Optimierungsprobleme *Set Covering*, *Set Packing* und *Set Partitioning*. In Abschnitt 3.2 beschäftigen wir uns in der Hauptsache mit 0/1 Programmierung, wofür wir in Abschnitt 3.3 ein Beispiel geben.

### 3.1 Set Covering, Set Packing, Set Partitioning

In diesem Abschnitt betrachten wir drei verwandte Probleme, das *set covering*-Problem, das *set partitioning*-Problem und das *set packing*-Problem. Alle drei Probleme werden üblicherweise als ganzzahlige Programme mit 0/1-Variablen formuliert. Durch logische Reduktionen kann man normalerweise sowohl Vereinfachungen als auch Verstärkungen der Problemformulierung erreichen. Wir stellen einige dieser Methoden vor.

Gegeben sei eine endliche Grundmenge  $S$  und eine Menge von Teilmengen  $P \subset 2^S$ . Das klassische *set covering*-Problem (SCP) besteht darin, die Menge  $S$  mit möglichst wenigen Elementen  $P$  zu überdecken. Das *set packing*-Problem (SSP) besteht darin, möglichst viele Mengen aus  $P$  auszuwählen, die sich nicht überschneiden. Das *set partitioning*-Problem (SPP) schließlich besteht darin, mit möglichst wenigen Elementen aus  $P$  eine Partition von  $S$  zu bilden. Üblicherweise wird den Mengen ein Gewicht zugeordnet, und man versucht, die Summe der Gewichte zu minimieren bzw. zu maximieren. Alle drei Probleme sind  $\mathcal{NP}$ -schwer (siehe z.B. [5]).

Trotz ihrer Ähnlichkeit unterscheiden sich SCP, SPP und SSP in wichtigen Details. So ist die Lösbarkeit beim SSP klar, beim SCP ist sie leicht überprüfbar – man muss nur prüfen, ob  $S$  abgedeckt ist, wenn *alle* Elemente aus  $P$  verwendet werden. Beim SPP hingegen ist die Klärung der Lösbarkeit eine schwierige Aufgabe.

Wir möchten die drei Probleme als ganzzahlige Optimierungsprobleme formulieren. Hierbei ordnen wir jeder Menge  $P_k$  aus  $P$  ein Gewicht  $\omega_k$  zu.

Sei  $A = (a_{ik}), a_{ik} \in \{0, 1\}$  eine  $m \times n$ -Matrix,  $\omega$  ein  $n$ -Vektor, dann lauten die Formulierungen

$$\begin{array}{lll}
 \text{(SSP)} & \max \omega^T x & \text{(SPP)} & \min \omega^T x & \text{(SCP)} & \min \omega^T x \\
 & Ax \leq 1 & & Ax = 1 & & Ax \geq 1 \\
 & x \geq 0 & & x \geq 0 & & x \geq 0 \\
 & x \in \{0, 1\}^n & & x \in \{0, 1\}^n & & x \in \{0, 1\}^n
 \end{array}$$

Diese Formulierung ist aus [4] übernommen. (Die Bedingungen  $x \geq 0$  sind eingefügt, damit das Weglassen der 0/1-Bedingungen zu einer für uns geeigneten Relaxierung führt.) Den Wert der Optimallösung bezeichnen wir mit  $\omega_0$ .

Befassen wir uns ein wenig genauer mit der IP-Formulierung. In dieser Arbeit beschäftigen wir uns hauptsächlich mit dem SCP, deswegen verwenden wir dieses Problem als Beispiel. Die Spalten der Matrix  $A$  korrespondieren zu Elementen aus  $P$ , die Zeilen zu Elementen aus  $S$ . Es gilt  $a_{ik} = 1$ , wenn das Element  $i$  aus  $S$  in der Teilmenge  $P_k \in P$  enthalten ist. Seien die Zeilen von  $A$  mit  $a_i^T$  bezeichnet. Die einzelnen Bedingungen  $a_i^T x \geq 1$  heißen *Abdeckungsbedingungen*. Je nachdem, ob in einer Teilmenge  $C$  von  $P$  eine Menge  $P_k$  existiert, die Element  $i$  enthält oder nicht, ist die  $i$ -te Abdeckungsbedingung erfüllt oder nicht. Entscheiden wir uns, Menge  $P_k$  in unsere Abdeckung mit einzubeziehen – setzen wir also  $x_k = 1$  –, so sind für alle Zeilen  $a_j$  mit  $j \in P_k$  die Abdeckungsbedingungen erfüllt.

Die im SPP enthaltenen Gleichheitsbedingungen nennen wir *Partitionsbedingungen*. Die in dieser Arbeit vorgestellten ganzzahligen Optimierungsprobleme enthalten beide Typen von Bedingungen. Die „ $\leq$ “-Bedingungen aus dem *set packing* kommen in unseren Modellen nicht vor.

Die Abdeckungs- und Partitionsbedingungen entsprechen logischen *Oder-* bzw. *Entweder-oder-*Verknüpfungen (siehe auch Abschnitt 3.2).

Daraus ergeben sich unter Umständen logische Implikationen, die man in neue Gleichungen und Ungleichungen transformieren kann. Diesen Vorgang bezeichnet man als *Vorbearbeitung (Preprocessing)*. Bei der Vorbearbeitung werden Variablen fixiert, redundante Ungleichungen entfernt und weitere Operationen durchgeführt. Die bei der Vorbearbeitung verwendeten Methoden heißen *Reduktionen (reductions, [4])*. Wir möchten hier ein paar Reduktionen angeben, die bei SCP und SPP für uns relevant sind. Eine umfangreichere Liste mit weiterführenden Literaturangaben findet man in [4].

**Beobachtung 3.1. (Leere Spalte.)** Existiert eine Spalte  $k$  mit  $a_{ik} = 0$  für alle  $i \in S$ , so wird

- Spalte  $k$  entfernt, falls  $\omega_k > 0$ ,
- Spalte  $k$  entfernt und  $\omega_k$  zu  $w_0$  hinzuaddiert, falls  $\omega_k \leq 0$ .

Variable  $x_k$  entspricht also der Menge  $P_k = \emptyset$ . Je nachdem, welches Gewicht die leere Menge hat, wird sie in die Abdeckung mit einbezogen oder nicht.

**Beobachtung 3.2. (Leere Zeile.)** Existiert eine Zeile  $a_i^T$  mit  $a_{ik} = 0$  für alle  $k = 1, 2, \dots, n$ , so ist das Problem ungültig.

Das Element  $i$  aus  $S$  ist in keiner der Mengen aus  $P$  enthalten. Folglich existiert keine Abdeckung.

Eine Variable  $x_k$  ist zur Variablen  $x_l$  benachbart, wenn eine Partitionsbedingung  $a_j^T x = 1$  existiert mit  $a_{jk} = a_{jl} = 1$ . Die Mengen  $P_k$  und  $P_l$  überschneiden sich also und schließen sich beim SPP gegenseitig aus.

**Beobachtung 3.3. (Einzelnes Element.)** Existiert eine Zeile  $a_i^T$ , sodass  $a_{ik} = 1$  für genau ein  $k$ , so wird Variable  $x_k$  auf eins fixiert. Das bedeutet:

- Lösche Spalte  $k$ , addiere  $\omega_k$  zu  $\omega_0$ .
- Lösche alle Zeilen  $a_j^T$ , für die gilt  $a_{jk} = 1$ .
- Falls gilt  $a_i^T x = 1$  (Partitionsbedingung), setze alle benachbarten Variablen  $x_l$  auf null.

Bisher wurde immer nur eine Gleichung bzw. Ungleichung betrachtet. Die folgende Regel bietet ein Beispiel dafür, was passiert, wenn ein Problem sowohl Abdeckungs- als auch Partitionsbedingungen enthält.

**Beobachtung 3.4. (Dominierte Zeile.)** Eine Zeile  $a_i^T$  dominiert eine Zeile  $a_j^T$ , falls gilt  $a_{ik} \geq a_{jk}$  für alle  $k \in \{1, \dots, n\}$ .

- Entsprechen beide Zeilen Abdeckungsbedingungen, gilt also  $a_i^T x \geq 1$ ,  $a_j^T x \geq 1$ , so kann Zeile  $a_j^T x \geq 1$  aus  $A$  entfernt werden.
- Enthält das System die Zeilen  $a_i^T x \geq 1$  und  $a_j^T x = 1$ , so kann Zeile  $a_i^T x \geq 1$  aus  $A$  entfernt werden.
- Entsprechen beide Zeilen Partitionsbedingungen, so kann Zeile  $a_j^T x = 1$  (!) aus  $A$  entfernt werden. Des Weiteren können alle Variablen  $x_k$  mit  $a_{ik} = 1$  und  $a_{jk} = 0$  auf den Wert null fixiert werden. Gleiches gilt, wenn  $a_i^T x = 1$  eine Partitionsbedingung und  $a_j^T x \geq 1$  eine Abdeckungsbedingung ist.

$$\begin{array}{l} a_i^T \\ a_j^T \end{array} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ & & \vdots & & \\ 0 & 1 & 0 & \cdots & 1 \end{pmatrix}$$

In allen vier Fällen, die in Beobachtung 3.4 untersucht werden, ist Element  $i \in S$  in allen Mengen  $P_k$  enthalten, die auch  $j$  enthalten. Im ersten Fall muss man also nur dafür sorgen, dass  $j$  abgedeckt ist. Im zweiten Fall reicht es, dafür zu sorgen, dass  $j$  in genau einer Menge enthalten ist. Element  $i$  ist dann automatisch in mindestens einer Menge enthalten. Im dritten Fall, ergeben sich durch das Element  $i$  stärkere Bedingungen an die Mengen aus  $P$ . Das Element  $j$  ist nur in Mengen enthalten, die auch  $i$  enthalten. Daher können wir keine Menge wählen, die  $i$  enthält aber nicht  $j$ , denn dann steht keine Menge mehr zur Wahl, die  $j$  enthält. Es reicht also, dafür zu sorgen, dass  $i$  in genau einer Menge vorkommt, und dass keine Menge gewählt wird, die  $i$  enthält aber nicht  $j$ . Das gleiche gilt, wenn  $j$  nicht in genau einer, sondern in mindestens einer Menge vorkommen soll.

In den von uns in Kapiteln 4 und 5 betrachteten Optimierungsproblemen sind viele Abdeckungsbedingungen und wenige Partitionsbedingungen enthalten. Die Partitionsmatrix ist dort so aufgebaut, dass auf sie keine Reduktionsregeln angewendet werden können. Vereinfachungen ergeben sich in der Abdeckungsmatrix und durch Verwendung von Beobachtung 3.4, wobei man die Abdeckungsbedingungen für die Vereinfachung der Partitionsbedingungen verwendet und umgekehrt.

Für das *set covering*-Problem und das *set packing*-Problem gibt es zahlreiche Klassen von Schnittebenen (siehe [4]). Sie verwenden Strukturen von Untermatrizen von  $A$ . Wir möchten hier einige Beobachtungen notieren, die man machen kann, wenn ein Problem sowohl Abdeckungs- als auch Partitionsbedingungen enthält. Betrachten wir im weiteren Verlauf das folgende Programm: Seien  $A \in \{0, 1\}^{m_1 \times n}$ ,  $B \in \{0, 1\}^{m_2 \times n}$ .

$$\min \omega^T x$$

sodass

$$\begin{aligned} Ax &\geq 1 \\ Bx &= 1 \\ x &\geq 0 \\ x &\in \{0, 1\}^n \end{aligned} \tag{3.1}$$

Dieses IP enthält sowohl Abdeckungs- als auch Partitionsbedingungen, und man kann die folgende Beobachtung machen:

**Beobachtung 3.5.** Sei  $b_i^T x = 1$  eine Partitionsgleichung aus  $B$ ,  $U := \{k \in 1, \dots, n \mid b_{ik} = 1\}$ . Sei  $A_{JU}$  eine Untermatrix von  $A$ , die Zeilen aus einer beliebigen Teilmenge  $J$  der Zeilenindexmenge  $\{1, \dots, m_1\}$ , und die Spalten aus  $U$  enthält. Hat das System

$$\min 1_U^T x$$

sodass

$$\begin{aligned} A_{JU}x &\geq 1 \\ x &\geq 0 \\ x &\in \{0, 1\}^{|U|} \end{aligned} \tag{3.2}$$

einen Wert  $\omega_{0_U} > 1$ , so ist

$$a_V^T x \geq 1$$

mit  $V := S \setminus U$  eine gültige Ungleichung für System 3.1.

$$\begin{pmatrix} \vdots & \vdots \\ A_{JV} & A_{JU} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \cdot \begin{pmatrix} x_V \\ x_U \end{pmatrix} \begin{matrix} \geq \\ = \end{matrix} \begin{pmatrix} \vdots \\ 1 \\ 1 \end{pmatrix}$$

**Beweis: (Beobachtung 3.5)** Die Beobachtung stimmt trivialerweise, wenn System 3.2 keine Lösung hat. Existiert eine Lösung, und nimmt man im Sinne eines Widerspruchs an, dass für alle  $x_k$  mit  $k \in V$  gilt  $x_k = 0$ , so existiert keine Variablenbelegung für die  $x_l$  mit  $l \in U$ , sodass gleichzeitig  $A_{JU}x \geq 1$  und  $b_i^T x = 1$  erfüllt sind. Also gilt für alle Lösungen  $x^*$  des Systems 3.2:  $a_V^T x^* \geq 1$ . ■

Man kann Beobachtung 3.5 auf einfache Weise verstärken, indem man die Aussage beschränkt auf die Variablen, die tatsächlich in  $A_{JV}$  eine Rolle spielen. Sei

$$V' := \{k \in V \mid a_{jk} = 1 \text{ für ein } j \in J\} \subset V.$$

Unter denselben Annahmen wie in Beobachtung 3.5 gilt:

**Corollar 3.1.** Die Ungleichung  $a_{V'}^T x \geq 1$  ist für System 3.2 eine gültige Ungleichung.

Beobachtung 3.5 bietet uns eine Möglichkeit, Schnittebenen zu finden. Allerdings stehen wir vor zwei Problemen: Zum einen müssen wir wissen, welche Mengen  $J$  sich zu betrachten lohnen. Es gibt dafür nämlich  $2^{m_1}$  Kandidaten. Zum anderen ist die Lösung von System 3.2 komplexitätstheoretisch nicht einfacher als die von System 3.1. Dem zweiten Problem kann man für bestimmte Teilmengen  $J$  beikommen:

**Beobachtung 3.6.** Die Bezeichnungen seien wie in Beobachtung 3.5. Hat das Programm

$$\min \omega_V^T x$$

sodass

$$\begin{aligned} A_{JU}x &\geq 1 \\ x &\geq 0 \end{aligned} \tag{3.3}$$

einen Wert  $\omega_{0_U} > 1$ , so ist  $a_{V'}^T x \geq 1$  eine gültige Ungleichung.

Diese Aussage hilft uns insofern, als das Lösen von LPs einerseits in der Praxis schneller geht als das Lösen von IPs und andererseits auch komplexitätstheoretisch einfacher ist.

### 3.2 0/1-Programmierung

Die von uns in dieser Arbeit betrachteten Probleme können als nichtlineare Optimierungsprobleme betrachtet werden. Wir beschreiben hier, wie man boolesche Ausdrücke in Polynome transformieren kann. Mit diesen Polynomen werden nichtlineare 0/1-Programme formuliert. Wir stellen zwei Ansätze vor, wie man diese Programme in gemischt-ganzzahlige Programme übersetzen kann. Diese Formulierung kann mit den von uns vorgestellten Schnittebenen verschärft werden.

Ein Boolescher Ausdruck hat den Wahrheitswert *wahr* oder *falsch*. Wir bilden den Wert *falsch* auf die Zahl Null ab und den Wert *wahr* auf eine Zahl größergleich eins. In einem booleschen Programm ist eine Menge boolescher Ausdrücke und eine Zielfunktion gegeben. Die Aufgabe besteht darin, eine Variablenbelegung zu finden, sodass die Zielfunktion einen optimalen Wert annimmt. Konkret sind Variablen

$$x_i, \quad i \in I$$

gegeben. Sie können den Wert null (falsch) und eins (wahr) annehmen. Einen booleschen Ausdruck stellen wir uns so vor, dass er aus Und- und Oder-Verknüpfungen besteht. Die Aussage „ $x_1$  oder  $x_2$ “ wird durch das Polynom

$$x_1 + x_2$$

und die Aussage „ $x_1$  und  $x_2$ “ wird durch das Polynom

$$x_1 \cdot x_2$$

wiedergegeben. Dabei fällt auf, dass der Wert des Polynoms  $x_1 + x_2$  auch zwei sein kann.

Der Ausdruck „*nicht*  $x_i$ “ wird offensichtlich durch das Polynom

$$1 - x_i$$

widergespiegelt.

Betrachten wir ein Beispiel. Wir möchten die Aussage

$$x_1 \text{ und } x_2 \text{ und nicht } x_3$$

in ein Polynom umwandeln. Wie man sich leicht klar macht, erfüllt das Polynom

$$x_1 \cdot x_2 \cdot (1 - x_3)$$

die Eigenschaft, dass sein Wert null ist, genau dann wenn die obige Aussage falsch ist, und eins sonst. In dieser Arbeit treten oftmals Ausdrücke der Form

$x_1$  und  $x_2$  und (kein  $x_k$  mit  $k \in Z$ )

auf. Die Menge  $Z$  ist dabei eine endliche Menge. Offenbar hat das entsprechende Polynom die Form

$$x_1 \cdot x_2 \cdot \prod_{k \in Z} (1 - x_k). \quad (3.4)$$

Es ist üblich, die Aussage „nicht  $x_i$ “ auch als  $\bar{x}_i$  zu codieren. Ausdruck (3.4) erhält dann die Form

$$x_1 \cdot x_2 \cdot \prod_{k \in Z} \bar{x}_k.$$

Ein nichtlineares 0/1-Programm mit Nebenbedingungen (*constrained 0 – 1 program*, siehe [12]) – im folgenden 0/1-Problem genannt – hat die Form

$$\min f(x) = \sum_{k=1}^p c_k T_k \quad (3.5)$$

sodass

$$g_i(x) = \sum_{k=1}^{p_i} a_{ik} T_{ik} \leq b_i, \quad i = 1, 2, \dots, m \quad (3.6)$$

$$x \in \{0, 1\}^n, \quad (3.7)$$

wobei

$$T_k = \prod_{j \in N_k} x_j, N_k \subset N = \{1, 2, \dots, n\}, \quad k = 1, 2, \dots, p \quad (3.8)$$

und

$$T_{ik} = \prod_{j \in N_{ik}} x_j, N_{ik} \subset N, k = 1, 2, \dots, p_i, \quad i = 1, 2, \dots, m. \quad (3.9)$$

Die Polynome  $T_k, T_{ik}$  sind hier Repräsentanten logischer Ausdrücke, wie sie oben beschrieben wurden. Die Optimierungsaufgabe besteht also darin, eine Belegung der  $x_i$  Variablen zu finden, sodass die gewichtete Summe der Polynome  $T_k$  möglichst gering ist. Dabei müssen die gewichteten Summen der  $T_{ik}$ -Ausdrücke bestimmte Eigenschaften erfüllen. In dieser Arbeit werden nur Probleme mit positiven  $c_k$ -Koeffizienten betrachtet.

Zur Lösung von 0/1-Problemen gibt es mehrere Klassen von Ansätzen. Sie sind zum Beispiel im Übersichtsartikel von Hansen, Jaumard und MATHON ([12]) beschrieben. Zur Lösung der in dieser Arbeit behandelten Aufgaben benutzten wir zwei dieser Methoden: die Formulierung als gemischt-ganzzahliges Programm und Heuristiken. Eine dritte Methode besteht in der Verwendung von Schnittebenenverfahren. Hierfür geben wir in diesem Abschnitt eine allgemein gültige Klasse von Schnittebenen.

Unser erster Ansatz besteht darin, das Problem in ein gemischt-ganzzahliges Programm (MIP) umzuwandeln. Als einer der ersten beschäftigte sich Fortet [9] mit dieser Idee. Jeder Ausdruck  $T_k$  (oder  $T_{ik}$ ) wird dabei durch eine Variable  $0 \leq x_{n+k} \leq 1$  ersetzt, die durch die Ungleichungen

$$\sum_{j \in N_k} x_j - |N_k| + 1 \leq x_{n+k} \quad (3.10)$$

und

$$|N_k| \cdot x_{n+k} \leq \sum_{j \in N_k} x_j \quad (3.11)$$

beschränkt ist. Diese Formulierung geht auf Fortet ([9]) zurück. Ungleichung 3.11 kann dabei durch Glovers Ungleichungen

$$x_j \geq x_{n+k} \quad \forall j \in N_k \quad (3.12)$$

ersetzt werden (siehe [10]). Die erste Formulierung hat dabei weniger Ungleichungen, die zweite ist allerdings stärker, da Ungleichung 3.11 eine konische Kombination der Ungleichungen 3.12 ist.

Ein Vorteil bei der Transformation in ein MIP ist, dass wir die existierenden Lösungsverfahren, namentlich die Software CPLEX [3], verwenden können. Selbst wenn eine Instanz zu groß ist, um optimal gelöst zu werden, erhalten wir eine Abschätzung für die Güte der Lösung.

Man erhält ein ganzzahliges Programm (IP), wenn man fordert

$$x_{n+k} \in \{0, 1\}.$$

Eine zweite zur Lösung von 0/1-Programmen häufig verwendete Methode sind Enumerationsverfahren. Sie durchlaufen die Lösungsmenge in einer vom Anwender definierten Reihenfolge. Ist diese Reihenfolge für die Problem Instanz gut geeignet, so können große Äste des Analysebaums abgeschnitten werden. Nach Terminierung erhält man die bewiesenermaßen optimale Lösung.

Sind die betrachteten Problem Instanzen sehr groß, so wird unter Umständen darauf verzichtet, alle Lösungen zu betrachten oder auszuschließen. Mit heuristischen Auswahlverfahren betrachtet man nur eine Teilmenge der möglichen Lösungen. Bei diesem Verfahren erhält man nach Terminierung keine Optimalitätsgarantie. Wir verwendeten Algorithmen der letzteren Art. Eine dritte Möglichkeit besteht in der Verwendung von Schnittebenenverfahren (siehe [11]). Dabei wird das Problem auf eine andere Weise linearisiert als bei Fortet. Zielfunktion und Nebenbedingungen werden hierbei durch eine Menge linearer Ungleichungen repräsentiert, deren Größe exponentiell ist im Verhältnis zur Zahl der Variablen. Daher wird immer nur ein Teil der Ungleichungen gleichzeitig betrachtet. Nachdem eine Optimallösung gefunden wurde, werden dem System neue Ungleichungen hinzugefügt und eventuell

alte entfernt, sofern die Optimallösung nicht ganzzahlig ist. Ein solches Verfahren heißt *Schnittebenenverfahren*. Balas und Mazzola haben in [2] solche Ungleichungen hergeleitet. Sie basieren auf dem folgenden

**Satz 3.1.** Sei  $f(x)$  eine Zielfunktion wie in 3.5, mit  $c_k > 0$  für alle  $k \in N := \{1, 2, \dots, p\}$ . Für jedes  $M \subset N$  sei

$$g_M(x) = \sum_{i \in M} \left( \sum_{k \in M | i \in T_k} c_k \right) x_i - \sum_{k \in M} (|T_k| - 1) a_k. \quad (3.13)$$

Dann gilt für alle  $x \in \{0, 1\}^p$

$$f(x) \geq g_M(x) \quad (3.14)$$

für alle  $M \subset N$ .

### 3.3 Ein Beispiel: Das MAX-CUT Problem

In diesem Abschnitt geben wir ein Beispiel für ein 0/1-Problem. Dafür benutzen wir das bekannte MAX-CUT-Problem auf einem Graphen. Wir geben zunächst die nichtlineare Formulierung und zeigen dann, wie man das Problem auf verschiedene Weisen in ein IP transformieren kann. Die verschiedenen Formulierungen werden uns den nachfolgenden Kapiteln helfen, die Struktur der dortigen Optimierungsprobleme zu verstehen. Außerdem helfen sie uns in Abschnitt 4.3.2 bei der Entwicklung von Schnittebenen für das Obermengenmodell.

Gegeben sei ein ungerichteter Graph  $G = (V, E)$  mit  $V = \{v_1, \dots, v_n\}$  und eine Gewichtsfunktion  $c : E \rightarrow \mathbb{R}^+$ . Wir wählen für ungerichtete Kanten  $e = \{u, v\}$  die kürzere Schreibweise  $e = uv$ . Das Problem

$$\max \sum_{e \in \delta(U)} c(e), \quad U \subset V$$

heißt MAX-CUT Problem. In Worten besteht es darin, aus der Knotenmenge  $V$  eine Teilmenge  $U$  so auszuwählen, dass die Summe der Gewichte der Kanten, die ein Ende in  $U$  und das andere in  $V \setminus U$  haben, maximal ist.

Seien  $x_i \in \{0, 1\}$ ,  $i \in \{1, \dots, n\}$  Entscheidungsvariablen. Es gelte  $x_i = 1$ , falls  $v_i \in U$  und  $x_i = 0$  sonst. Es sei  $\bar{x}_i := 1 - x_i$  für alle  $i = 1, 2, \dots, n$ . Dann lautet eine Formulierung des MAX-CUT Problems als 0/1-Problem:

$$\begin{aligned} \max \quad & \sum_{e=uv \in E} c(e)(x_u \cdot \bar{x}_v + \bar{x}_u \cdot x_v) \\ & x_v + \bar{x}_v = 1 \quad \forall v \in V \\ & x \in \{0, 1\}^n. \end{aligned} \quad (3.15)$$

Die Linearisierung mit den Ungleichungen vom Typ 3.10 und 3.12 von Fortet bzw. Glover (siehe [9] bzw. [10]) lautet

$$\max \sum_{e \in E} (c(e) \cdot y_e + c(e) \cdot y'_e) \quad (3.16)$$

sodass

$$\begin{aligned} x_u + \bar{x}_v - 1 &\leq y_e \quad \forall e = uv \in E \\ \bar{x}_u + x_v - 1 &\leq y'_e \\ x_u &\geq y_e \\ \bar{x}_v &\geq y_e \\ \bar{x}_u &\geq y'_e \\ x_v &\geq y'_e \\ x_v + \bar{x}_v &= 1 \quad \forall v \in V \\ x, \bar{x} &\in \{0, 1\}^n. \end{aligned} \quad (3.17)$$

Eine IP-Formulierung erhält man durch Hinzufügen der Ungleichungen

$$y, y' \in \{0, 1\}^{|E|}$$

Wir erhalten ein Minimierungsproblem, wenn wir versuchen, die Summe der Gewichte aller Kanten, die nicht im Schnitt sind, zu minimieren. Ein zum MAX-CUT Problem äquivalentes Problem ist also

$$\min \sum_{e \notin \delta(U)} c(e), \quad U \subset V.$$

Eine Formulierung dieses Problems als 0/1-Problem lautet

$$\begin{aligned} \min \sum_{e=uv \in E} c(e)(x_u \cdot x_v + \bar{x}_u \cdot \bar{x}_v) \\ x_v + \bar{x}_v &= 1 \quad \forall v \in V \\ x &\in \{0, 1\}^n. \end{aligned} \quad (3.18)$$

Dabei erhalten wir den Wert des maximalen Schnittes durch Abzug des Optimalwertes von Programm 3.18 vom Term  $\sum_{e \in E} c(e)$ . Transformieren wir das Problem wie oben, so erhalten wir das IP

$$\min \sum_{e \in E} (c(e) \cdot y_e + c(e) \cdot y'_e)$$

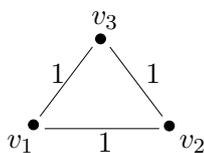


Abbildung 3.1: Ein Beispielgraph

sodass

$$\begin{aligned}
 x_u + x_v - 1 &\leq y_e \quad \forall e = uv \in E \\
 \bar{x}_u + \bar{x}_v - 1 &\leq y'_e \\
 x_u &\geq y_e \\
 x_v &\geq y_e \\
 \bar{x}_u &\geq y'_e \\
 \bar{x}_v &\geq y'_e \\
 x_v + \bar{x}_v &= 1 \quad \forall v \in V \\
 x, \bar{x} &\in \{0, 1\}^n \\
 y, y' &\in \{0, 1\}^{|E|}.
 \end{aligned} \tag{3.19}$$

Da wir in unserer Definition des MAX-CUT Problems nur nicht-negative Kantengewichte zugelassen haben, können wir die Ungleichungen weglassen, die die  $y$ - und  $y'$ -Variablen nach oben beschränken.

Diese Formulierung des MAX-CUT Problems mag nicht die einfachste sein. Wir betrachten sie hier jedoch ein wenig eingehender, weil die Optimierungsprobleme der nachfolgenden Kapitel diese Form aufweisen. Betrachten wir den Graphen  $C_3$  in Abbildung 3.1 als Beispiel.

Alle Kantengewichte seien eins. Es ist klar, dass das MAX-CUT Problem hier den Wert zwei hat. Jede Menge, die einen oder zwei Knoten enthält, liefert einen solchen Schnitt. Problem 3.18 hat also den Optimalwert eins. Diesen Wert erhalten wir auch aus der IP-Formulierung des Problems:

$$\min y_{12} + y_{13} + y_{23} + y'_{12} + y'_{13} + y'_{23}$$

unter den Nebenbedingungen

$$\begin{array}{rcll}
 x_1 + x_2 & & -1 & \leq y_{12} \\
 x_1 & & + x_3 & -1 \leq y_{13} \\
 & x_2 + x_3 & & -1 \leq y_{23} \\
 \bar{x}_1 + \bar{x}_2 & & & -1 \leq y'_{12} \\
 \bar{x}_1 & & + \bar{x}_3 & -1 \leq y'_{13} \\
 & \bar{x}_2 + \bar{x}_3 & & -1 \leq y'_{23}
 \end{array}$$

$$\begin{aligned}x_1 + \bar{x}_1 &= 1 \\x_2 + \bar{x}_2 &= 1 \\x_3 + \bar{x}_3 &= 1 \\x &\in \{0, 1\}^3 \\y, y' &\in \{0, 1\}^3\end{aligned}$$

Der Optimalwert des Systems ist eins. Das ist leicht zu sehen, denn es haben mindestens zwei  $x$ -Variablen oder zwei  $\bar{x}$ -Variablen den Wert eins. Seien das o.B.d.A die Werte  $x_1$  und  $x_2$ . Dann hat die Variable  $y_{12}$  auch den Wert eins. Eins ist also eine untere Schranke. Wählt man  $x_1 = x_2 = \bar{x}_3$ , so hat diese Lösung den Wert eins und ist somit optimal.

Es zeigt sich allerdings, dass die Relaxierung des Programms den Wert null hat. Das ergibt sich, wenn man für alle  $x$ -Variablen den Wert  $\frac{1}{2}$  wählt. Hier zeigt sich das Problem dieser linearen Formulierung. Die LP-Relaxierung liefert eine schlechte Abschätzung des Optimalwerts des IPs.

## Kapitel 4

# Planung interferenzarmer GSM-Netze

In diesem Kapitel übertragen wir den in Kapitel 2 beschriebenen Ansatz zur Planung von Funknetzen auf die GSM-Technologie. Wir entwickeln ein mathematisches Modell zur Berechnung des Interferenzpotenzials einer Netzkonfiguration und beschreiben daraufhin die von uns entwickelten und angewendeten Lösungsmethoden. Schließlich stellen wir am Ende des Kapitels unsere Ergebnisse vor.

Abschnitt 4.1 enthält unseren auf GSM-Netze umformulierten Planungsansatz. Wir beschreiben zunächst, wie für eine Netzkonfiguration ein Frequenzplan ermittelt wird. Hierzu erläutern wir, wie durch die Bestimmung *best server*-Bereiche mit Hilfe von Verkehrskarten Zellkapazitäten festgelegt werden. Sind die Kapazitäten bestimmt, wird für jeden Sender ein Funkkanal festgelegt. Dieser Prozess heißt Frequenzzuweisung. Wir beschreiben in Abschnitt 4.1 unser Modell zur Frequenzzuweisung, das aus [6] übernommen wurde. Ziel der Frequenzzuweisung ist es, die Funkkanäle so zuzuweisen, dass im Netz keine oder nur sehr wenig Interferenz entsteht. In dem von uns verwendeten Modell werden die Sendefrequenzen auf Basis von Interferenzmatrizen ermittelt. Für unseren Planungsansatz verwenden wir die von uns eingeführten *Kopplungsmatrizen*. Aus ihnen können die Interferenzmatrizen abgeleitet werden. Unsere Planungsansatz besteht darin, diese Kopplungsmatrizen zur Beurteilung von Netzkonfigurationen zu verwenden. Wir formulieren drei Thesen, welche Eigenschaften von Kopplungsmatrizen für die Frequenzplanung wünschenswert sind. Unser Ziel wird sein, die Summe der in der Kopplungsmatrix notierten Kopplungswerte zu minimieren.

In Abschnitt 4.2 geben wir eine erste mathematische Formulierung unseres Planungsansatzes. Er basiert auf der Auswertung der Kopplung zwischen Installationen auf Pixelbasis. Daher heißt unser erstes Modell *Pixel-Modell*. Die erste Formulierung hat eine nichtlineare Zielfunktion. Wir linearisieren es in der Folge, und stellen fest, dass die IPs sehr groß sind, und die uns verfügbaren IP-Löser sie nicht lösen können. Wir stellen zwei Klassen von

Schnittungleichungen vor, mit denen die Problemformulierung verbessert wird. Das Lösungsverhalten ändert sich jedoch nur geringfügig, weswegen wir uns nach anderen Problemformulierungen umsehen.

Eine solche Formulierung stellen wir in Abschnitt 4.3 vor. Das dort entwickelte *Obermengen-Modell* ist kompakter, denn es geht aus dem in Abschnitt 4.2 vorgestellten nichtlinearen Programm durch Zusammenfassung von Termen der Zielfunktion hervor. Dieses Modell basiert nicht mehr auf der Bewertung von Kopplung auf Pixeln, sondern beurteilt die Kopplung zwischen Antennenpaaren danach, wie Teilmengen der restlichen Antennen konfiguriert sind. Die Systeme, die wir mit dem Obermengen-Modell erzeugen sind kleiner als die pixelbasierte, aber nicht besser lösbar. Die in Abschnitt 4.2 entwickelten Schnittungleichungen sind für das Obermengen-Modell nicht verwendbar, und so entwickeln wir neue Klassen von Schnittungleichungen. Auch im Obermengen-Modell gelingt es uns nicht, das Lösungsverhalten des IP-Algorithmus entscheidend zu verbessern.

In Abschnitt 4.4 stellen wir zwei Möglichkeiten vor, das Obermengen-Modell zu vereinfachen. Die erste Vereinfachung besteht darin, bestimmte Kopplungsbeziehungen, die wir für „wenig wahrscheinlich“ halten, nicht zu berücksichtigen. Diese Methode führt zu kleineren Programmen, die schneller lösbar sind. Die ermittelten Lösungen sind für das vollständige Programm jedoch nicht mehr optimal. Wir untersuchen, welche Kopplungsbeziehungen wir vernachlässigen dürfen, um trotzdem noch Lösungen angemessener Qualität zu erhalten.

Die zweite Vereinfachung des Obermengen-Modells besteht darin, den Schwellwert zu verändern, ab wann wir sagen, dass eine Antenne eine andere stört. Auch hierdurch erhalten wir Programme, die schneller lösbar sind, deren Lösungen jedoch ebenfalls nicht notwendigerweise optimal sind für das Ausgangsprogramm.

In Abschnitt 4.5 wenden wir uns von der IP-Formulierung ab und gehen das Optimierungsproblem auf heuristische Weise an. Wir stellen zunächst eine Heuristik vor, die auf der Auswertung von Antennenausrichtungen basiert, ohne Berücksichtigung von Pfadverlust-Prädiktionen. Mit ihr erzielen wir gute Ergebnisse, die wir als Ausgangslösungen für eine zweite Heuristik, eine lokale Suche, verwenden. Der lokale Suchalgorithmus liefert uns von allen vorgestellten Lösungsverfahren die besten Lösungen. Auf kleinen Probeinstanzen, die mit dem IP-Modell gelöst werden können, findet er optimale Lösungen. Auf größeren Instanzen ermittelt er auf den von uns untersuchten Instanzen Netzkonfigurationen, die ein deutlich besseres Interferenzverhalten aufweisen als die Ausgangskonfigurationen.

In Abschnitt 4.6 tragen wir alle Ergebnisse zusammen. Dabei werden die Ausgangsnetze mit den optimierten Konfigurationen verglichen. Die Analyse erfolgt in zwei Stufen. Zunächst vergleichen wir die Kopplungsmatrizen der Ausgangsnetze mit denen der optimierten Konfigurationen. Es zeigt sich, dass die Kopplungswerte der optimierten Netze im Allgemeinen deutlich

kleiner sind als die der Ausgangsnetze. Die optimierten Konfigurationen weisen also ein deutlich geringeres Interferenzpotenzial auf. Ein Indikator dafür, wie gut eine Netzkonfiguration für die Frequenzzuweisung geeignet ist, sind die Kopplungsgrade der Antennen. Hier wird beurteilt, wie viele andere Antennen bei der Frequenzzuweisung einer Antenne als potenzielle Störer berücksichtigt werden müssen. Die Knotengrade der optimierten Netze sind deutlich kleiner als die der Ausgangsnetze.

Im zweiten Analyseschritt bewerten wir die Frequenzpläne, die aus der Frequenzzuweisung resultieren. Wir zeigen, dass die optimierten Netze nach der Zuweisung deutlich weniger Interferenz aufweisen als die Ausgangsnetze. Unser in Abschnitt 4.1 vorgestellter, auf Kopplungsmatrizen basierender Optimierungsansatz ist also für die integrierte Planung von Abdeckung und Interferenz für GSM-Netze geeignet.

In Abschnitt 4.7 fassen wir dieses Kapitel zusammen.

## 4.1 Optimierungsansatz

In diesem Abschnitt zeigen wir, wie wir unseren allgemeinen Planungsansatz aus Kapitel 2 für das GSM-System formulieren. In Unterabschnitt 4.1.1 zeigen wir zunächst, wie mit Hilfe von Verkehrsdaten die Kapazitätsplanung erfolgt. In Unterabschnitt 4.1.2 definieren wir dann Kopplungsmatrizen, welche das im Netz vorhandene Interferenzpotenzial beschreiben. Die in den Kopplungsmatrizen enthaltenen Kopplungswerte verwenden wir zur Ermittlung von Interferenzwerten, die als Eingangsdaten für einen Frequenzzuweisungs-Algorithmus dienen. Weitere Eingangsdaten für diesen Algorithmus sind die in Unterabschnitt 4.1.2 ebenfalls eingeführten Separierungsbedingungen. In Unterabschnitt 4.1.3 formulieren wir Thesen darüber, welche Eigenschaften die Kopplungsmatrizen haben sollten, damit sie in Interferenzmatrizen transformiert werden können, die eine gute Frequenzzuweisung ermöglichen. Der aus der Frequenzzuweisung resultierende Lösungswert kann direkt als im Netz auftretende Interferenz interpretiert werden. Aus den in Unterabschnitt 4.1.3 formulierten Thesen leiten wir unseren Optimierungsansatz zur Planung interferenzarmer GSM-Netze ab. Eine detaillierte Modellierung und die Vorstellung des daraus hergeleiteten mathematischen Optimierungsproblems folgt in Abschnitt 4.2.

### 4.1.1 Kapazitätsplanung

In diesem Unterabschnitt zeigen wir, wie wir für eine gegebene Netzkonfiguration die Kapazitäten der Zellen bestimmen. Dafür benutzen wir die Erlang-B Formel. Nachdem die Zellkapazitäten festgelegt sind, werden den Sendern nach dem in Unterabschnitt 4.1.2 vorgestellten Modell Frequenzen zugewiesen.

Die Kapazität bestimmt man durch die Zahl der installierten Sender/Empfänger-Einheiten, den TRX (*transmitter/receiver unit*) für eine Antenne. Eine Antenne ist mit bis zu sechzehn dieser Einheiten ausgestattet. Ein TRX verfügt über acht zeitlich separierte *Kommunikationskanäle*. Bis zu zwei dieser Kanäle werden für die Übertragung von Protokollinformationen verwendet. Die übrigen sechs bis acht Kanäle werden für Sprachverbindungen verwendet. Die Anzahl der TRX auf einer Antenne richtet sich nach dem prognostizierten Verkehr im *best server*-Bereich. Vom Verkehr wird angenommen, dass er um einen bestimmten Wert schwankt. Die Anzahl der Kanäle wird so berechnet, dass die Wahrscheinlichkeit, dass ein Kunde keine Verbindung zum Netz herstellen kann – die *Blockierwahrscheinlichkeit* –, einen bestimmten Wert nicht überschreitet.

Wir verwenden die in der Funknetzplanung übliche Erlang-B Formel (siehe [20]). Mit ihrer Hilfe berechnet man die Blockierwahrscheinlichkeit  $B$  in Abhängigkeit von der Zahl der Kunden und der Zahl der zur Verfügung stehenden Kommunikationskanäle  $N$ . Es wird angenommen, dass die Zahl der Kunden um den Wert  $A$  schwankt, daher benötigt man mehr als  $A$  Verbindungen. Die Formel lautet

$$B(A, N) = \frac{1}{\sum_{i=0}^N \frac{A^i}{i!}} \cdot \frac{A^N}{N!}.$$

Wir lassen in unseren Modellen eine Blockierwahrscheinlichkeit von 2% zu (vgl. hierzu [6]). In der mittleren Spalte von Tabelle 4.1 ist zu sehen, wie viele Kommunikationskanäle notwendig sind, um eine gegebene Zahl von Kunden mit einer Blockierwahrscheinlichkeit von weniger als 2% bedienen zu können. In der rechten Spalte sehen wir, wie viele TRX zur Versorgung der Kunden notwendig sind. Wir erkennen, dass das Verhältnis von Kanälen zu Kunden  $N/A$  mit steigender Zahl der Kunden kleiner wird. Am effektivsten kann eine gegebene Zahl von Kunden also mit möglichst wenig Zellen bedient werden.

Kunden $A$	Kanäle $N$	TRX
1	4	1
2	6	1
3	8	2
5	10	2
10	17	3
20	28	4
30	39	6
50	61	8

Tabelle 4.1: Anzahl von Kunden und benötigten Kanälen.

Zur Verminderung der Interferenz wird das Frequenzband, das einem Anbieter bei GSM zur Verfügung steht, in *Funkkanäle* unterteilt. Signale stören sich nur, wenn sie die gleichen oder eng benachbarte Frequenzen verwenden. Jedem TRX wird eine feste Frequenz zugewiesen, auf der es mit den Kunden kommuniziert. In unseren Planungsszenarien ist die Anzahl der verwendeten Sender in einem Netz sehr viel größer als die Anzahl der zur Verfügung stehenden Kanäle. Daher existieren Sender, die dieselbe Frequenz teilen, und die sich somit stören. Die Aufgabe, für die Sender im Netz einen guten Frequenzplan zu ermitteln, heißt *Frequenzzuweisungsproblem* (*frequency assignment problem*, FAP).

#### 4.1.2 Kopplungsmatrizen und Frequenzzuweisung

In diesem Unterabschnitt stellen wir das Frequenzplanungsproblem vor. Es handelt sich dabei um die Aufgabe, für eine gegebene Netzkonfiguration eine Kanalbelegung der Sender zu finden. Hierzu erläutern wir das in [6] entwickelte Modell zur Frequenzzuweisung. Wir führen Dämpfungsfaktoren für Signale ein und beschreiben, wie mit ihnen die Zellen und die *best server*-Bereiche eines GSM-Netzes ermittelt werden können. Danach definieren wir Kopplungswerte und leiten aus ihnen Interferenzwerte ab, die wir als Eingangsdaten für die Frequenzzuweisung verwenden. Nebenbedingungen für die Frequenzzuweisung sind sogenannten *Separierungsbedingungen* für Sender derselben Zelle oder desselben Standortes.

Die Kopplungswerte werden in Kopplungsmatrizen zusammengefasst. Mit Hilfe der Kopplungswerte werden wir in Unterabschnitt 4.1.3 unseren Planungsansatz für GSM-Netze formulieren.

Zur Bestimmung der Interferenzwerte wird das Planungsgebiet zunächst in Pixel eingeteilt. Für jeden Pixel  $p$  wird mit Hilfe der Pfadverlust-Prädiktionen für jede Antenne  $i$  ein *Dämpfungsfaktor* (*attenuation factor*, siehe [7])

$$\gamma_{i,p} = \frac{\text{Empfangsleistung Antenne } i \text{ auf Pixel } p}{\text{Sendeleistung Antenne } i}$$

ermittelt. Man beachte, dass der in Kapitel 2 eingeführte Pfadverlust das Inverse des Dämpfungsfaktors ist. Der Pfadverlust ist also groß, wenn der Dämpfungsfaktor klein ist. Die Dämpfungsfaktoren werden linear dargestellt. Bei ihnen betrachten wir also *Verhältnisse*, wohingegen wir beim Pfadverlust zumeist von *Differenzen* reden.

Ist Antenne  $i$  auf Pixel  $p$  *best server*, so besteht zwischen Antenne  $i$  und Antenne  $j$  Gleichkanal-Kopplung, wenn gilt

$$\frac{\gamma_{i,p}}{\gamma_{j,p}} \leq 21\text{dB} \approx 125.$$

Antenne  $j$  ist dann ein *Störer*. Den Wert 21 dB bezeichnen wir auch als  $\delta$ -Schwellwert. Haben Antennen  $i$  und  $j$  auf  $p$  denselben Pfadverlust, und sind sie beide *best server*, so stören sie sich nach der obigen Definition gegenseitig.

Sendet ein Störer auf einem Pixel auf derselben Frequenz wie ein *best server*, so entsteht *Gleichkanal-Interferenz* (*co-channel interference*). Im GSM-Standard [8] des *European Telecommunications Standards Institute* (ETSI) ist festgelegt, dass ein Empfänger, in der Lage sein soll, ein Signal zu entschlüsseln, wenn das Verhältnis der Signalstärke zur Rauschleistung 9 dB übertrifft. Das bedeutet beispielsweise, dass ein Empfänger etwa sechzehn Störsignale tolerieren kann, deren Stärke um genau 21 dB geringer ist als die des *best servers*. Sind die Störer stärker, so können entsprechend weniger von ihnen toleriert werden.

Es ist mitunter schwierig, Gleichkanal-Interferenz völlig auszuschließen, wenn *best server* und Störer jeweils mehrere TRX, also auch mehrere Kanäle belegen. Sei  $P^i$  die Menge der Pixel, wo Antenne  $i$  *best server* ist. Der *Kopplungswert*

$$y_{i,j} := |\{p \in P^i \mid \frac{\gamma_{i,p}}{\gamma_{j,p}} \leq 21\text{dB}\}|$$

ist die Anzahl der Pixel, auf denen Antenne  $j$  Antenne  $i$  stört, falls sie gleiche Kanäle benutzen. Wir sagen, je größer der Kopplungswert ist, desto größer ist die *Kopplung* zwischen den Antennen. Die Kopplungsmatrix  $Y = (y_{i,j})_{i,j}$  enthält alle Kopplungswerte. In den Einträgen  $y_{i,i}$  der Hauptdiagonale ist die Größe der *best server*-Bereiche der Antennen notiert.

Ist ein Störer fast ebenso stark wie der *best server*, liegt sein Pfadverlust also weniger als 3 dB über dem des *best servers*, so tritt *Nachbarkanal-Interferenz* (*adjacent channel interference*) auf, wenn er eine Frequenz verwendet, die einer Sendefrequenz des *best servers* direkt benachbart ist. Die Werte

$$a_{i,j} := |\{p \in P^i \mid \frac{\gamma_{i,p}}{\gamma_{j,p}} \leq 3\text{dB}\}|$$

geben die *Nachbarkanal-Kopplung* wieder.

In Abbildung 4.1 ist die Kopplung zweier Antennen dargestellt. Auf den weißen Pixeln ist Antenne  $i$  *best server*, auf den hellgrauen Antenne  $j$ . Im linken Bild sind diejenigen Pixel dunkelgrau gezeichnet, auf denen Antenne  $i$  *best server* ist, Antenne  $j$  aber noch so stark sendet, dass potenziell Gleichkanal-Interferenz auftritt. Auf dem rechten Bild gilt dasselbe für die Nachbarkanal-Interferenz. In diesem Beispiel ist  $y_{i,j} = 26$  und  $a_{i,j} = 9$ .

Die für die Frequenzzuweisung bestimmte Interferenz-Matrix wird mit Hilfe der Kopplungswerte bestimmt. Jeder Eintrag oberhalb der Hauptdiagonalen besteht aus zwei Werten,  $c_{i,j}^{co}$  und  $c_{i,j}^{ad}$ . Hier steht  $c^{co}$  für Gleichkanal-Interferenz und  $c_{i,j}^{ad}$  für Nachbarkanal-Interferenz. Es gilt

$$c_{i,j}^{co} = \frac{y_{i,j}}{y_{i,i}} + \frac{y_{j,i}}{y_{j,j}} \quad \text{und} \quad c_{i,j}^{ad} = \frac{a_{i,j}}{y_{i,i}} + \frac{a_{j,i}}{y_{j,j}}.$$

Nehmen wir für das Beispiel von Abbildung 4.1 die Werte  $y_{j,i} = 10$  und  $a_{j,i} = 5$  an, so erhalten wir die Einträge

$$c_{i,j}^{co} = \frac{26}{131} + \frac{10}{113} \approx 0,29 \quad \text{und} \quad c_{i,j}^{ad} = \frac{9}{131} + \frac{10}{113} \approx 0,11.$$

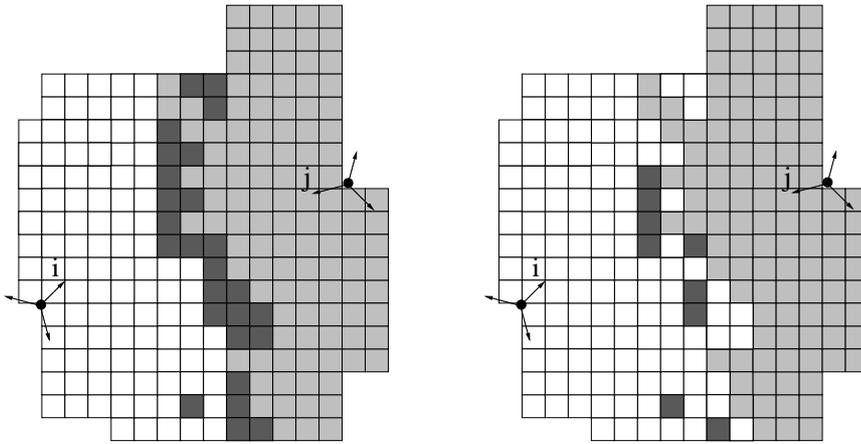


Abbildung 4.1: Überlappende Sendebereiche zweier Antennen.

Abgesehen von den Werten für Gleichkanal- und Nachbarkanal-Interferenz müssen Separationsbedingungen berücksichtigt werden. Damit ist gemeint, dass Sender, die auf derselben Antenne oder am selben Standort installiert wurden, aus technischen Gründen unterschiedliche Funkkanäle benutzen müssen. In unseren Szenarien fordern wir, dass die Funkkanäle von Sendern desselben Standortes (*co-site separation*) um mindestens zwei, von Sendern derselben Zelle (*co-cell separation*) um mindestens drei auseinanderliegen.

Diese Separierungsbedingungen werden in der Interferenz-Matrix unterhalb der Hauptdiagonale notiert. Muss ein Antennenpaar  $i, j$  einer solchen Bedingung genügen, so werden in der Interferenz-Matrix die entsprechenden Werte der Gleichkanal- und Nachbarkanal-Interferenz mit  $\infty$  notiert.

Weitere Separierungsbedingungen kommen hinzu, wenn zwischen Sendern *handover* möglich ist. Das bedeutet, dass die Zellen benachbart sind, und ein Nutzer von einer Zelle in die andere wechseln kann. Der Sender, mit dem der Kunde kommuniziert, kann sich dabei im Laufe einer bestehenden Verbindung ändern, ohne dass der Kunde davon etwas merkt. Ein solcher Transfer ist natürlich nur möglich, wenn ein Bereich von ausreichender Ausdehnung existiert, wo beide Sender eine ausreichende Empfangsleistung gewährleisten. Unser Kriterium dafür, dass zwischen zwei Sendern *handover* möglich ist und daher eine *handover*-Separation notwendig, ist dass ihr Signal auf einer Fläche von 10000 Quadratmetern, wo einer der beiden Sender *best server* ist, eine Differenz von 5dB nicht überschreitet. Ein ähnliches Kriterium wurde in [13] verwendet.

Nach der Ermittlung der Interferenz-Matrix erfolgt die Frequenzzuweisung. Die Aufgabe lautet, die Frequenzen so zuzuweisen, dass insgesamt wenig Störung auftritt. Sei also  $y : V \rightarrow C$  die Zuordnungsfunktion von

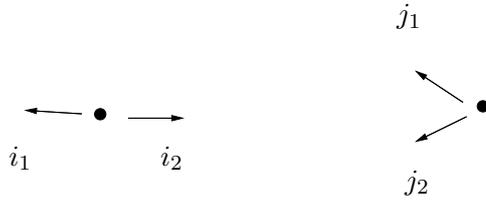


Abbildung 4.2: Ein sehr kleines Funknetz.

Kanälen  $c \in C$  zu Sendern  $v \in V$ , so lautet die zu minimierende Zielfunktion

$$\sum_{\substack{v,w \in V: \\ y(v)=y(w)}} c_{i,j}^{co} + \sum_{\substack{v,w \in V: \\ |y(v)-y(w)|=1}} c_{i,j}^{ad}. \tag{4.1}$$

In Abbildung 4.2 ist ein Funknetz mit vier Antennen an zwei Standorten zu sehen. Aus der Interferenzmatrix in Abbildung 4.3 gehen Separationsbedingungen und Kopplungswerte hervor. Antenne  $j_2$  hat zwei TRX, alle anderen Antennen eines. Die Funkkanäle der Sender  $j_{2,A}$  und  $j_{2,B}$  müssen sich um mindestens drei unterscheiden. Die Antennen  $i_1$  und  $i_2$  sowie  $j_1$  und  $j_2$  befinden sich jeweils am selben Standort. Sie müssen Separationsbedingungen von zwei erfüllen. Zwischen Antenne  $i_2$  und Antenne  $j_2$  besteht eine *handover*-Beziehung, daher die Separationsbedingung von zwei.

Die Antennen  $i_1$  und  $j_2$  sind nicht gekoppelt. Den dortigen Sendern kann man also denselben Funkkanal zuweisen, ohne dass dem Modell zufolge Interferenz entsteht. Anders ist das beispielsweise bei den Antennen  $i_2$  und  $j_1$ . Dieses Paar weist mit 0,35 einen relativ hohen Gleichkanal-Interferenzwert auf. Ihren Sendern sollten nach Möglichkeit unterschiedliche Frequenzen zugewiesen werden.

### 4.1.3 Thesen zur Planung interferenzarmer GSM-Netze

In diesem Unterabschnitt stellen wir Thesen zur Planung interferenzarmer GSM-Netze auf. Sie basieren auf den im vorigen Unterabschnitt definierten Kopplungswerten. Aus diesen Thesen leiten wir am Ende des Unterabschnittes unseren Optimierungsansatz ab. Die mathematische Formulierung dieses Problems wird in Abschnitt 4.2 erfolgen.

Der Optimierungsansatz zielt darauf ab, eine Netzkonfiguration mit einer möglichst günstigen Interferenz-Matrix zu ermitteln. Das FAP ist ein kompliziertes Problem, und so ist es nicht einfach zu beurteilen, wann eine Interferenz-Matrix „günstig“ ist, wann sie also einen Frequenzplan mit wenig Interferenz zulässt. Unser Planungsansatz gründet sich auf die folgenden drei Thesen:

**These 4.1.** *Je kleiner die Einträge der Interferenz-Matrix sind, desto besser.*

	$i_1$	$i_2$	$j_1$	$j_{2,A}$	$j_{2,B}$
$i_1$		$\infty$	0,01 0,00	0,01 0,00	
$i_2$	2		0,35 0,15	$\infty$	$\infty$
$j_1$				$\infty$	$\infty$
$j_{2,A}$		2	2		$\infty$
$j_{2,B}$		2	2	3	

Abbildung 4.3: Interferenzmatrix für das Netz in Abbildung 4.2.

Wir sagen, zwei Antennen sind nicht gekoppelt, wenn die gegenseitigen Gleichkanal-Interferenzwerte einen bestimmten Schwellwert – beispielsweise 0,01 – nicht überschreiten. Für ein gegebenes Netz können wir damit für jede Antenne  $i$  einen *Kopplungsgrad* definieren. Er entspricht der Zahl

$$|\{k \in I \setminus i \mid c_{ik}^{co} \geq 0,01 \vee c_{ki}^{co} \geq 0,01\}|.$$

**These 4.2.** *Je kleiner die Kopplungsgrade der Antennen sind, desto besser.*

**These 4.3.** *Es genügt, bei der Optimierung die Kopplungswerte zu verwenden. Eine Betrachtung der Größen der best server-Bereiche und damit der Werte von Gleichkanal- und Nachbarkanal-Interferenz ist nicht notwendig.*

Der Unterschied in der Betrachtung von Kopplungswerten und Interferenzwerten liegt darin, dass bei den Kopplungswerten die Größen der *best server*-Bereiche nicht eingehen. Für verschiedene Netzkonfigurationen sind diese Größen unterschiedlich, und in der Regel ist ihre Ermittlung aufwändig. Aus der Definition der Interferenzwerte  $c^{co}$  und  $c^{ad}$  geht hervor, dass diese in nichtlinearer Weise von der Netzkonfiguration abhängen. Da wir in dieser Arbeit das Ziel verfolgen, eine lineare Formulierung des Optimierungsmodells zu entwickeln, sind die Interferenzwerte also ungeeignet.

Die letzte mag die strittigste der drei Thesen sein. Gerade bei kleinen Zellen spielt die Größe eine wichtige Rolle für die eigentlich interessanten Interferenzwerte. Ein Unterschied in der Zellgröße von wenigen Pixeln kann hier zu großen Unterschieden der Interferenzwerte führen. In den Kopplungswerten führen diese Veränderungen zu geringen Unterschieden. Die in dieser Arbeit ermittelten Ergebnisse (siehe Abschnitt 4.6) stützen jedoch These 4.3. Die Kopplungswerte eignen sich zur Ermittlung interferenzarmer GSM-Netze.

Unser Planungsziel lautet: Finde ein gültiges Netz  $N$ , sodass die Summe der Nebendiagonaleinträge der Kopplungsmatrix minimal ist.

## 4.2 Das pixelbasierte Modell

Nachdem wir das Optimierungsproblem in Abschnitt 4.1 formuliert haben, leiten wir hier eine erste mathematische Formulierung her. Sie enthält alle Informationen darüber, welche Antenne durch welche andere Antenne auf welchem Pixel gestört wird. Daher nennen wir das Modell auch *Pixel-Modell*. Wir werden das Modell linearisieren und erste Eindrücke über die Größe erhalten. Das Pixel-Modell enthält so viele Informationen, dass die resultierenden IPs zu groß sind. Wir werden es in Abschnitt 4.3 durch das kompakter formulierte Obermengen-Modell ersetzen.

In Unterabschnitt 4.2.1 führen wir die dem Modell zugrundeliegenden Mengen ein. Daraus leiten wir eine erste, nichtlineare Problemformulierung ab. Diese Formulierung lässt sich mit einfachen Methoden deutlich verbessern. Dies tun wir in Unterabschnitt 4.2.2. Hier verwenden wir unter anderem die in Abschnitt 3.1 dargestellten Methoden zur besseren Formulierung von Abdeckungsmatrizen. Die Zahl der verwendeten Terme wird durch die Verbesserung der Formulierung deutlich verringert. In Unterabschnitt 4.2.3 geben wir eine erste Linearisierung des nichtlinearen Problems an. Die Summanden der Zielfunktion werden hierbei in Ungleichungen und Hilfsvariablen überführt. Wir werden sehen, dass die Größe der so erhaltenen IPs beträchtlich ist. Der uns zur Verfügung stehende IP-Löser CPLEX konnte sie nicht in befriedigender Weise bearbeiten. Daher versuchen wir in Unterabschnitt 4.2.4, die lineare Problemformulierung mit Schnittungleichungen zu verbessern.

### 4.2.1 Erste Formulierung

In diesem Unterabschnitt führen wir zunächst die unseren Modellen zugrundeliegenden Bezeichnungen ein. Diese verwenden wir sowohl im hier vorgestellten Pixel-Modell als auch im später erläuterten Obermengen-Modell. Danach geben wir eine erste mathematische Formulierung des Optimierungsproblems aus Abschnitt 4.1 an. Diese wird in den nachfolgenden Unterabschnitten sukzessive verbessert.

Gegeben sei ein Parameter  $c$ , der *cutoff*. Er beträgt in unseren Szenarien  $-125$  dB. Gilt für den Dämpfungsfaktor  $\gamma_{ip} \geq c$ , so gehört Pixel  $p$  zu Zelle  $i$ . Gegeben sei außerdem der in Abschnitt 4.1 eingeführte  $\delta$ -Schwellwert.

- $P$  : Die Menge der betrachteten Pixel.
- $I$  : Die Menge der Installationen. Als *Installation* bezeichnen wir eine Antenne zusammen mit einer möglichen Ausrichtung. Eine Antenne wird also durch verschiedene Installationen repräsentiert. Die Wahl

einer Installation für ein Netz bedeutet die Festlegung auf eine Ausrichtung der zugehörigen Antenne.

- $\mathcal{S}$  : Eine Partition von  $I$ . Die Elemente dieser Partition heißen *Sektoren*. Die Menge der Elemente eines Sektors heißt *Sektormenge*. Eine Sektormenge enthält die möglichen Ausrichtungen einer Antenne. Wir definieren eine Abbildung  $S : I \rightarrow \mathcal{S}$ , wobei  $S(i)$  die Sektormenge ist, die Installation  $i$  enthält.
- $I_p$  : Die Menge der Installationen, deren Dämpfungsfaktor  $\gamma$  auf Pixel  $p$  größer oder gleich einem *cutoff*  $c$  ist.

$$I_p := \{i \in I \mid \gamma_{i,p} \geq c\}$$

- $P^i$  : Die Menge der Pixel  $p \in P$  mit  $i \in I_p$ .
- $Z^u(i, p)$  : Die Menge der Installationen, die Installation  $i$  auf Pixel  $p$  stören, für die also gilt

$$Z^u(i, p) := \{j \in I \setminus S(i) \mid c \leq \gamma_{i,p}, 1 \leq \frac{\gamma_{i,p}}{\gamma_{j,p}} \leq \delta\}.$$

Man beachte, dass für die Sektormenge  $S(i)$  gilt  $Z^u(i, p) \cap S(i) = \emptyset$ . Die Menge  $Z^u(i, p)$  heißt auch *Untermenge* der Installation  $i$  auf Pixel  $p$ .

- $Z^o(i, p)$  : Die Menge der Installationen, deren Empfangsstärke auf Pixel  $p \in P^i$  größer ist als die von Installation  $i$ .  $Z^o(i, p)$  ist nur definiert für Pixel, die in  $P^i$  liegen.

$$Z^o(i, p) := \{j \in I_p \setminus i \mid c \leq \gamma_{i,p} < \gamma_{j,p}\}$$

Wir bezeichnen  $Z^o(i, p)$  als *Obermenge* von Installation  $i$  auf Pixel  $p$ .

Betrachten wir zur Veranschaulichung Abbildung 4.4. Dargestellt sind die Pfadverluste der Installationen  $i_1, \dots, i_6$  auf einem Pixel  $p$ . Wir können erkennen, dass Installation  $i_1$  mit 60 dB den geringsten Pfadverlust hat. Der  $\delta$ -Schwellwert ist die Differenz der Pfadverluste, bis zu der ein anderer Sender noch als Störer gewertet wird. Er lautet hier 21 dB. Da Installation  $i_2$  mit 105 dB den zweitbesten Pfadverlustwert aufweist, hat Installation  $i_1$  auf  $p$  keinen Störer. Das ist bei Installation  $i_2$  anders. Sie wird durch Installation  $i_3$  gestört, denn deren Pfadverlust ist nur 112 dB. Installationen  $i_4, i_5$  und  $i_6$  werden auf  $p$  jedoch so schwach empfangen, dass sie nicht mehr als Störer von  $i_2$  gelten. Es gilt  $Z^o(i_2, p) = \{i_1\}$  und  $Z^u(i_2, p) = \{i_3\}$ .

Nur die Installationen  $i_1, i_2$  und  $i_3$  haben einen Pfadverlust, der kleiner ist als der (negative) *cutoff* von  $-125$  dB. Pixel  $p$  liegt also in den Zellen von  $i_1, i_2$  und  $i_3$ . Er liegt in den Störungsbereichen der Installationen  $i_4$

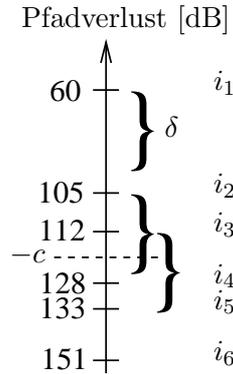


Abbildung 4.4: Pfadverlustwerte auf einem Pixel.

und  $i_5$ . Installation  $i_6$  ist so schwach zu empfangen, dass  $p$  außerhalb des Störungsbereichs von Antenne  $i_6$  liegt.

Die Entscheidung, ob eine Installation  $i$  ausgewählt ist oder nicht, wird durch die Variable

$$z_i \in \{0, 1\} \quad \forall i \in I$$

wiedergegeben. Es gilt  $z_i = 1$ , wenn  $i$  ausgewählt ist. Ein Netz  $N \subset I$  wird also repräsentiert durch einen Inzidenzvektor  $z = \{z_1, \dots, z_{|I|}\}$ .

Wir fordern, dass alle Pixel der Menge  $P$  abgedeckt sind. Also muss gelten

$$\sum_{i \in I_p} z_i \geq 1 \quad \forall p \in P. \quad (4.2)$$

Eine solche Ungleichung bezeichnen wir als *Abdeckungsbedingung*. Aus jeder Sektormenge wählen wir genau eine Installation aus. Also fordern wir

$$\sum_{i \in S} z_i = 1 \quad \forall S \in \mathcal{S} \quad (4.3)$$

Eine solche Gleichung bezeichnen wir als *Partitionsbedingung* (oder *Sektorgleichung*). Wir bezeichnen ein Netz als *gültig*, wenn sein Inzidenzvektor die Abdeckungs- und die Partitionsbedingungen erfüllt.

Die Störung beschreiben wir zunächst in nichtlinearer Weise mit Termen, die in die Zielfunktion eingehen. Eine Installation  $j$  stört eine Installation  $i$  auf Pixel  $p$  genau dann, wenn  $j \in Z^u(i, p)$  und  $z_k = 0$  für alle  $k \in Z^o(i, p)$ . Dieser Sachverhalt läßt sich durch das Polynom

$$z_i \cdot z_j \cdot \prod_{k \in Z^o(i, p)} (1 - z_k)$$

beschreiben. Der Wert des Polynoms ist für ganzzahlige Werte eins genau dann, wenn die oben beschriebenen Forderungen zutreffen. Ansonsten hat

das Polynom den Wert null. Es drückt eine *Kopplungsbeziehung* der Installationen  $i$  und  $j$  aus. Wir können die Werte der Kopplungsmatrix  $Y$  mit Hilfe dieser Terme ausdrücken. Zunächst definieren wir

$$A(i, j) := \{p \in P \mid i \in I_p \wedge j \in Z^u(i, p)\}.$$

Den Wert der Kopplungsvariablen  $y_{i,j}$  aus der Kopplungsmatrix  $Y$  erhalten wir durch den Ausdruck

$$y_{i,j} = \sum_{p \in A(i,j)} z_i \cdot z_j \cdot \prod_{k \in Z^o(i,p)} (1 - z_k)$$

Unsere erste, nichtlineare Formulierung des Optimierungsproblems lautet

$$\min \sum_{p \in P} \sum_{i \in I_p} \sum_{j \in Z^u(i,p)} z_i \cdot z_j \cdot \prod_{k \in Z^o(i,p)} (1 - z_k)$$

sodass

$$\begin{aligned} \sum_{i \in I_p} z_i &\geq 1 \quad \forall p \in P \\ \sum_{i \in S} z_i &= 1 \quad \forall S \in \mathcal{S} \\ z &\in \{0, 1\}^n. \end{aligned} \tag{4.4}$$

#### 4.2.2 Verbesserung der Formulierung

Es gibt zahlreiche Möglichkeiten, Formulierung 4.4 der Optimierungsaufgabe zu verbessern. Wir zeigen in diesem Unterabschnitt zunächst, wie die Zahl der Ungleichungen und der Variablen reduziert werden kann. Dazu verwenden wir die in Abschnitt 3.1 eingeführten Methoden zur besseren Formulierung von Abdeckungs- und Partitionsmatrizen.

Anschließend untersuchen wir, wie man die Zielfunktion kompakter ausdrücken kann, indem man Terme weglässt, deren Wert durch die Abdeckungs- und Partitionsbedingungen von vornherein feststeht. Wir zeigen, dass die Zahl der möglichen *best server* auf einem Pixel deutlich geringer ist, als die Zahl der Zellen, in der der Pixel liegt. Dadurch entfallen eine große Zahl von Kopplungsbeziehungen, die die erste Formulierung von NOPT noch berücksichtigte.

Zunächst wenden wir Beobachtung 3.4 auf Seite 17 auf das Problem an.

**Beobachtung 4.1.** *Sei  $p \in P$ . Existiert ein  $p' \in P$ , sodass  $I_{p'} \subset I_p$ , so ist die zu  $p$  gehörige Abdeckungsungleichung redundant.*

Das liegt daran, dass die zu  $p$  gehörige Ungleichung die zu  $p'$  gehörige dominiert. Diese Beobachtung entspricht dem ersten Fall von Beobachtung 3.4.

**Beobachtung 4.2.** Sei  $p \in P$ . Existiert ein  $S \in \mathcal{S}$ , sodass gilt  $S \subset I_p$ , so ist die zu  $p$  gehörige Abdeckungsungleichung redundant.

Das ist klar, wenn man die zu  $S$  gehörige Gleichung in die zu  $p$  gehörige Ungleichung substituiert. Diese Beobachtung entspricht dem zweiten Fall von Beobachtung 3.4.

**Beobachtung 4.3.** Existiert ein  $p \in P$ , sodass  $I_p \subsetneq S$  für ein  $S \in \mathcal{S}$ , so kann die Gleichung

$$\sum_{i \in S} z_i = 1 \quad \forall S \in \mathcal{S} \quad (4.5)$$

durch die Gleichungen

$$\sum_{i \in I_p} z_i = 1 \quad (4.6)$$

$$z_i = 0 \quad \forall i \in S \setminus I_p \quad (4.7)$$

ersetzt werden.

Dieses Resultat erhält man durch Substitution der zu  $S$  gehörigen Gleichung in die zu  $p$  gehörige Ungleichung. Es handelt sich dabei um eine Anwendung des zweiten Teils des dritten Falles in Beobachtung 3.4.

Beobachtung 4.3 erlaubt es uns, die Menge der betrachteten Installationen zu reduzieren. Manche Installationen können also nicht gewählt werden, weil die Abdeckung in manchen Gebieten ansonsten nicht gewährleistet werden kann. Ein Sonderfall, der in unseren Szenarien oftmals vorkam, ist der, dass ein Pixel nur von einer einzigen Installation abgedeckt ist. Dann ist diese Installation *notwendig*, und wir können alle anderen Installationen aus der zugehörigen Sektormenge entfernen. Bei den weiteren Betrachtungen gehen wir davon aus, dass unsere Installationsmenge so gewählt ist, dass auf die oben beschriebene triviale Weise keine Installationen aus  $I$  entfernt werden können. Man beachte, dass es durchaus möglich ist, aus den Abdeckungsungleichungen weitere Reduzierungen der Menge  $I$  abzuleiten. Ein Beispiel sei hier gegeben.

**Beobachtung 4.4.** Seien  $j, k$  zwei Installationen derselben Sektormenge und  $i$  eine Installation in einer anderen Sektormenge. Enthält die Abdeckungsmatrix die Ungleichungen

$$z_i + z_j \geq 1 \quad (4.8)$$

$$z_i + z_k \geq 1, \quad (4.9)$$

so ist  $i$  eine notwendige Installation.

Dieses einfache Beispiel für nichttriviale Notwendigkeitsbedingungen soll an dieser Stelle genügen. Es handelt sich hierbei um eine Anwendung von Beobachtung 3.5. Die Menge  $U$  ist  $\{j, k\}$ , die Menge  $J$  ist die Menge der dargestellten Zeilen. Die Matrix  $A_{JU}$  hat die Form

$$\begin{array}{cc} & \begin{array}{c} j \quad k \end{array} \\ \begin{array}{c} 1 \\ 0 \end{array} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array}$$

und das System 3.2 hat in diesem Beispiel den Wert 2. Da die Menge  $I'$  nur aus  $i$  besteht, folgt  $a_i \geq 1$ .

Wenden wir uns der Zielfunktion zu. Sie besteht in unseren Szenarien aus Millionen von Summanden. Daher liegt es in unserem Interesse, ihre Zahl zu verringern. Die Summanden haben die Werte eins oder null.

**Beobachtung 4.5.** *Der Term*

$$z_i \cdot z_j \cdot \prod_{k \in Z^o(i,p)} (1 - z_k)$$

*hat den Abdeckungs- und Partitionsbedingungen zufolge den Wert eins, wenn*

- *die Installationen  $i$  und  $j$  notwendig sind und*
- *alle Installationen aus  $Z^o(i,p)$  nicht wählbar sind.*

Die letzte Bedingung ist in der Regel nur durch eine eingehende Analyse der Abdeckungsmatrix überprüfbar, wie Beobachtung 4.4 deutlich macht. Interessanter sind Terme, die für jede gültige Lösung den Wert null annehmen.

**Beobachtung 4.6.** *Der Term*

$$z_i \cdot z_j \cdot \prod_{k \in Z^o(i,p)} (1 - z_k)$$

*nimmt für jede gültige Lösung des Systems 4.16 den Wert null an, wenn gilt:*

- *Eine der Variablen  $z_i, z_j$  ist durch die Abdeckungsungleichungen implizit auf den Wert null fixiert, oder*
- *es existiert eine Sektormenge  $S \subset \mathcal{S}$ , sodass  $S \subset Z^o(i,p)$ , oder*
- *es existiert ein Pixel  $\bar{p}$  mit  $I_{\bar{p}} \subset Z^o(i,p)$ .*

In den beiden letzten Fällen existiert in der Menge  $Z^o(i, p)$  zwangsläufig ein Index  $k$ , dessen Variable  $z_k$  den Wert eins annimmt. Das bedeutet nichts anderes, als dass Installation  $i$  auf Pixel  $p$  niemals *best server* sein kann. Wir bezeichnen die Menge der Installationen, die auf Pixel  $p$  *best server* sein können, als  $\bar{I}_p$ .

Durch Weglassen derjenigen Summanden in der Zielfunktion, die zu Installationen aus  $I_p \setminus \bar{I}_p$  gehören, wird in der Regel eine deutliche Reduzierung der Zahl aller Summanden erreicht. So konnte beispielsweise die Summe der Größen aller  $I_p$ -Mengen auf einer der untersuchten Instanzen von etwa 875000 auf 106000 gesenkt werden.

Die verbesserte Formulierung unseres Optimierungsproblems lautet

$$\min \sum_{p \in P} \sum_{i \in \bar{I}_p} \sum_{j \in Z^u(i, p)} z_i \cdot z_j \cdot \prod_{k \in Z^o(i, p)} (1 - z_k) \quad (4.10)$$

sodass

$$\begin{aligned} \sum_{i \in \bar{I}_p} z_i &\geq 1 \quad \forall p \in \bar{P} \\ \sum_{i \in S} z_i &= 1 \quad \forall S \in \mathcal{S} \\ z &\in \{0, 1\}^n. \end{aligned} \quad (4.11)$$

Wir bezeichnen dieses Optimierungsproblem als NOPT. Es unterscheidet sich äußerlich nicht sehr von der ersten Problemformulierung. Der Unterschied besteht darin, dass die Menge der Installationen und die Menge der Abdeckungsungleichungen sich verändert haben. Außerdem werden auf jedem Pixel  $p$  nur die dort infrage kommenden Installationen  $\bar{I}_p$  betrachtet. Die Pixelmenge  $\bar{P}$  besteht aus den Pixeln,

- die nicht bei jeder beliebigen Auswahl von Installationen abgedeckt sind,
- und deren Abdeckungsbedingungen nicht durch die Bedingungen auf einem anderen Pixel redundant geworden sind.

Wir bezeichnen die Menge  $\bar{P}$  auch als Menge der *interessanten* Pixel.

### 4.2.3 Linearisierung

Die Linearisierung von System 4.11 erfolgt durch Einführung von Ungleichungen und Variablen für die Summanden der nichtlinearen Zielfunktion 4.10. Wir benutzen die Methode von Fortet [9], die wir in Abschnitt 3.2 vorgestellt haben.

Sei

$$z_i \cdot z_j \cdot \prod_{k \in Z} (1 - z_k) = z_i \cdot z_j \cdot \prod_{k \in Z} \bar{z}_k \quad (4.12)$$

einer der Summanden aus der Zielfunktion, wobei  $Z = Z^o(i, p)$  der Obermenge von Installation  $i$  auf Pixel  $p$  entspricht. Wir führen die Variable

$$y(i, j, p) \in \{0, 1\}$$

ein, die den Wert von Ausdruck 4.12 annehmen soll, wenn die  $z$ -Variablen Werte aus der Menge  $\{0, 1\}$  annehmen. Der Wert der Variable  $y(i, j, p)$  lässt sich durch lineare Nebenbedingungen gemäß Fortet [9] festlegen. Diese Bedingungen lauten

$$\begin{aligned} & (z_i + z_j - 2) + \left( \sum_{k \in Z} \bar{z}_k - |Z| \right) + 1 \\ = & z_i + z_j - \sum_{k \in Z} z_k - 1 \leq y(i, j, p) \end{aligned} \quad (4.13)$$

und

$$(|Z| + 2) \cdot y(i, j, p) \leq z_i + z_j + \sum_{k \in Z} \bar{z}_k \quad (4.14)$$

Ungleichung 4.13 entspricht dabei 3.10 und Ungleichung 4.14 entspricht 3.11. Ungleichung 4.14 wird in unserer IP-Formulierung nicht benötigt, da wir die Summe der  $y(i, j, p)$ -Variablen minimieren. Ungleichungen vom Typ 4.13 nennen wir in der Folge *Beziehungsungleichungen*, denn sie geben die Koppelungsbeziehungen zwischen den Installationen wieder.

Das resultierende IP hat die Form

$$\min \sum_{p \in P} \sum_{i \in \bar{I}_p} \sum_{j \in Z^u(i, p)} y(i, j, p)$$

sodass

$$\begin{aligned} z_i + z_j - \sum_{k \in Z} z_k - 1 &\leq y(i, j, p) && \forall i \in I, p \in A(i, j), j \in Z^u(i, p) \\ \sum_{i \in \bar{I}_p} z_i &\geq 1 && \forall p \in \bar{P} \\ \sum_{i \in S} z_i &= 1 && \forall S \in \mathcal{S} \\ y(i, j, p) &\in \{0, 1\} && \forall i \in I, p \in \bar{P}, j \in Z^u(i, p) \\ z &\in \{0, 1\}^n. \end{aligned} \quad (4.15)$$

#### 4.2.4 Schnittungleichungen

In diesem kurzen Unterabschnitt definieren wir zwei Klassen von Schnittungleichungen für System (4.16). Der Grund hierfür ist, dass wir die IPs, die sich für Instanzen durchschnittlicher Größe ergaben, nicht in befriedigender Weise lösen können. Der Wert der LP-Relaxierung von (4.16) weicht deutlich ab von den Werten der gefundenen Lösungen. Das Konvergenzverhalten

des IP-Lösers ist unbefriedigend. Wir formulieren die Schnittungleichungen mit zwei Zielen: Zum einen soll sich der Wert der LP-Lösung erhöhen und damit die Lücke zur IP-Lösung verringert werden. Zum anderen soll sich das Konvergenzverhalten des IP-Lösers verbessern.

Die Aussage der ersten Klasse von Ungleichungen lässt sich so formulieren: Auf Pixel  $p$  gibt es eine Zahl von Störern, die mindestens so groß ist wie die Zahl von Störern, die jeder mögliche *best server* auf  $p$  in jedem Fall hat.

**Beobachtung 4.7.** Sei  $s(i, p)$  die Mindestzahl der Störer von Installation  $i$  auf Pixel  $p$ . Sei

$$s(p) := \min \{s(i, p) \mid i \in \bar{I}_p\},$$

die Mindestzahl an Störern, die jeder beliebige *best server* auf Pixel  $p$  hat. So ist

$$\sum_{i \in \bar{I}_p} \sum_{j \in Z^u(i, p)} y(i, j, p) \geq s(p) \quad (4.16)$$

eine gültige Ungleichung für System (4.16).

Die zweite Klasse von Ungleichungen beschreibt die Zahl von Störern, die die Installation  $i$  auf Pixel  $p$  mindestens hat, wenn es dort keine Installation mit geringerem Pfadverlust gibt.

**Beobachtung 4.8.** Sei  $i$  eine Installation und  $p$  ein Pixel. Es gelte

$$\gamma_{ip} \geq \gamma_{jp} \quad \forall j \in \bar{I}_p.$$

Dann ist

$$\sum_{j \in Z^u(i, p)} y(i, j, p) \geq s(i, p) \cdot z_i \quad (4.17)$$

eine gültige Ungleichung für System (4.16).

Die Implementierung dieser Schnittungleichungen führte nicht zu wesentlichen Fortschritten. Sie waren so gering, dass wir uns dem Obermengen-Modell zuwendeten, welches eine kompaktere lineare Formulierung von System 4.16 darstellt.

### 4.3 Das Obermengen-Modell

In Abschnitt 4.2 stellten wir eine Methode vor, das in Abschnitt 4.1 hergeleitete Optimierungsproblem mathematisch zu formulieren. Wir gaben zunächst eine nichtlineare Formulierung an, die wir dann linearisierten. Das entstandene Modell bezeichneten wir als Pixel-Modell, weil die Kopplungsverhältnisse der Netzkonfigurationen in Variablen festgehalten wurde, die auf Pixelebene definiert waren. In diesem Abschnitt führen wir das Obermengen-Modell ein. Es ist so bezeichnet, weil die Kopplungsvariablen davon

abhängen, ob ein *server* in der *Obermenge* einer Installation aktiv ist oder nicht.

Das Modell ist dem Pixel-Modell aus Abschnitt 4.2 relativ ähnlich. Es entsteht aus NOPT durch Linearisierung, nachdem man Zielfunktionskoeffizienten zusammengefasst hat. Diese Linearisierung führen wir in Unterabschnitt 4.3.1 durch. Dort zeigen wir auch, in welchem Ausmaß sich die Zahl von Ungleichungen und Hilfsvariablen gegenüber dem Pixel-Modell reduziert. Außerdem zeigen wir, dass das sich ergebende Optimierungsproblem OPT selbst ohne Abdeckungsbedingungen  $\mathcal{NP}$ -schwer ist, denn es existiert eine Reduktion des MAX-CUT Problems auf OPT.

Ein Nachteil der neuen Formulierung ist, dass wir die in Unterabschnitt 4.2.4 für das Pixel-Modell entwickelten Schnittungleichungen nicht mehr verwenden können. Nachdem wir das Obermengen-Modell auf einigen Instanzen getestet hatten, stellten wir fest, dass die Laufzeiten des IP-Lösers auch mit der neuen Formulierung zu groß waren. Wir versuchten daher wie im Pixel-Modell, die Formulierung durch Schnittungleichungen zu verschärfen. In Unterabschnitt 4.3.2 werden wir diese Schnittungleichungen herleiten. Für diese Herleitung transformieren wir das linearisierte Problem in ein Minimierungsproblem auf Graphen. Damit erhalten wir auch eine neue Sichtweise auf unsere Aufgabe. Das Optimierungsproblem lässt sich nach der Transformation als ein MAX-CUT Problem mit positiven Kantengewichten interpretieren, wobei das Gewicht der Schnittkanten auf komplexe Weise von der Wahl des Schnittes abhängt. Die graphische Sichtweise erleichtert uns die Definition einiger Klassen von Schnittungleichungen.

Die in diesem Abschnitt eingeführte verbesserte Problemformulierung und die anschließende Präzisierung durch Schnittungleichungen führt uns nicht zu Programmen, die wir in angemessener Zeit lösen können. Wir werden daher im nachfolgenden Abschnitt 4.4 Vereinfachungen des Obermengen-Modells durchführen, um die Lösungsgeschwindigkeit zu erhöhen.

### 4.3.1 Linearisierung

In diesem Unterabschnitt möchten wir das Obermengen-Modell analog zum Pixel-Modell in Abschnitt 4.2.3 linearisieren. Zunächst führen wir vor, wie sich die Zielfunktion von NOPT durch Zusammenfassen gleicher Koeffizienten kompakter darstellen lässt. Das nach der Linearisierung entstehende IP bezeichnen wir mit OPT. Wir zeigen, wie sich die Zahl der Beziehungsgleichungen durch die Einführung des neuen Modells reduziert. Schließlich zeigen wir, dass OPT  $\mathcal{NP}$ -schwer ist, selbst wenn es keine Abdeckungsbedingungen enthält. Wir tun das, indem wir das MAX-CUT Problem mit positiven Kantengewichten auf OPT zurückführen.

Das Obermengen-Modell erhalten wir, wenn wir in der Zielfunktion von System 4.16 gleichartige Summanden zusammenfassen. Summanden der Form

$$z_i \cdot z_j \cdot \prod_{k \in Z^o(i,p)} (1 - z_k)$$

können durch Störungsverhältnisse auf vielen Pixeln entstehen. Besonders häufig kommen Ausdrücke derselben Klasse vor, wenn  $Z^o(i, p)$  die leere Menge ist. Sei

$$\begin{aligned} A(i, j, Z) &:= \{p \in P^i \mid Z^o(i, p) = Z, j \in Z^u(i, p)\} \subset A(i, j), \\ c_{i,j}^Z &:= |A(i, j, Z)| \end{aligned} \quad (4.18)$$

und

$$Z(i) := \{Z^o(i, p) \mid p \in P\}.$$

Dann lautet die neue Formulierung des Optimierungsproblems

$$\min \sum_{i \in I} \sum_{Z \in Z(i)} \sum_{j \mid A(i,j,Z) \neq \emptyset} c_{i,j}^Z \cdot z_i \cdot z_j \cdot \prod_{k \in Z} (1 - z_k)$$

sodass

$$\begin{aligned} \sum_{i \in \bar{I}_p} z_i &\geq 1 \quad \forall p \in \bar{P} \\ \sum_{i \in S} z_i &= 1 \quad \forall S \in \mathcal{S} \\ z &\in \{0, 1\}^n. \end{aligned} \quad (4.19)$$

In der Praxis wurde die Zielfunktion damit auf einen Bruchteil der Summanden verringert (siehe auch Tabelle 4.2). Für die Koeffizienten der Kopplungsmatrix gilt mit den obigen Definitionen

$$\begin{aligned} y_{i,j} &= \sum_{p \in A(i,j)} z_i \cdot z_j \cdot \prod_{k \in Z^o(i,p)} (1 - z_k) \\ &= \sum_{Z \in Z(i)} \sum_{p \in A(i,j,Z)} z_i \cdot z_j \cdot \prod_{k \in Z} (1 - z_k) \\ &= \sum_{Z \in Z(i)} c_{i,j}^Z \cdot z_i \cdot z_j \cdot \prod_{k \in Z} (1 - z_k). \end{aligned} \quad (4.20)$$

Substituieren wir  $y(i, j, Z)$  mit diesem Wissen in Gleichung 4.20, so erhalten wir

$$y_{i,j} = \sum_{Z \in Z(i)} c_{i,j}^Z \cdot y(i, j, Z).$$

Die Linearisierung von (4.19) nach Fortet [9] liefert das folgende IP:

$$\min \sum_{i \in I} \sum_{Z \in Z(i)} \sum_{j \in I \setminus S(i)} c_{i,j}^Z \cdot y(i, j, Z)$$

Name	Standorte	Installationen	Pixel	Ungl.	
				Pixel-Modell	Obermengen-Modell
HG1/4	36	72	6241	317992	34002
HG3/1	60	120	6241	404956	53730
NET01/4	148	238	22199	1108160	67509
H01/3	193	313	22122	594216	93882

Tabelle 4.2: Größenvergleich der IPs des Pixel- und des Obermengen-Modells.

sodass

$$\begin{aligned}
z_i + z_j - \sum_{k \in Z} z_k - 1 &\leq y(i, j, Z) && \forall i \neq j \in I, Z \in Z(i) \\
\sum_{i \in \bar{I}_p} z_i &\geq 1 && \forall p \in \bar{P} \\
\sum_{i \in S} z_i &= 1 && \forall S \in \mathcal{S} \\
y(i, j, Z) &\in \{0, 1\} && \forall i \neq j \in I, Z \in Z(i) \\
z_i &\in \{0, 1\} && \forall i \in I.
\end{aligned} \tag{4.21}$$

Diesem Problem geben wir den Namen OPT. Man beachte, dass eine Beschränkung der  $y$ -Variablen nach unten genau so wie im Pixel-Modell nicht erfolgen muss, da wir minimieren, und die  $c_{ij}^Z$ -Werte positiv sind.

Tabelle 4.2 zeigt, wie sich die Größe des IPs durch Einführung der Obermengenvariablen verringert. Dargestellt sind Kennzahlen für vier Szenarien: die Zahl der Standorte, die Zahl der Installationen, die zur Wahl stehen, die Zahl der Pixel und die Zahl der Ungleichungen der beiden Modelle.

Wir sehen, dass die Zahl der Ungleichungen im Obermengen-Modell immer deutlich geringer ist als die im Pixel-Modell. Der Faktor liegt zwischen etwa fünf und fünfzehn. Das Pixel-Modell hat im Szenario NET01/4 bereits über eine Million Ungleichungen, und das, obwohl dieses Szenario mit 238 Installationen noch recht klein ist. Wir sehen, dass die Zahl der Ungleichungen auch im Obermengen-Modell noch beträchtlich ist. Wir wollen Instanzen mit etwa fünf- bis achthundert Installationen optimieren. Die einzige Instanz, die in etwa an diese Größe heranreicht, ist H01/3. Hier beträgt die Zahl der Ungleichungen schon fast hunderttausend.

OPT enthält ein *set covering*-Problem, und ist daher  $\mathcal{NP}$ -schwer (siehe [5]). Es kommt jedoch häufiger vor, dass die Netzabdeckung so gut ist, dass keine Abdeckungsungleichungen notwendig sind. Wir zeigen nun, dass unser Optimierungsproblem trotzdem  $\mathcal{NP}$ -schwer ist.

**Satz 4.1.** *Das Optimierungsproblem OPT ist  $\mathcal{NP}$ -schwer, wenn die Menge  $\bar{P}$  leer ist.*

**Beweis: (Satz 4.1)** Wir haben in Abschnitt 3.3 eine Formulierung des MAX-CUT Problems als IP hergeleitet. Sie sei hier noch einmal notiert:

$$\begin{aligned}
\min \quad & \sum_{e \in E} c(e) \cdot (y_e + y'_e) \\
x_u + x_v - 1 & \leq y(e) \quad \forall e = uv \in E \\
\bar{x}_u + \bar{x}_v - 1 & \leq y'(e) \quad \forall e = uv \in E \\
x_v + \bar{x}_v & = 1 \quad \forall v \in V \\
x_v, \bar{x}_v & \in \{0, 1\} \quad \forall v \in V \\
y_e, y'_e & \in \{0, 1\} \quad \forall e \in E
\end{aligned} \tag{4.22}$$

Dieses Problem stellt sich als Spezialfall von OPT dar, wobei alle Obermengen leer sind. Der Wert des MAX-CUT Problems ist als die Summe

$$\sum_{e \in E} c(e) - 1 \cdot \sum (y_e + y'_e)$$

aus der Optimallösung erhältlich.

■

Das Problem OPT ist kann also ein verallgemeinertes MAX-CUT Problem mit Nebenbedingungen und positiven Kantengewichten betrachtet werden, die auf komplexe Weise von der Wahl des Schnittes abhängen. Diese Abhängigkeit der Kantengewichte wird wiedergegeben durch die Beziehungsungleichungen, die die Summe  $\sum_{k \in Z} z_k$  enthalten.

Selbst wenn wir ein Problem betrachten, das nur leere Obermengen enthält, ist OPT eine Verallgemeinerung des MAX-CUT Problems, denn jeder der vier Zustände der beiden Knotenvariablen einer Kante ruft eine andere Bewertung der Kante hervor. Beim MAX-CUT Problem gibt es für Kanten nur zwei Zustände, nämlich geschnitten oder nicht geschnitten. Hier ist der Schnitt zusätzlich gerichtet, und es ist ein Unterschied, ob beide Endknoten in  $U$  sind oder keiner. Kommen Abdeckungsbedingungen zu OPT hinzu, so muss die gewählte Menge  $U$  diesen genügen. Auch dadurch wird eine Richtung in das MAX-CUT Problem induziert, denn ist hier die Menge  $U$  optimal, so ist es auch  $V \setminus U$ . Wegen der *set covering*-Bedingungen ist  $V \setminus U$  in OPT in der Regel keine gültige Lösung mehr.

Ein Problem von Formulierung 4.22 von OPT ist, dass die LP-Relaxierung sehr schlecht ist. Betrachten wir eine Instanz, die

- keine Abdeckungsbedingungen hat, und
- wo jede Sektormenge mindestens zwei Installationen enthält.

Für die LP-Relaxierung ist die Lösung

$$z_i = \frac{1}{|S(i)|} \quad \forall i \in I \quad (4.23)$$

$$y(i, j, Z) = 0 \quad \forall i, j, i \neq j \in I, Z \subset I \quad (4.24)$$

gültig und optimal. Sie hat den Wert null. Offensichtlich liegt das daran, dass die Beziehungsungleichungen nicht sehr streng formuliert sind. Die Summe der Werte von  $z_i$  und  $z_j$  muss größer als eins sein, damit die  $y(i, j, Z)$ -Variablen einen Wert annehmen, der größer ist als null. Es wäre wünschenswert, dass die  $y$ -Variablen schon bei kleineren  $z$ -Werten Kopplung anzeigen. Hierfür entwickeln wir im folgenden Unterabschnitt verschiedene Klassen von Schnittungleichungen.

### 4.3.2 Schnittungleichungen

In diesem Unterabschnitt befassen wir uns mit der Ermittlung mehrerer Klassen von Schnittungleichungen für OPT. Unsere Zielsetzung ist, die LP-Relaxierung des Problems und das Konvergenzverhalten des IP-Lösers zu verbessern. Bei der Bestimmung der Schnittebenen konzentrieren wir uns auf die Verbesserung der Formulierung der Beziehungsungleichungen. Wir haben im vorigen Unterabschnitt gesehen, dass die LP-Relaxierung so schlecht ist, weil die Beziehungsungleichungen in der Relaxierung erst bei relativ großen Werten der  $z$ -Variablen Kopplung anzeigen.

Bei unseren Untersuchungen bemerkten wir früh, dass ein Weglassen der Abdeckungsbedingungen für die von uns betrachteten Instanzen keine bedeutende Verbesserung des IP-Werts zur Folge hat. Daher vermuten wir, dass nicht viel zu erreichen wäre, wenn man versucht, die Formulierung des in OPT enthaltenen *set covering*-Problems zu verschärfen. Die *Hinzunahme* von Abdeckungsbedingungen hat vermutlich ebenso geringe Auswirkungen auf den Wert der LP-Relaxierung und das Konvergenzverhalten des IP-Lösers wie das *Weglassen*. Außerdem haben wir ebenfalls Instanzen betrachtet, die von vornherein keine Abdeckungsbedingungen enthalten und ebenfalls ein schlechtes Konvergenzverhalten aufweisen.

Zur Definition der Schnittungleichungen definieren wir zunächst Störungsgraphen. Ihre Kanten werden mit komplexen Störungsfunktionen bewertet. Durch die Störungsgraphen gewinnen wir eine neue Sichtweise auf OPT. Mit den Störungsgraphen leiten wir Schnittungleichungen her, die die Kopplung zwischen Installationen und Sektormengen, zwischen Paaren von Sektormengen und schließlich zwischen mehreren Sektormengen genauer beschreiben als die bisherigen Beziehungsungleichungen. Am Ende formulieren wir eine Klasse von Ungleichungen, die wir aus den in Abschnitt 3.1 eingeführten Schnitten für das *set covering* ableiten.

Unsere Bemühungen führten nicht zu den angestrebten Erfolgen. Der LP-Wert ändert sich nur geringfügig, eine Verbesserung des Konvergenzver-

haltens ist nicht feststellbar. In den nachfolgenden Abschnitten 4.4 und 4.5 widmen wir uns daher heuristischen Planungsansätzen.

### Graphische Ungleichungen

Wir führen hier den *Störungsgraphen* einer Menge von Installationen ein. An ihm kann man verschiedene Klassen von Schnittebenen direkt ablesen. Diese Schnittebenen bezeichnen wir deshalb als *graphische* Schnittebenen.

**Definition 4.1. (Störungsfunktion)** Sei  $I$  eine Menge von Installationen,  $\mathcal{S}$  eine Partition von  $I$  und  $i, j \in I$  mit  $S(i) \neq S(j)$ . Sei  $\mathcal{I} = \{N_0, N_1, \dots\}$  die Menge der gültigen Netze. Für jedes Netz  $N$  sei  $Y(N) = (y_{i,j})$  die dazugehörige Störungsmatrix. Dann heißt die Funktion

$$f_{ij} : \mathcal{I} \rightarrow \mathbb{Q}$$

$$f_{ij}(N) = \begin{cases} 0 & \text{falls } i \notin N \text{ oder } j \notin N, \\ y_{i,j} + y_{j,i} & \text{sonst} \end{cases} \quad (4.25)$$

die Störungsfunktion von  $\mathcal{I}$  bezüglich  $i$  und  $j$ . Die Reihenfolge von  $i$  und  $j$  spielt keine Rolle, es gilt  $f_{ji} = f_{ij}$ .

Die Störungsfunktion ist also für Paare von Installationen definiert. Sie misst die Störung, die ein Paar in einem gegebenen Netz insgesamt erzeugt. Ist eine der Installationen nicht im Netz enthalten, so ist der Wert von  $f$  trivialerweise null.

**Definition 4.2. (Störungsgraph)** Sei  $I$  eine Menge von Installationen,  $\mathcal{S}$  eine Partition von  $I$ . Sei  $G = (V, E)$  ein ungerichteter Graph, sodass  $V = I$  und

$$e = ij \in E \Leftrightarrow S(i) \neq S(j),$$

mit einer Gewichtsfunktion  $c : (E(G), \mathcal{I}) \rightarrow \mathbb{Q}^+$  mit

$$c(ij, N) := f_{ij}(N)$$

für alle  $ij \in E$ . Dann heißt  $(G, c)$  der Störungsgraph von  $I$ .

Ein Auszug aus einem Störungsgraphen ist in Abbildung 4.5 zu sehen. Mit dem Störungsgraphen können wir den Vorgang der Optimierung visuell darstellen. Ein gültiges Netz enthält immer genau einen der Knoten aus den eingekreisten Bereichen, welche die Sektoren repräsentieren. Wählen wir ein Netz  $N$ , so entspricht das der Wahl jeweils eines Knotens aus jedem eingekreisten Bereich. Um die Gewichtsfunktion  $c$  darzustellen, wird jede Kante  $e$  des Graphen mit den Werten  $f_e$  des Netzes  $N$  belegt. Der Wert des Netzes entspricht dann der Summe aller Kantengewichte des Graphen.

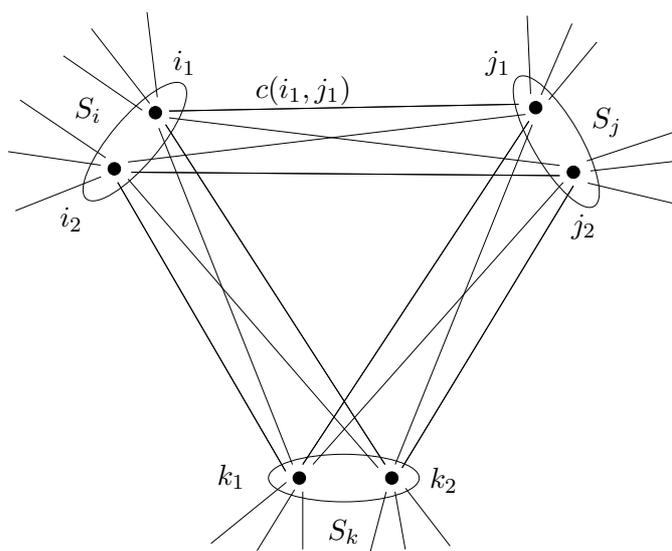


Abbildung 4.5: Auszug aus einem Störungsgraphen.

Wir haben damit eine neue Sichtweise auf unser Optimierungsproblem gewonnen. Die Aufgabe besteht darin, ein Netz so zu wählen, dass die Summe aller Werte der Gewichtsfunktion  $c$  minimal ist. Formell lautet die Aufgabe wie folgt:

**Problem 4.1.** Gegeben seien ein ungerichteter Graph  $G = (V, E)$ , eine Partition  $P = \{P_0, \dots, P_k\}$  der Knotenmenge  $V$  und eine Menge  $\mathcal{V} \subset 2^V$  von Teilmengen aus  $V$ . Eine Menge  $A \subset V$  mit

$$|A \cap P_i| = 1 \quad \forall i \in \{1, \dots, k\}$$

heiße Auswahl. Die Menge aller Auswahlen heiße  $\mathcal{A}$ . Eine Auswahl  $A \in \mathcal{A}$  heiße gültig, wenn außerdem gilt

$$|A \cap U| \geq 1 \quad \forall U \in \mathcal{V}.$$

Die Menge aller gültigen Auswahlen heiße  $\mathcal{A}' \subset \mathcal{A}$ . Gegeben sei weiterhin eine Gewichtsfunktion  $c : (E, \mathcal{A}) \rightarrow \mathbf{Q}^+$ .

Finde eine gültige Auswahl  $A \in \mathcal{A}'$ , sodass

$$\sum_{e \in E} c(e, A) = \min \left\{ \sum_{e \in E} c(e, B) \mid B \in \mathcal{A}' \right\}$$

Wir haben kaum Anhaltspunkte darüber, wie sich die Störungsfunktionen – und damit die Gewichtsfunktionen der Kanten – verhalten. Sind zwei Knoten  $i$  und  $j$  gewählt, so hängt das Gewicht der Kante  $ij$  von der Wahl weiterer Knoten ab.

Im Folgenden werden wir für die Gewichtsfunktion untere Grenzen zu ermitteln. Diese benutzen wir zur Ermittlung von Schnittgleichungen, die in das Problem OPT eingefügt werden. Unsere Idee ist, den Wert der LP-Relaxierung anzuheben und die Laufzeit des IP-Lösers zu verringern.

**Definition 4.3. (Minimaler Störwert)** Sei  $f_{ij}$  die Störungsfunktion der Installationen  $i$  und  $j$ . Der minimale Störwert  $M(i, j)$  ist wie folgt definiert:

$$M(i, j) := \min \{ f_{ij}(N) \mid N \in \mathcal{I} : i \in N, j \in N \}.$$

Der minimale Störwert ist nicht einfach zu ermitteln. Allerdings ist es einfach, ihn nach unten abzuschätzen:

**Lemma 4.1.** Seien  $i, j \in I$  zwei Installationen mit  $S(i) \neq S(j)$ . Dann gilt

$$M(i, j) \geq (c_{i,j}^\emptyset + c_{j,i}^\emptyset).$$

**Beweis: (Lemma 4.1)** Jeder Pixel in einer der Mengen  $A(i, j, \emptyset)$ ,  $A(j, i, \emptyset)$  trägt in jedem Netz  $N \in \mathcal{I}$ , das  $i$  und  $j$  enthält, zur Störung bei. Es gilt also

$$y_{i,j} \geq |A(i, j, \emptyset)| \cdot z_i \cdot z_j, \quad (4.26)$$

$$y_{j,i} \geq |A(j, i, \emptyset)| \cdot z_i \cdot z_j. \quad (4.27)$$

Daher gilt für jedes gültige Netz  $N$ , das  $i$  und  $j$  enthält:

$$f_{ij}(N) \geq (c_{i,j}^\emptyset + c_{j,i}^\emptyset).$$

Daraus folgt die Behauptung. ■

Wir bezeichnen die Abschätzung  $c_{i,j}^\emptyset + c_{j,i}^\emptyset$  des minimalen Störwerts mit  $m(i, j)$ :

$$m(i, j) := c_{i,j}^\emptyset + c_{j,i}^\emptyset$$

Auf die gleiche Weise kann man obere Schranken für die Kopplung finden. Dies ist für unser Problem aufgrund der Zielfunktion allerdings nicht interessant. Das Problematische an der von uns gefundenen Schranke ist, dass es keine lineare Nebenbedingung ist. Im nachfolgenden Abschnitt benutzen wir die Sektorgleichungen, um lineare Nebenbedingungen zu ermitteln.

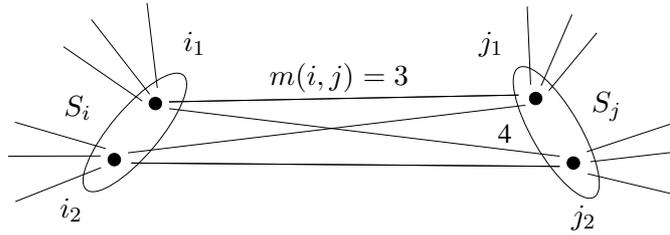


Abbildung 4.6: Kopplung zwischen einer Installation und einem Sektor.

### Kopplung zwischen Installationen und Sektormengen

Wir möchten nun den folgenden Sachverhalt als Schnittungleichung festhalten: Ist Installation  $i_1$  eingeschaltet, so ist der Betrag der Störung mit der Sektormenge  $S(j) = \{j_1, j_2, \dots\}$  mindestens so groß, wie das Minimum der minimalen Störwerte  $M(i_1, j_1), M(i_1, j_2), \dots$ .

Hierzu führen wir eine Variable ein, die die Störung einer Installation durch alle Installationen einer Sektormenge wiedergibt:

**Definition 4.4.** Sei  $i$  eine Installation und  $S \in \mathcal{S}$  eine Sektormenge. Die Variable  $y(i, S)$  sei folgendermaßen definiert:

$$y(i, S) := \sum_{j \in S} (y_{i,j} + y_{j,i})$$

Es ist klar, dass genau eine Installation aus  $S$  zur Störung beiträgt, falls alle  $z_k$  ganzzahlig sind. In der Relaxierung tragen jedoch mehrere Installationen  $j \in S$  zu den Kopplungswerten bei.

Betrachten wir das Beispiel in Abbildung 4.6. Für die Kopplung zwischen Installation  $i_1$  und dem Sektor  $S_j$  gelten die Abschätzungen

$$M(i_1, j_1) = 3, \quad M(i_1, j_2) = 4,$$

also gilt

$$y(i_1, S_j) \geq 3 \cdot z_i. \quad (4.28)$$

**Lemma 4.2.** Seien  $i, j \in I$  Installationen mit  $S(i) \neq S(j)$ . So ist

$$y(i, S(j)) \geq \min \{M(i, j) \mid j \in S(j)\} \cdot z_i. \quad (4.29)$$

eine gültige Ungleichung.

**Beweis: (Lemma 4.2)** Sei  $N$  ein gültiges Netz. Ist  $i \notin N$ , so gilt  $z_i = 0$  und Ungleichung 4.29 ist trivialerweise erfüllt. Sei also  $z_i = 1$ . Für genau ein  $j' \in S$  gilt  $z_{j'} = 1$ . Also folgt

$$y(i, S) = \sum_{j \in S} (y_{i,j} + y_{j,i}) = y_{i,j'} + y_{j',i} \geq M(i, j') \geq \min \{M(i, j) \mid j \in S\}.$$

■

Es ist uns nicht möglich, die minimalen Störwerte ohne großen Aufwand zu ermitteln. Wir können aber die oben definierte Abschätzung verwenden:

**Corollar 4.1.** *Es gilt*

$$y(i, S) \geq \min \{m(i, j) \mid j \in S\} \cdot z_i. \quad (4.30)$$

Damit haben wir eine Klasse gültiger Ungleichungen gefunden.

Betrachten wir abermals unser Beispiel. Nehmen wir an, dass die Kopplung zwischen den beiden Sektormengen unabhängig von der Wahl der restlichen Installationen ist. Betrachtet man eine Lösung mit  $z_{i_1} = z_{i_2} = z_{j_1} = z_{j_2} = \frac{1}{2}$ , so geht aus den linearen Nebenbedingungen von OPT nur hervor, dass  $y(i, S) \geq 0$ . Unsere neue Schnittgleichung 4.28, die aus Corollar 4.1 hervorgeht gibt uns aber  $y(i, S) \geq \frac{3}{2}$ .

Diese Installation-Sektor Ungleichungen werden nur in günstigen Fällen überhaupt von einem Netz  $N$  mit Gleichheit erfüllt. In diesem Fall müssten für  $N$  alle  $y(i, j, Z)$  mit  $Z \in Z(i) \setminus \emptyset$  den Wert null annehmen. Die Abschätzung in Corollar 4.1 ist also ziemlich grob. Die Schnittungleichungen führten bei der Implementierung allerdings zu einer im Anfang etwas schnelleren Anhebung der unteren Schranke.

Aus den oben definierten  $y(i, S)$ -Variablen leiten wir nun Ungleichungen her, die die paarweise Kopplung zwischen Sektormengen beschreiben. Hierzu definieren wir neue Kopplungsvariablen:

**Definition 4.5.** *Seien  $S_1, S_2$  zwei Sektormengen. Wir definieren die Variable  $y(S_1, S_2)$  folgendermaßen:*

$$y(S_1, S_2) := \sum_{i \in S_1} y(i, S_2)$$

Wegen der Symmetrie der Kopplung gilt

$$y(S_1, S_2) = \sum_{i \in S_1} y(i, S_2) = \sum_{j \in S_2} y(j, S_1) = y(S_2, S_1).$$

Die neu eingeführten Variablen beschreiben die Kopplung zwischen zwei Sektormengen.

Betrachten wir das Beispiel in Abbildung 4.7. Die folgenden Ungleichung

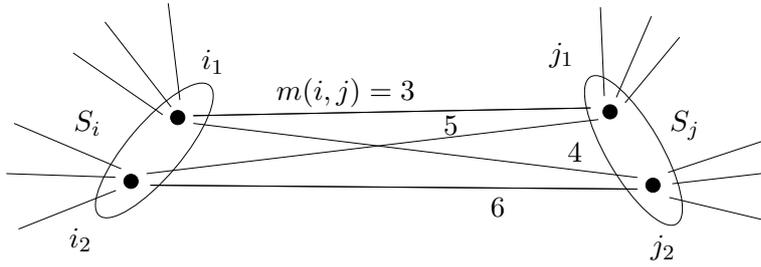


Abbildung 4.7: Störung zwischen zwei Sektormengen.

ist gültig:

$$y(S_1, S_2) \geq 3 + 2z_{i_2} + z_{j_2}. \quad (4.31)$$

Wie gelangt man zu dieser Ungleichung? Die Kopplung zwischen den Sektormengen kann man mit nichtlinearen Termen einfach nach unten abschätzen. Sei  $N$  ein beliebiges Netz und  $i'$  und  $j'$  die Vertreter der Sektormengen  $S_1$  beziehungsweise  $S_2$ . Für die  $y(S_1, S_2)$ -Variablen gilt

$$\begin{aligned} y(S_1, S_2) &= \sum_{i \in S_1} y(i, S_2) \\ &= \sum_{i \in S_1} \sum_{j \in S_2} m(i, j) z_j z_i \end{aligned} \quad (4.32)$$

Diese Ungleichung ist eine gültige Ungleichung für System 4.22. Allerdings ist es keine lineare Ungleichung. Die folgenden Betrachtungen dienen dazu, eine lineare Ungleichung für die Kopplung zwischen Sektormengen herzuleiten. Dazu betrachten wir ein  $i \in S_1$ . Sei  $j^*(i) \in S_2$ , sodass gilt

$$m(i, j^*(i)) = \min_{j \in S_2} \{m(i, j)\}.$$

Dann gilt

$$\sum_{i \in S_1} \sum_{j \in S_2} m(i, j) z_j z_i \quad (4.33)$$

$$= \sum_{i \in S_1} \sum_{j \in S_2} [m(i, j^*(i)) + m(i, j) - m(i, j^*(i))] z_j z_i \quad (4.34)$$

$$= \sum_{i \in S_1} \left[ \sum_{j \in S_2} m(i, j^*(i)) z_j + \sum_{j \in S_2} \underbrace{(m(i, j) - m(i, j^*(i)))}_{\geq 0} z_j \right] z_i \quad (4.35)$$

Wegen  $\sum_{j \in S_2} z_j = 1$  gilt

$$\sum_{j \in S_2} m(i, j^*(i)) z_j = m(i, j^*(i)).$$

Definieren wir

$$d(i, j) := m(i, j) - m(i, j^*(i)),$$

so vereinfacht sich Ausdruck 4.35 zu

$$\sum_{i \in S_1} m(i, j^*(i)) z_i + \sum_{i \in S_1} \sum_{j \in S_2} d(i, j) z_i z_j. \quad (4.36)$$

Sei

$$m(i^*, j^*) := \min \{m(i, j) \mid i \in S_1, j \in S_2\},$$

so gilt

$$\sum_{i \in S_1} m(i, j^*(i)) z_i \quad (4.37)$$

$$= \sum_{i \in S_1} [m(i^*, j^*) + m(i, j^*(i)) - m(i^*, j^*)] z_i \quad (4.38)$$

$$= \sum_{i \in S_1} m(i^*, j^*) z_i + \sum_{i \in S_1} [m(i, j^*(i)) - m(i^*, j^*)] z_i \quad (4.39)$$

$$= m(i^*, j^*) + \sum_{i \in S_1} (m(i, j^*(i)) - m(i^*, j^*)) z_i. \quad (4.40)$$

Sei

$$d(i^*(j), j) := \min_{i \in S_1} \{d(i, j)\}$$

so gilt

$$\sum_{j \in S_2} \sum_{i \in S_1} d(i, j) z_i z_j \quad (4.41)$$

$$= \sum_{j \in S_2} \sum_{i \in S_1} (d(i^*(j), j) + d(i, j) - d(i^*(j), j)) z_i z_j \quad (4.42)$$

$$= \sum_{j \in S_2} \left[ \sum_{i \in S_1} d(i^*(j), j) z_i + \sum_{i \in S_1} (d(i, j) - d(i^*(j), j)) z_i \right] z_j \quad (4.43)$$

$$= \sum_{j \in S_2} d(i^*(j), j) z_j + \sum_{i \in S_1} \underbrace{(d(i, j) - d(i^*(j), j))}_{\geq 0} z_i z_j. \quad (4.44)$$

Durch Substitution der Ausdrücke 4.40 und 4.44 in Ausdruck 4.36 erhalten wir die Ungleichung

$$\begin{aligned} y(S_1, S_2) &\geq m(i^*, j^*) + \sum_{i \in S_1} (m(i, j^*(i)) - m(i^*, j^*)) z_i \\ &\quad + \sum_{j \in S_2} d(i^*(j), j) z_j + \sum_{i \in S_1} (d(i, j) - d(i^*(j), j)) z_i z_j \\ &\geq m(i^*, j^*) + \sum_{i \in S_1} (m(i, j^*(i)) - m(i^*, j^*)) z_i \\ &\quad + \sum_{j \in S_2} d(i^*(j), j) z_j \end{aligned}$$

Mit diesen Betrachtungen haben wir das folgende Lemma bewiesen:

**Lemma 4.3.** *Seien  $S_1 = \{i_1, i_2, \dots\}$ ,  $S_2 = \{j_1, j_2, \dots\}$  zwei Sektormengen, und  $y(S_1, S_2)$  die dazugehörige Kopplungsvariable. Seien  $m(i, j)$  die unteren Schranken für die minimalen Störwerte. Seien*

$$\begin{aligned} m(i, j^*(i)) &:= \min_{j \in S_2} \{m(i, j)\}, \\ m(i^*, j^*) &:= \min \{m(i, j) \mid i \in S_1, j \in S_2\}, \\ \text{und} \\ d(i^*(j), j) &:= \min_{i \in S_1} \{d(i, j)\}. \end{aligned}$$

So gilt:

$$\begin{aligned} y(S_1, S_2) &\geq m(i^*, j^*) + \sum_{i \in S_1} (m(i, j^*(i)) - m(i^*, j^*)) z_i \\ &\quad + \sum_{j \in S_2} d(i^*(j), j) z_j. \end{aligned} \tag{4.45}$$

Die Anwendung dieses Lemmas auf das Beispiel in Abbildung 4.7 ergibt Ungleichung 4.31.

### Dreiecksungleichungen und andere graphische Ungleichungen

Für die IP-Formulierung des MAX-CUT Problems existiert die Klasse der Dreiecksungleichungen. Diese besagen, dass in jedem Dreieck von Kanten maximal zwei im Schnitt sind (oder mindestens eine im „Nicht-Schnitt“). Man mache sich dies an Abbildung 3.1 auf Seite 25 klar. Wir übertragen diese Dreiecksungleichungen nun auf OPT. Betrachten wir hierzu Abbildung 4.5. Zwischen den dort dargestellten Sektormengen  $S_i, S_j$  und  $S_k$  gibt es für jedes Tripel von Installationen dieser Mengen ein Dreieck von Kanten. Die aus dem MAX-CUT Problem übertragene Dreiecksungleichung besagt, dass zwischen den Sektormengen  $S_i, S_j$  und  $S_k$  mindestens soviel Kopplung besteht, wie die Summe der minimalen Störwerte auf den Kanten eines der Dreiecke.

**Beobachtung 4.9.** *Für die Störung zwischen den drei Sektormengen gilt*

$$\begin{aligned} y(S_i, S_j) + y(S_j, S_k) + y(S_k, S_i) \\ \geq \min \{m(u, v) + m(v, w) + m(w, u) \mid u \in S_i, v \in S_j, w \in S_k\}. \end{aligned} \tag{4.46}$$

Das ist leicht zu sehen, denn eine Lösung von OPT induziert in dem gezeigten Auszug des Störungsgraphen ein Dreieck. Die Werte der Störungsfunktionen auf den Kanten des betreffenden Dreiecks sind größer als die  $m(i, j)$ -Werte.

Die Dreiecksungleichungen können in polynomieller Zeit ermittelt werden, da die Zahl der Dreiecke polynomiell von der Zahl der Installationen in  $S_i, S_j$  und  $S_k$  abhängt. Wir möchten das Konzept aus Ungleichung 4.46 nun auf bestimmte Klassen von Teilgraphen von  $G$  übertragen. Hierzu wählen wir eine Teilmenge der Sektormengen und treffen Aussagen über die dort mindestens vorhandene Kopplung. Wir betrachten wie bei den Dreiecksungleichungen die Menge der möglichen Auswahlen (siehe Seite 51). Da ihre Anzahl nicht unbedingt polynomiell von der Zahl der Sektormengen abhängt, sind die Ungleichungen durch Enumeration nicht notwendigerweise in polynomieller Zeit ermittelbar. Wir möchten diesen Sachverhalt in der Folge präziser formulieren.

Die in der folgenden Beobachtung formulierte Klasse von Schnitten besagt, dass die Kopplung innerhalb einer Teilmenge der Sektormengen immer einen bestimmten Wert überschreitet, unabhängig davon, welche Netzkonfiguration gewählt wurde.

**Beobachtung 4.10.** *Sei  $S' \subset \mathcal{S}$  eine Menge von Sektoren. Sei  $G' = (V', E')$  mit  $V' = \bigcup_{S \in S'} S$  ein induzierter Untergraph von  $G$ . Sei  $\mathcal{A}'$  die Menge aller Auswahlen auf  $G'$  und*

$$M := \min_{A \in \mathcal{A}'} \left\{ \sum_{i,j \in A} m(i,j) \right\}.$$

So gilt:

$$\sum_{S_i, S_j \in S'} y(S_i, S_j) \geq M. \quad (4.47)$$

Wir haben diese Ungleichungen nicht implementiert, da die  $m(u, v)$ -Werte in der Regel deutlich kleiner sind als die Werte der später tatsächlich auftretenden Kopplung. Die Dreiecksungleichungen aus Beobachtung 4.9 sind ein Spezialfall der Ungleichungen aus Beobachtung 4.10. Dort hat die Menge  $S'$  drei Elemente.

### Ungleichungen aus dem *Set Partitioning*

In Abschnitt 3.1 führten wir bereits einige Schnittebenen für das *set covering*-Problem ein. Wir möchten diese Schnittebenen nun auf das Problem OPT übertragen.

**Lemma 4.4.** *Sei  $S \in \mathcal{S}$  eine Sektormenge und  $J$  eine Teilmenge der Zeilenindexmenge von  $A$ . Sei  $A_{JS}$  die von  $J$  und  $S$  induzierte Submatrix der Abdeckungsmatrix  $A$ . Hat das System*

$$\min 1_S^T x$$

unter den Nebenbedingungen

$$\begin{aligned} A_{JS}x &\geq 1 \\ x_S &\in \{0, 1\}^{|S|} \end{aligned} \quad (4.48)$$

einen Wert  $w_{0_S} < 1$ , so ist

$$a_{I'}^T x \geq 1$$

eine gültige Ungleichung, wobei

$$I' := \{k \mid a_{jk} = 1 \text{ für ein } j \in J, k \in I \setminus S\}.$$

**Beweis: (Lemma 4.4)** Der Beweis ergibt sich direkt aus Beobachtung 3.5. ■

Für einige Teilmengen  $J$  der Zeilenmenge ist der Wert der LP-Relaxierung von System 4.48 größer als eins. Diese sind komplexitätstheoretisch einfacher zu ermitteln.

**Corollar 4.2.** *Ungleichung  $a_{I'}^T x \geq 1$  gilt auch, wenn die Ganzzahligkeits-Bedingungen in System 4.48 weggelassen werden.*

Mit diesem Verfahren wird nur eine Teilmenge der in Lemma 4.4 definierten Schnitte ermittelt.

## 4.4 Vereinfachungen des IP-Modells

In den Abschnitten 4.2.3 und 4.3.1 stellten wir zwei Methoden vor, das aus unserem Planungsansatz zur Ermittlung interferenzarmer GSM-Netze hervorgehende Optimierungsproblem NOPT zu linearisieren. Das Pixel-Modell aus Abschnitt 4.2.3 wies deutlich mehr Ungleichungen auf als das Obermengen-Modell aus Abschnitt 4.3.1. Für beide Systeme entwickelten wir Schnittungleichungen mit dem Ziel, die LP-Relaxierung und das Konvergenzverhalten des IP-Lösers zu verbessern. Beide Ansätze kann man als gescheitert betrachten, denn für die von uns untersuchten Instanzen waren die auftretenden Laufzeiten sehr groß.

In diesem Abschnitt stellen wir zwei Verfahren vor, die ganzzahligen Programme zu vereinfachen. Weil es kleiner ist, konzentrieren wir uns auf das Obermengen-Modell. Die Vereinfachung erfolgt in beiden Fällen durch Weglassen bestimmter Klassen von Kopplungsbeziehungen.

In Unterabschnitt 4.4.1 stellen wir einen Ansatz vor, bei dem wir große Obermengen aus der Menge der betrachteten Obermengen entfernen. Dieser Ansatz basiert auf der Beobachtung, dass durch Obermengen großer

Kardinalität in den resultierenden Netzen wenig Kopplung entsteht. Die zu diesen Obermengen korrespondierenden Ungleichungen und Variablen werden in unserem Ansatz aus dem System gestrichen. Es wird sich zeigen, dass diese Idee mit Einschränkungen tragfähig ist. Der IP-Löser weist auf dem reduzierten System nach wie vor lange Laufzeiten auf, weil wir nach wie vor viele Kopplungsbeziehungen berücksichtigen müssen, um Lösungen mit akzeptabler Qualität zu erhalten. Daher konnten wir nur für kleine Instanzen Ergebnisse erhalten.

In Unterabschnitt 4.4.2 stellen wir einen Ansatz vor, in dem der  $\delta$ -Schwellwert herabgesetzt wird, um die Zahl der Ungleichungen zu verringern. Schwache Störer werden bei diesem Ansatz also ignoriert. Es zeigt sich, dass der Ansatz im Prinzip gut ist, denn Netzkonfigurationen, die bei einem geringen  $\delta$ -Schwellwert verhältnismäßig wenig Interferenz aufweisen, tun dies in der Regel auch bei großen  $\delta$ -Schwellwerten. Allerdings sind die entstehenden IPs nach wie vor schwer lösbar. Die Laufzeiten sind viel größer als die der Heuristiken, die wir in Abschnitt 4.5 vorstellen.

#### 4.4.1 Reduktion der Zahl der Obermengen

In diesem Abschnitt beschreiben wir ein Verfahren, mittels eines aus dem Obermengen-Modell hervorgehenden Ansatzes zu guten Lösungen zu kommen. Die Idee ist, nur Obermengen kleiner Kardinalität zu berücksichtigen. Hintergrund ist, dass die „Wahrscheinlichkeit“, dass ein Ausdruck der Form

$$z_i \cdot z_j \cdot \prod_{k \in Z} \bar{z}_k$$

den Wert null hat, umso größer ist, je größer die Menge  $Z$  ist. Unser Kalkül ist, dass wir diese Ausdrücke aus der nichtlinearen Zielfunktion entfernen können, ohne dass sich das Problem bedeutend verändert. Transformiert man das nichtlineare Problem danach in ein IP, so führt das Weglassen der Zielfunktionskoeffizienten zur Reduktion der Zahl Ungleichungen und Hilfsvariablen. Wir ermittelten Ergebnisse auf einer Probeinstanz. Sie sind in Tabelle 4.3 dargestellt.

In den Zeilen sind Eigenschaften der IPs aufgeführt, die Kopplungsbeziehungen bis zu einer bestimmten Kardinalität der Obermengen enthalten. So wurden in der ersten Zeile nur Kopplungsbeziehungen mit leerer Obermenge berücksichtigt. In der drittletzten Zeile wurde alle Kopplungsbeziehungen berücksichtigt. Die vorletzte Zeile enthält den Wert des Ausgangsnetzes HG1, die letzte den Wert einer mittels einer später vorgestellten Heuristik ermittelten Lösung.

In Spalte zwei ist aufgeführt, wie viele Beziehungsungleichungen das Obermengen-Modell aufweist. Die dritte Spalte gibt das Verhältnis der Zahl der berücksichtigten Obermengen zur Zahl aller Obermengen wieder. In den

max. Obermen- ge	#Ungleichungen		IP-Lösung		Wert
	Oberm.	Anteil %	Wert	Dauer s	
0	657	5,7	1834	2	14065
1	3079	26,7	5645	63	10661
2	6100	52,8	7829	615	9615
3	8451	73,2	8443	1574	8645
4	9947	86,1	8558	3546	8645
5	10809	93,6	8642	4877	8645
alle	11547	100,0	8645	5492	8645
HG1					30366
LS					8645

Tabelle 4.3: Ergebnisse der Obermengenheuristik.

nächsten beiden Spalten sind Werte und Lösungsdauer der IP-Lösungen dargestellt. Die letzte Spalte enthält die wirklichen Werte der Lösungen, das heißt unter Berücksichtigung aller Obermengen. Die IPs wurden mit dem Obermengen-Modell und einem  $\delta$ -Schwellwert von 9 dB aufgestellt.

Die Zahl der betrachteten Obermengen im Verhältnis zur Gesamtzahl wächst zunächst steil und später weniger steil an. Nur 6,4 Prozent aller Obermengen haben eine Kardinalität von mehr als fünf. Dies ist jedoch auf die Größe der Instanz zurückzuführen. Untersuchungen auf größeren Instanzen ergaben, dass es dort der Anteil von Obermengen der Größe sechs oder mehr weit größer ist als hier. Um einen vergleichbaren Anteil wie hier von 93,6 Prozent zu erreichen, müssen auf Instanzen normaler Größe Obermengen bis etwa zur Größe fünfzehn berücksichtigt werden.

Die Lösungsdauer nimmt mit steigender Zahl der berücksichtigten Obermengen rapide zu. Bei Berücksichtigung aller Obermengen liegt die Lösungsdauer bei etwa anderthalb Stunden. Für größere Instanzen ist das IP praktisch nicht lösbar.

Wir sehen in den letzten drei Spalten, dass die Lösungen mit steigender Kardinalität der Obermengen erwartungsgemäß besser werden. Bei Kardinalität drei finden wir bereits eine Lösung mit optimalem Wert, obwohl nur 73,6 Prozent aller Obermengen berücksichtigt werden. Es handelt sich hierbei allerdings um eine andere Lösung als bei den Kardinalitäten vier und fünf und dem IP mit allen Obermengen. Die dortigen Lösungen sind gleich.

Zum Vergleich haben wir in den letzten beiden Zeilen zwei weitere Lösungen aufgeführt. Die vorletzte Zeile enthält den Wert des Ausgangsnetzes. Dieser ist deutlich größer als die Werte der anderen Lösungen. Mit dem vereinfachten System ist es also möglich, Netzkonfigurationen mit deutlich besseren Netzeigenschaften zu ermitteln. Die letzte Zeile enthält den Wert

einer Lösung, die mit einer Heuristik ermittelt wurde, die wir in Abschnitt 4.5.2 vorstellen werden. Diese Heuristik findet die Optimallösung in wenigen Sekunden.

Die vorgestellte Vereinfachung durch Weglassen ist als Heuristik nur bedingt verwendbar. Es konnten zwar optimale Lösungen ermittelt werden, aber die Laufzeiten waren sehr groß. Der schnellste Durchlauf, bei dem eine optimale Lösung gefunden wurde, dauerte 1574 Sekunden. Für größere Instanzen waren die Laufzeiten selbst bei ausschließlicher Betrachtung der leeren Obermenge so groß, dass keine Optimallösung ermittelt werden konnte. Dabei ist bei größeren Instanzen zu erwarten, dass man optimale Lösungen erst ermitteln kann, wenn die Kardinalität der betrachteten Obermengen deutlich größer ist als drei. Wir schließen daraus, dass andere Verfahren entwickelt werden müssen, um zu guten Lösungen zu gelangen.

#### 4.4.2 Die Rolle des $\delta$ -Schwellwerts

In diesem Unterabschnitt stellen wir unsere zweite Idee vor, das aus dem Optimierungsprogramm NOPT abgeleitete ganzzahlige Programm OPT zu vereinfachen. Nachdem wir im vorigen Unterabschnitt die Menge der betrachteten Obermengen dadurch eingeschränkt haben, dass wir Mengen großer Kardinalität nicht berücksichtigten, reduzieren wir nun den verwendeten  $\delta$ -Schwellwert. Dadurch werden ebenfalls viele Kopplungsverhältnisse nicht ins System aufgenommen.

Zunächst zeigen wir, dass Netze, die bezüglich eines kleinen  $\delta$ -Schwellwerts von 9 dB verhältnismäßig wenig Kopplung aufweisen, auch für den normalen  $\delta$ -Schwellwert von 21 dB gute Kopplungseigenschaften vorweisen können. Allerdings werden wir sehen, dass Netze, die bezüglich eines kleinen  $\delta$ -Schwellwerts eher schlecht sind, bezüglich eines  $\delta$ -Schwellwerts sehr gut sein können. Daher sollte man sich bei der Optimierung nicht ausschließlich auf einen kleinen  $\delta$ -Schwellwert stützen.

Im Anschluss zeigen wir, wie sich die Zahl der Gleichungen und Ungleichungen im Pixel- und im Obermengen-Modell ändern, wenn der  $\delta$ -Schwellwert herabgesetzt wird. Die resultierenden Systeme sind nicht sehr groß. Schließlich werden wir jedoch zeigen, dass auch für das Lösen der kleinen Systeme viel Zeit investiert werden muss, sodass der Ansatz, das IP zu vereinfachen, letzten Endes als gescheitert betrachtet werden muss.

Ob man mit Hilfe eines kleinen  $\delta$ -Schwellwerts Aussagen machen kann über die Qualität des Netzes mit einem Schwellwert von 21 dB wollen wir am Beispiel der in Unterabschnitt 4.4.1 betrachteten Instanz klären. Dort wurden die Ergebnisse mit einem  $\delta$ -Schwellwert von 9 dB ermittelt.

In Tabelle 4.4 ist aufgeführt, welche Werte die Netze für die verschiedenen  $\delta$ -Schwellwerte aufweisen. In der zweiten und dritten Spalte sehen wir die Werte der Netze für die  $\delta$ -Schwellwerte 9 und 21 dB. Die vierte und fünfte Spalte zeigen die durchschnittlichen Knotengrade an. Die letzte Zeile enthält

max. Obermenge	Wert		$\emptyset$ Knotengrad	
	9 dB	21 dB	9 dB	21 dB
0	14065	53844	11,05	21,27
1	10661	46466	11,16	21,72
2	9615	42018	10,00	21,00
3	8645	40912	9,05	20,05
4	8645	40617	9,05	20,05
5	8645	40617	9,05	20,05
alle	8645	40617	9,05	20,05
LS	9049	40424	9,83	20,05

Tabelle 4.4: Gegenüberstellung der IP-Lösungen für  $\delta = 9$  dB und  $\delta = 21$  dB.

die Werte für ein Netz, das wir mit der später vorgestellten lokalen Suche ermittelt haben.

Es ist zu sehen, dass es einen Zusammenhang gibt zwischen den Werten für  $\delta = 9$  dB und  $\delta = 21$  dB. Je besser der Lösungswert für den Schwellwert 9 ist, desto besser ist er für den Schwellwert 21. Da bei einem  $\delta$ -Schwellwert von 9 dB weniger Kopplung im Netz vorhanden ist, sind die Knotengrade hier kleiner. Auch an ihnen ist eine Verbesserung abzulesen. Je besser der Lösungswert bezüglich 9 dB ist, desto besser sind die durchschnittlichen Knotengrade bezüglich 9 und 21 dB. Wir sehen, dass das Netz, was mit Obermengen der Größe bis drei ermittelt wurde, für 21 dB einen schlechteren Wert aufweist, als die Netze, die mehr Obermengen berücksichtigten. Dieses Resultat hätte aber genauso gut andersherum ausfallen können.

Die letzte Zeile zeigt eine Lösung der lokalen Suche, die wir in Abschnitt 4.5.2 vorstellen werden. Sowohl ihr Lösungswert als auch ihr durchschnittlicher Knotengrad bezüglich 9 dB sind größer als die Werte der anderen Netze. Trotzdem ist die heuristisch ermittelte Lösung der lokalen Suche besser bezüglich des großen  $\delta$ -Schwellwerts. Es reicht also nicht, Netze nur unter Berücksichtigung eines kleinen Schwellwerts zu ermitteln. Die Differenz zwischen den Lösungswerten beträgt nur wenige Prozent, sodass man durchaus einen kleinen  $\delta$ -Schwellwert verwenden kann, wenn man nicht an optimalen, sondern nur an guten, schnellen Lösungen interessiert ist. Welches für die oben betrachtete Instanz der Optimalwert bezüglich 21 dB ist, konnten wir nicht ermitteln.

Die Verkleinerung des  $\delta$ -Schwellwerts führt im IP-Ansatz zu einer deutlichen Reduktion der Zahl der Ungleichungen und Variablen. Hierzu betrachte man Abbildung 4.5. Hier sind für vier Netze die Größen der IPs für Pixel- und Obermengen-Modell mit den jeweiligen  $\delta$ -Schwellwerten von 9 und 21 dB aufgeführt. Wir sehen, dass für  $\delta = 21$  dB sowohl im Pixel- als auch im Obermengen-Modell zwischen etwa drei- bis fünfmal mehr Ungleichungen

Name	Pixel-Modell		Obermengen Modell	
	$\delta = 9$ dB	$\delta = 21$ dB	$\delta = 9$ dB	$\delta = 21$ dB
HG1/4	85592	317992	11547	34002
HG3/1	101996	404956	16627	53730
NET01/4	241564	1108160	20069	67509
H01/3	140709	594216	26084	93882

Tabelle 4.5: Vergleich der IP-Größen für die  $\delta$ -Schwellwerte 9 und 21 dB.

aufgestellt werden als für  $\delta = 9$  dB. Die Zahl der Ungleichungen ist abgesehen vom Pixel-Modell für  $\delta = 21$  dB von einer Größenordnung, dass man durchaus gute Laufzeiten erwarten könnte.

Um zu sehen, wie groß die Verbesserung des Lösungsverhaltens des IP-Lösers bei Reduzierung des  $\delta$ -Schwellwerts ist, betrachte man Tabelle 4.6.

$\delta$	Ungleichg.	IP-Lösung		
		Wert	Dauer $s$	Lücke %
9	11547	8645	5491	0
21	34002	41060	6858	54,87

Tabelle 4.6: Gegenüberstellung der Laufzeiten im Obermengen-Modell für  $\delta = 9$  dB und  $\delta = 21$  dB.

Für 9 dB konnte die Optimallösung ermittelt werden, für 21 dB wich der Wert der gefundenen Lösung nach vergleichbarer Laufzeit von der unteren Schranke noch um etwa 55 Prozent ab. Es ist also zu erwarten, dass der IP-Algorithmus noch lange laufen würde.

Die Instanz in Tabelle 4.6 ist sehr klein. Sie hat 36 Sektoren, und pro Sektor stehen 2 Installationen zur Auswahl. Die Menge der Pixel ist etwa 6000. Die Instanzen, für die wir Ergebnisse erzielen wollen, weisen 150 bis 250 Sektoren auf. Wir möchten mit 4 Installationen pro Sektor rechnen. Die Anzahl der betrachteten Pixel soll bei 25000 liegen. Für die größeren in Tabelle 4.5 aufgeführten Instanzen zeigte der IP-Löser nach langen Laufzeiten noch große Lücken an.

Anhand der Größen der lösbaren Instanzen sehen wir, dass deutliche Verbesserungen der IP-Formulierung nötig wären, um mit diesem Ansatz für Instanzen normaler Größe taugliche Ergebnisse zu erzielen. Die hier vorgestellten Versuche, die IP-Formulierung zu vereinfachen, betrachten wir als gescheitert. Wir wenden uns in Abschnitt 4.5 daher Verfahren zu, die keine IP-Formulierung verwenden und auf heuristischem Wege gute Lösungen erzielen sollen.

## 4.5 Heuristiken

In den Abschnitten 4.2 und 4.3 wurde beschrieben, wie man den in Abschnitt 4.1 hergeleiteten Planungsansatz für GSM-Netze als Optimierungsproblem formulieren kann. In Abschnitt 4.2 gaben wir zunächst eine nichtlineare Formulierung an. Die dort vorgestellte Umformulierung als IP wurde in Abschnitt 4.3 verbessert. Beide Formulierungen führten zu Programmen, die mit der zur Verfügung stehenden Software und Rechnerleistung nicht auf befriedigende Weise gelöst werden konnten. Daher verbesserten wir die Problemformulierung zunächst mit Schnittungleichungen. Dieser Ansatz brachte keine wesentlichen Fortschritte, und so versuchten wir in Abschnitt 4.4, durch Weglassen von Ungleichungen und Variablen ganzzahlige Programme zu erhalten, die in befriedigender Zeit gelöst werden konnten. Die ermittelten Lösungen waren nicht mehr notwendigerweise optimal für das Ausgangsproblem.

In diesem Abschnitt verfolgen wir die Idee weiter, gute Lösungen zu ermitteln, die nicht notwendigerweise optimal sind. Wir stellen hier Algorithmen vor, die schnell arbeiten und gute Lösungen ermitteln. Unseren Ehrgeiz, bewiesenermaßen optimale Lösungen zu finden, stellen wir damit zurück.

Zunächst beschreiben wir in Unterabschnitt 4.5.1 Startheuristiken. Diese beruhen auf der Beobachtung, dass gute Netze in der Regel große Tilts aufweisen. Wir stellen ein ganzzahliges Programm so auf, dass eine Lösung umso besser ist, je größer die Summe aller Neigungen ist. Die optimale Lösung muss die Abdeckungs- und Partitionsbedingungen erfüllen. Mit dieser Heuristik erreichen wir teilweise gute Verbesserungen.

Unterabschnitt 4.5.1 heißt „Startheuristiken“, weil wir die dort ermittelten Lösungen als Startlösungen für die in Unterabschnitt 4.5.2 vorgestellte lokale Suche verwenden können. Diese lokale Suche hat von allen von uns entwickelten Lösungsmethoden die besten Ergebnisse geliefert. In Unterabschnitt 4.5.2 erläutern wir ausführlich unsere Überlegungen zu diesem Algorithmus, beschreiben ihn mit Pseudocode und vergleichen erste Ergebnisse mit den Resultaten der IP-Optimierung. Bei der Implementierung verwendeten wir unter anderem die Erkenntnisse über die  $\delta$ -Schwellwerte, die wir in Abschnitt 4.4 gemacht hatten.

### 4.5.1 Tiltheuristik

Wir stellen hier eine Startheuristik vor. Sie ermittelt eine gültige Lösung für OPT, mit der weitergerechnet wird im lokalen Suchalgorithmus, der in Unterabschnitt 4.5.2 erläutert wird.

Sei  $t : I \rightarrow \mathbb{N}$  eine Funktion, die jeder Installation  $i$  aus  $I$  ihren elektrischen Tilt in Grad zuordnet. Wir gehen hier davon aus, dass die elektrischen Tilts ganzzahlig sind, und dass die Antenne nicht nach oben ausgerichtet wird. Sei  $c : \mathbb{N} \rightarrow \mathbb{N}$  eine Gewichtsfunktion, die jedem elektrischen Tilt einen

Wert zuordnet. Unsere Tiltheuristik ermittelt eine Startlösung durch Lösen des ganzzahligen Programmes

$$\max \sum_{i \in I} c(t(i)) \cdot z_i$$

sodass

$$\begin{aligned} \sum_{i \in \bar{I}_p} z_i &\geq 1 \quad \forall p \in \bar{P} \\ \sum_{i \in S} z_i &= 1 \quad \forall S \in \mathcal{S} \\ z_i &\in \{0, 1\} \quad \forall i \in I. \end{aligned} \tag{4.49}$$

Nach unseren Erfahrungen erzeugen Installationen mit einem großen Tilt tendenziell am wenigsten Kopplung. Daher ist es sinnvoll, die Gewichtsfunktion  $c$  der elektrischen Tilts so zu definieren, dass sie monoton steigt. Die Netze in Tabelle 4.7 enthalten Installationen mit elektrischen Tilts von zwei, vier, sechs und acht Grad. Die Lösungen wurden mit der Gewichtsfunktion  $c(2) = 1, c(4) = 3, c(6) = 7$  und  $c(8) = 15$  ermittelt.

Name	Netz		TRX			Knotengrad		Wert
	Zellen	Inst.	$\sum$	$\emptyset$	max	$\emptyset$	max	
BPN1	148	592	303	2,04	6	39,53	83	347970
opt			303	2,03	6	23,39	54	123358
BPNe	193	553	360	1,86	3	51,96	105	361248
opt			354	1,83	3	28,83	66	121062
HAG1	36	144	82	2,27	4	28,61	35	103483
opt			83	2,30	6	21,77	35	52162
HAG3	60	240	107	1,78	4	44,86	58	154287
opt			112	1,86	4	26,63	45	43599

Tabelle 4.7: Ergebnisse der Tiltheuristik.

Es ist zu sehen, dass beim Kriterium der Summe der Kopplungswerte in der letzten Spalte deutliche Verbesserungen erreicht werden konnten. Die maximalen und durchschnittlichen Knotengrade haben sich ebenfalls deutlich verringert. Wir möchten eine genauere Analyse auf Abschnitt 4.6 verschieben. Tabelle 4.9 auf Seite 73 ist so aufgebaut wie Tabelle 4.7. Dort erfolgt eine eingehende Analyse. Hier sei nur soviel gesagt, dass die Ergebnisse der Startheuristik deutliche Verbesserungen gegenüber den Ausgangsnetzen darstellten.

#### 4.5.2 Lokale Suche

In diesem Unterabschnitt stellen wir ein heuristisches Verfahren zur Ermittlung von Lösungen vor. Es handelt sich um ein lokales Suchverfahren. Seine

Vorteile sind, dass es mit hoher Geschwindigkeit Lösungen ermittelt, und dass es auf sehr großen Instanzen anwendbar ist. Die Heuristik arbeitet mit verschiedenen veränderbaren Parametern. Sie arbeitet so schnell, dass der Anwender die Durchläufe beobachten und die Parameter dem Resultat entsprechend verändern kann. Ein Nachteil der Heuristik ist, dass sie keinen Hinweis auf die Optimalität der gefundenen Lösungen gibt.

Zunächst machen wir einige einleitende Überlegungen zur Implementierung der lokalen Suche. Wir definieren dort die Nachbarschaften und stellen fest, dass der Suchraum nach dieser Definition nicht zusammenhängend ist. Deshalb entwickeln wir eine Methode, die es ermöglicht, von einer Zusammenhangskomponente in eine andere zu gelangen. Weiterhin stellen wir Überlegungen zur Zielfunktion an. Wir verwenden unsere Erfahrungen aus Abschnitt 4.4.1 dahingehend, dass wir den verwendeten  $\delta$ -Schwellwert zur Bestimmung des Lösungswertes dynamisch anpassen.

Der Algorithmus wird daraufhin mit Hilfe von Pseudocode beschrieben und ausführlich erläutert. Schließlich vergleichen wir Laufzeiten und Ergebnisse der Heuristik mit denen des IP-Algorithmus. Dort wird gezeigt, dass die Heuristik auf den von uns untersuchten kleineren Instanzen immer optimale Lösungen ermittelt hat. Für Instanzen, die mit dem IP-Ansatz nicht gelöst werden konnten, zeigen wir, dass die lokale Suche Lösungen findet, die gleich gut oder besser sind als die vom IP-Löser ermittelten.

### Einleitende Überlegungen

Die von uns verwendete Heuristik ist eine *lokale Suche*. Zu jedem Zeitpunkt liegt eine aktuelle Lösung vor, deren Nachbarschaft untersucht wird. Zwei Lösungen sind dabei benachbart, wenn sie sich nur in der Ausrichtung einer einzigen Antenne unterscheiden. Befindet sich in der Nachbarschaft der aktuellen Lösung eine bessere, so wird in der Nachbarschaft der besseren Lösung weitergesucht. Das Qualitätskriterium ist hierbei die Summe der Nebendiagonal-Einträge der Kopplungsmatrix des Netzes. Der Wechsel der aktuellen Lösung heißt *lokaler Schritt*.

Mitunter erreicht man Lösungen, die keinen besseren Nachbarn aufweisen. Durch die Struktur des Optimierungsproblems ist keine einfache Beurteilung möglich, ob man sich in einem lokalen oder in einem globalen Minimum befindet. An einem solchen Punkt muss also entschieden werden, ob der Algorithmus terminieren soll, oder ob ein großer Schritt zu einem anderen Ausgangspunkt gemacht wird. Wir bezeichnen einen solchen Schritt als *Sprung*. Nachdem die Startheuristik eine Anfangslösung ermittelt hat, beginnt die lokale Suche nach besseren Lösungen.

Zunächst möchten wir den Begriff *benachbart* für gültige Netze präzise definieren. In unseren Algorithmen verwenden wir den *Hamming-Abstand* der Inzidenzvektoren als Metrik. Seien

$$z = \{z_1, \dots, z_n\}, z' = \{z'_1, \dots, z'_n\} \in \{0, 1\}^n$$

die Inzidenzvektoren zweier Netze. Sie sind *benachbart*, wenn ihr Hamming-Abstand kleiner oder gleich zwei ist, also wenn gilt

$$\sum_{i=1}^n |z_i - z'_i| \leq 2,$$

und wenn beide Netze gültig sind. Man beachte, dass wir bei jedem lokalen Schritt darauf achten, dass die aktuelle Lösung gültig ist, also wenn aus jeder Sektormenge genau eine Installation gewählt wurde, und wenn das Netz die Abdeckungsbedingungen erfüllt. Da in jeder Sektormenge genau eine Installation verwendet wird, bedeutet das, dass zwei Lösungen genau dann benachbart sind, wenn sich die Installationen in genau einem Sektor unterscheiden. Bildlich gesprochen betrachten wir bei einem lokalen Schritt, was passiert, wenn wir die Ausrichtung genau einer Antenne verändern.

Ein wichtiger Aspekt bei der Verwendung von lokalen Suchalgorithmen ist der Zusammenhang des Suchraums. Man sollte die Frage klären, ob für jedes Paar von Lösungen eine Sequenz lokaler Schritte existiert, mit der man von der einen Lösung zur anderen gelangt. Weil in OPT sowohl Abdeckungs- als auch Partitionsbedingungen vorkommen, ist der Suchraum unter Umständen nicht zusammenhängend. Man betrachte hierzu das folgende Beispiel:

$$\min x_2 + x_4$$

sodass

$$x_1 + x_4 \geq 1$$

$$x_2 + x_3 \geq 1$$

$$x_1 + x_2 = 1$$

$$x_3 + x_4 = 1$$

$$x \in \{0, 1\}^4.$$

Die gültigen Lösungen dieses Problems sind  $\{1, 0, 1, 0\}$  und  $\{0, 1, 0, 1\}$ . Sie sind nach unserer Definition nicht benachbart. Wählen wir  $\{0, 1, 0, 1\}$  als Startlösung, so gelangen wir mit lokalen Schritten nicht zur optimalen Lösung. Mit diesem Beispiel haben wir außerdem gezeigt, dass die lokale Suche auf unseren Probleminstanzen unter Umständen nicht gut arbeitet. Der von uns implementierte Algorithmus verwendet Sprünge, die es erlauben, von einer Zusammenhangskomponente des Suchraums in eine andere zu gelangen.

Zur Bewertung einer Lösung ziehen wir wie bei der ganzzahligen Optimierung die Kopplungsmatrix heran. Eine schnelle Ermittlung der Kopplungsmatrizen ist entscheidend für die Laufzeit des Algorithmus. Die Dauer der Berechnung hängt stark von dem verwendeten  $\delta$ -Schwellwert ab, dem Wert der Pfadverlust-Differenzen, ab dem eine Installation als Störer gewertet wird. Daher wird in unserem Algorithmus der  $\delta$ -Schwellwert dynamisch

verändert. Je schlechter die Chancen eingeschätzt werden, in der Nachbarschaft der aktuellen Lösung gute Lösungen zu finden, desto kleiner ist der  $\delta$ -Schwellwert. Hat die aktuelle Lösung einen Wert, der dem der momentan besten gefundenen Lösung nahe ist, so wird mit einem hohen  $\delta$ -Schwellwert mehr Zeit in die Analyse der Nachbarschaft investiert. Dass dieses Verfahren sinnvoll ist, kann man Tabelle 4.4 auf Seite 63 entnehmen. Netze, die bezüglich eines kleinen Schwellwertes einen guten Wert aufweisen, tun das auch tendenziell bei der Beurteilung mit einem großen Schwellwert.

Bei lokalen Such-Algorithmus gibt es zwei klassische Strategien, die aktuelle Lösung zu verändern: Die erste sieht vor, zum ersten gefundenen Nachbarn zu wechseln, der einen besseren Lösungswert hat. Die zweite Strategie sieht vor, alle Nachbarn zu betrachten und den besten gültigen Nachbarn als Ziel des lokalen Schrittes auszuwählen. Es ist in der Regel nicht möglich zu sagen, welche Strategie für die gegebene Situation die bessere ist. Die zweite Strategie hat den Vorteil, dass sich die Lösung bei den meisten Schritten stärker verbessert als bei der ersten Strategie, aber den Nachteil, dass jeder Schritt im Vergleich zur ersten Strategie länger dauert.

Der von uns implementierte Algorithmus arbeitet auf vielen Instanzen am erfolgreichsten, wenn zunächst nur ein Teil der Nachbarschaft untersucht wird, zum Beispiel jeder dritte Nachbar. Aus dieser Menge wird der beste Nachbar als neue Lösung gewählt. Falls jedoch kein besserer Nachbar gefunden wurde, so wird die Menge aller Nachbarn untersucht. Der Anteil der zunächst untersuchten Nachbarn wird wie die  $\delta$ -Schwellwerte dynamisch an die augenblickliche Situation angepasst. Je besser die aktuelle Lösung ist, desto größer ist die von Anfang an untersuchte Teilmenge der Nachbarn. Die Idee ist abermals, dass nur Rechenzeit investiert wird, wenn es sich lohnt.

Mitunter gelangt der Algorithmus an einen Punkt, wo kein besserer Nachbar existiert. In unseren Algorithmen verwendeten wir – abgesehen von der Terminierung des Algorithmus – zwei Methoden, in einem solchen Fall vorzugehen.

Bei der ersten Methode wird die Ausrichtung einer zufälligen Menge von Antennen auf zufällige Weise verändert. Das resultierende Netz ist nicht immer gültig, weswegen die Ausrichtungen in einer Serie von Validierungsschritten teilweise nachkorrigiert werden müssen. Ziel dieses zufälligen Sprunges ist es, lokalen Minima der Zielfunktion zu entfliehen. Unter Umständen haben uns die lokalen Schritte in eine Sackgasse geführt, obwohl in der näheren Umgebung noch Lösungen mit besseren Zielwerten vorhanden sind. Führt man einen zufälligen Sprung aus, so verlässt man die Sackgasse und gelangt mitunter zu diesen besseren Lösungen. Außerdem führen uns die lokalen Sprünge unter Umständen in eine neue Zusammenhangskomponente mit besseren Lösungen.

Bei der zweiten Methode werden für einen Teil der Antennen neue Tiltvariationen ermöglicht. Manchmal führte eine bestimmte Ausrichtung einer Antenne immer zu schlechten Lösungen, weil die Ausrichtung in vielen Zellen

```

(1) findStartSolution();
(2) while goAhead() {
    precision = smallDelta;
(3)   do{
        candidates = allCandidates.getRandomSubset();
        betterNeighborFound = findBetterNeighbor(candidates);
        if !betterNeighborFound
            findBetterNeighbor(allCandidates);
    } while betterNeighborFound;

    precision = largeDelta;
(4)   do{
        candidates = allCandidates.getRandomSubset();
        betterNeighborFound = findBetterNeighbor(candidates);
        if !betterNeighborFound
            findBetterNeighbor(allCandidates);
    } while betterNeighborFound;

(5)   if !goAhead()
        break;

(6)   if makeRandomStep()
        randomStep();
        else
        update();

    }

```

Abbildung 4.8: Lokale Suche.

große Störungen verursachte. Eine solche Ausrichtung wird aus der Menge  $I$  der Installationen entfernt. Neue, bisher nicht betrachtete Ausrichtungen werden untersucht und in die Menge  $I$  mit aufgenommen, wenn sie sich als vielversprechend erweisen. Die Erneuerung der Menge der Tiltvariationen erweist sich in der Regel als aufwändig, aber lohnenswert.

Der Algorithmus im nachfolgenden Abschnitt noch einmal mit Pseudocode beschrieben.

### Der Algorithmus

Abbildung 4.8 zeigt den Pseudocode des Algorithmus. In Schritt (1) findet die Methode `findStartSolution()` eine akzeptable Startlösung. Diese existiert, da die Pixelmenge so gewählt ist, dass das Ausgangsnetz sie abdeckt. Es ist auch möglich, die Methode `findStartSolution()` so zu implemen-

tieren, dass sie das Ausgangsnetz als Startlösung festlegt. Dann wird von Anfang an mit jedem lokalen Schritt eine Lösung erreicht, die besser ist als das Ausgangsnetz. Alternativ kann man die Tiltheuristik aus Abschnitt 4.5.1 verwenden.

Mit der Startlösung wird die `while`-Schleife (2) betreten und erst wieder verlassen, wenn die Methode `goAhead()` den Wert `false` zurückgibt. Die Methode `goAhead()` liest hierzu verschiedene Daten ein, die während des Algorithmus gesammelt werden. Üblicherweise brechen wir die Optimierung ab, wenn die `update()`-Methode keine Installationen findet, die die aktuelle Optimallösung verbessern. Die `update()`-Methode ist für die Erneuerung der Menge der untersuchten Tiltvariationen zuständig.

In der `do`-Schleife (3) wird eine Serie von lokalen Schritten ausgeführt. In der Methode `findBetterNeighbor()` wird zunächst immer nur ein kleiner, zufälliger Teil der Nachbarschaft untersucht. Der Rückgabewert dieser Methode zeigt an, ob ein besserer Nachbar gefunden wurde. Ist das nicht der Fall, wird die gesamte Nachbarschaft analysiert. Wird auch dort kein besserer Nachbar gefunden, wird die `do`-Schleife verlassen. Die Bewertung der Lösungen erfolgt hier mit kleiner Präzision.

Die `do`-Schleife (4) macht dieselben Schritte wie Schleife (3), außer dass dieses Mal mit großer Präzision, also der exakten Bewertung der Lösungen durch die Kopplungsmatrix gerechnet wird.

In (5) wird durch nochmaligen Aufruf der Methode `goAhead()` entschieden, ob der Algorithmus terminieren soll. Ist das nicht der Fall, so wird in (6) durch den Aufruf der Methode `makeRandomStep()` entschieden, ob ein Zufallssprung gemacht werden soll oder die Menge der möglichen Ausrichtungen durch Aufruf der Methode `update()` verändert werden soll. Beide Methoden sind so implementiert, dass nach ihrer Terminierung ein gültiges Netz vorliegt.

### Qualität der gefundenen Lösungen

Hier vergleichen wir die Ergebnisse der zuvor beschriebenen lokalen Suche mit den Ergebnissen des IP-Lösers. Man betrachte Tabelle 4.8.

Name	Netz		IP			lokale Suche
	Zellen	Inst.	Wert	Lücke %	Dauer <i>s</i>	
HG1/4	36	72	8645	0	5492	8645
BPN1/2	148	169	64952	0	9	64952
BPN1/3	148	187	60069	0	123	60069
HG3/2	60	87	16835	2,75	11038	16835
BPN1/4	148	238	53023	11,38	125518	52894

Tabelle 4.8: Vergleich der Ergebnisse der Heuristik und des IP-Lösers.

Es sind Resultate für fünf Netze dargestellt. Spalten zwei und drei führen auf, wie viele Antennen die Netze jeweils enthalten, und wie viele Ausrichtungs-Variationen wir bei der Optimierung berücksichtigten. Die mit IP überschriebenen Spalten enthalten Informationen über Wert, Lücke und Laufzeiten der Optimierung. Die oberen drei Instanzen konnten optimal gelöst werden, die unteren beiden weisen einen Lücke auf. Die letzte Spalte schließlich gibt den Wert der Lösung an, die mit der lokalen Suche ermittelt wurde.

Man sieht, dass die Laufzeiten der drei gelösten Szenarien sehr unterschiedlich sind. Sie hängen nicht in erster Linie von der Zahl der Antennen ab. Der IP-Löser arbeitet umso besser, je größer der Anteil von Sektormengen ist, wo nur eine Installation zur Verfügung steht, und je größer das Verhältnis von Antennen zu Installationen ist. Ist dieses Verhältnis wie in BPN1/2 groß, so ist der Lücke in der Regel von Anfang an ziemlich klein.

Die Laufzeiten der lokalen Suche sind nicht aufgetragen. Sie liegen für alle gelösten Instanzen im Sekundenbereich und für die nicht optimal gelösten Instanzen bei etwa einer Minute. Für die letzten zwei Instanzen wurde die lokale Suche mehrfach mit verschiedenen gewählten internen Parametern durchgeführt. Die Laufzeiten hängen sowohl von den gewählten Parametern als auch von der verwendeten Startlösung ab. Die angezeigten Werte sind die besten, die mit der lokalen Suche ermittelt werden konnten. Im Szenario HG3/2 hat sie keine bessere Lösung ermitteln können als der IP-Löser. Im Szenario BPN1/4 hingegen eine geringfügig bessere Lösung gefunden.

Wir deuten die Ergebnisse dahingehend, dass die Werte der beim Lösungsvorgang des IP-Lösers ermittelten Lösungen nahe am Optimalwert liegen. Dass das Konvergenzverhalten der unteren Schranke im IP-Löser so schlecht ist, führen wir darauf zurück, dass das Obermengen-Modell keine gute Linearisierung von NOPT darstellt. Da die Implementierung der ersten Klassen von Schnittungleichungen keine deutlichen Fortschritte ermöglichten, beschlossen wir, den Schrittabenansatz nicht weiterzuverfolgen.

Wir konzentrierten uns bei den größeren Instanzen daher darauf, gute Lösungen mit dem lokalen Suchalgorithmus zu ermitteln.

## 4.6 Ergebnisse

Wir möchten nun unsere Optimierungsergebnisse vorstellen. Zunächst zeigen wir, welche Fortschritte wir bezüglich unseres Optimierungskriteriums erreicht haben. Man wird erkennen, dass die Summe der Kopplungswerte der optimierten Netze im Vergleich zu den Ausgangsnetzen deutlich reduziert werden konnte. Es folgt eine Analyse der Frequenzzuweisung. Hierzu transformieren wir die Kopplungsmatrizen in Interferenzmatrizen, wie sie in Abschnitt 4.1 beschrieben wurden. Daraufhin führen wir eine Frequenzzuweisung mit verschiedenen zur Verfügung stehenden Mengen von Funk-

kanälen durch. Es zeigt sich, dass die optimierten Netze bei gleicher Kanalzahl deutlich weniger Interferenz aufweisen. Diese Tatsache kann man benutzen, um Kanäle einzusparen. Wir führen eine zweite Frequenzzuweisung mit weniger Kanälen durch. Die resultierende Interferenz in den optimierten Netzen ist dann noch immer geringer als die Interferenz der Ausgangsnetze bei Verwendung vieler Kanäle. Durch die Optimierung ist es uns also möglich, Frequenzen einzusparen.

Während unserer Untersuchungen konzentrierten wir uns auf die Betrachtung von vier Szenarien. Alle Ausgangsnetze wurden mehrfach untersucht, wobei die Erweiterung der Antennenmenge  $N$  zur Installationsmenge  $I$  immer unterschiedlichen Gesichtspunkten folgte. Die in Tabelle 4.9 vorgestellten Ergebnisse stellen die erfolgreichsten Durchläufe dar. In jedem

Name	Netz		TRX			Knotengrad		Wert
	Zellen	Inst.	$\sum$	$\emptyset$	max	$\emptyset$	max	
BPN1	148	592	303	2,04	6	39,53	83	347970
opt			304	2,05	6	22,93	54	117116
BPNe	193	553	360	1,86	3	51,96	105	361248
opt			351	1,81	3	27,70	66	118416
HAG1	36	144	82	2,27	4	28,61	35	103483
opt			85	2,36	6	21,33	35	40010
HAG3	60	240	107	1,78	4	44,86	58	154287
opt			113	1,86	4	24,83	44	38796

Tabelle 4.9: Instanzen der Optimierung.

der horizontalen Abschnitte ist an oberer Stelle das Ausgangsnetz und an unterer Stelle das gemäß dem Optimierungskriterium aus Abschnitt 4.2.3 optimierte Netz enthalten.

In der ersten Spalte ist der Name des Szenarios enthalten. In der zweiten sehen wir, wie viele Antennen die Netze enthalten. Ausgangsnetz und optimiertes Netz haben natürlich immer die gleiche Zahl von Antennen. In der dritten Spalte ist festgehalten, wie viele Installationen bei der Optimierung herangezogen wurden. Außer im Szenario BPNe betrachteten wir immer vier mögliche Antennenausrichtungen. Bei BPNe betrachteten wir für einen Teil der Antenne vier Tiltvariationen, die restlichen Antennen blieben fixiert.

Die drei nächsten, mit TRX überschriebenen Spalten, enthalten Informationen darüber, wie viele TRX im Netz installiert werden müssen, wie viele sich im Durchschnitt auf jeder Antenne befinden, und wie groß die maximale Anzahl von TRX pro Antenne ist. Diese Zahlen wurden mit der in Abschnitt 4.1 erläuterten Erlang-B Formel ermittelt.

In den beiden nächsten Spalten sehen wir Informationen über die ebenfalls in Abschnitt 4.1 eingeführten Knotengrade. Die linke der beiden Spalten gibt an, welchen Grad die Antennen im Durchschnitt haben, die rechte, welches der maximale Knotengrad im Netz ist. In der letzten Zeile schließlich

ist aufgeführt, welche Werte die Ausgangsnetze und die optimierten Netze haben. Aufgeführt ist die Summe der Kopplungswerte im Netz.

Die Zahl der notwendigen TRX bleibt bei der Optimierung relativ unverändert. Im Szenario BPNe sinkt ihre Zahl um 2,5 Prozent, im Szenario HAG3 ist die größte prozentuale Steigerung um 5,6 Prozent festzustellen. Da der Verkehr für Ausgangs- und optimiertes Netz gleich ist, ist die Gesamtzahl der TRX ausschließlich von der Form der *best server*-Bereiche und ihrer Größe abhängig. Aus Tabelle 4.1 auf Seite 30 geht hervor, dass die Zahl der TRX steigen würde, wenn sich bei der Optimierung viele kleine Zellen bilden würden, da das Verhältnis der Zahl der TRX im Verhältnis zur Zahl der Nutzer umso kleiner ist, je kleiner die Zahl der Nutzer ist. Dass die Gesamtzahl der TRX bei der Optimierung relativ konstant ist, werten wir als positives Zeichen. Wir schließen daraus, dass die Verteilung der Zellgrößen im optimierten Netz ähnlich ist, wie im Ausgangsnetz. Negative Folgen der Optimierung hätten sein können, dass sich wenige „Riesenzellen“ herausbilden und viele sehr kleine Zellen ohne Verkehr. Dieser Effekt tritt nicht auf. Damit sehen wir These 4.3 aus Abschnitt 4.1 bestätigt, dass die Zellgrößen nicht stark schwanken, und es ausreicht, bei der Optimierung nur die Kopplungswerte zu beachten und nicht die  $c^{co}$ -Werte.

Die durchschnittlichen Knotengrade haben sich bei der Optimierung teilweise deutlich verändert. Im Szenario BPNe ist mit 46 Prozent die größte Verbesserung erkennbar. Der Knotengrad ist ein Maß für die Dichte des dem FAP zugrundeliegenden Graphen. Das FAP ist ein Färbungsproblem, und der maximale Knotengrad zeigt an, wie viele Farben man mindestens verwenden muss, damit im Graphen keine Kante vorkommt, die ein größeres Gewicht als 0,01 hat, und deren Endknoten dieselbe Farbe haben.

Die Kopplungswerte aller Ausgangsnetze konnten deutlich verbessert werden. In allen vier Szenarien gelang eine starke Reduzierung der Summe der Nebendiagonaleinträge. Die größte Reduktion wurde im mit 75 Prozent im Szenario HAG3 erreicht, die geringste im Szenario HAG1 mit 61 Prozent. Tabelle 4.10 wird darüber Aufschluss geben, welche Auswirkungen die Reduktion der Kopplungswerte auf die nach der Frequenzzuweisung auftretende Interferenz hat.

Für die Zuweisung wurde Software verwendet, in der die in [6] vorgestellten Frequenzzuweisungs-Verfahren implementiert sind. Da es sich um heuristische Verfahren handelt, unterscheiden sich zuweilen die Ergebnisse der einzelnen Durchläufe für eine Instanz. Daher haben wir für jede Instanz mehrere Durchläufe vorgenommen.

Man betrachte Tabelle 4.10. Jede Zeile entspricht einer Netzkonfiguration. Unter den Ausgangskonfigurationen sind immer die optimierten Netze aufgeführt. Für jedes Szenario wurden Frequenzzuweisungen mit zwei verschiedenen Anzahlen von Kanälen durchgeführt. Jedem Szenario entsprechen also vier Zeilen in Tabelle 4.10.

Die obere Zahl der Kanäle richtet sich danach, wie viele Kanäle in etwa für das Ausgangsnetz benötigt werden, damit die resultierende Interferenz

Name	Kanäle	Interferenz			Ungültige Zuweisungen
		max	min	$\varnothing$	
BPN1	45	7,45	5,68	6,47	0/10
opt	45	0,01	0,00	0,00	0/10
BPN1	35	–	–	–	10/10
opt	35	1,41	0,72	0,95	0/10
BPNe	30	7,56	5,48	6,42	0/10
opt	30	0,01	0,00	0,00	0/10
BPNe	25	25,18	21,63	23,69	4/10
opt	25	0,57	0,32	0,44	0/10
HAG1	45	3,08	1,44	2,37	0/10
opt	45	0,00	0,00	0,00	0/10
HAG1	31	–	–	–	10/10
opt	31	0,97	0,84	0,96	0/10

Tabelle 4.10: Ergebnisse der Frequenzzuweisung.

akzeptable Werte annimmt. Dabei hielten wir Interferenzwerte zwischen fünf und zehn für akzeptabel. Die untere Zahl richtet sich nach dem optimierten Netz. Wir führten Zuweisungen mit noch geringeren Kanalzahlen durch und erhielten als Resultat oftmals ungültige Zuweisungen. Die unteren Kanalzahlen liegen also nahe an der Zahl von Kanälen, die mindestens notwendig sind, damit die Separierungsbedingungen erfüllt werden können.

Die dritte bis fünfte Spalte geben Interferenzwerte wieder. Die Zahlen resultieren aus jeweils zehn Durchläufen der Frequenzzuweisung. In der dritten und vierten Spalte sind die maximalen und minimalen Interferenzsummen aller Durchläufe festgehalten. Die fünfte Spalte enthält den Durchschnittswert aller gültigen Durchläufe. In der letzten Spalte ist die Anzahl der Versuche festgehalten, bei denen der Zuweisungs-Algorithmus keine gültige Lösung fand. Er terminiert in einem solchen Fall mit einer Lösung, die Separierungsbedingungen verletzt.

Die optimierten Netze weisen bei der gleichen Zahl von Kanälen bedeutend weniger Interferenz auf als die Originalnetze. Betrachten wir die Zuweisungen mit 45 Kanälen für BPN1, 30 Kanälen für BPNe und 45 Kanälen für HAG1. In jedem der betrachteten Szenarios wurde für das optimierte Netz eine Frequenzzuweisung ermittelt, wo keine Interferenz auftritt. Die Ausgangsnetze wiesen dahingegen minimale Interferenzwerte zwischen 1,44 und 5,68 auf. Dass die interferenzfreien Zuweisungen für die optimierten Netze nicht „durch Zufall“ gefunden wurde, zeigen die Durchschnittswerte. Bei den optimierten Netzen liegen diese auf zwei Nachkommastellen gerundet bei null, bei den Ausgangsnetzen zwischen 2,37 und 6,42. Wir schließen daraus, dass die Optimierung ein Erfolg war. Ausgangsnetze, die für ei-

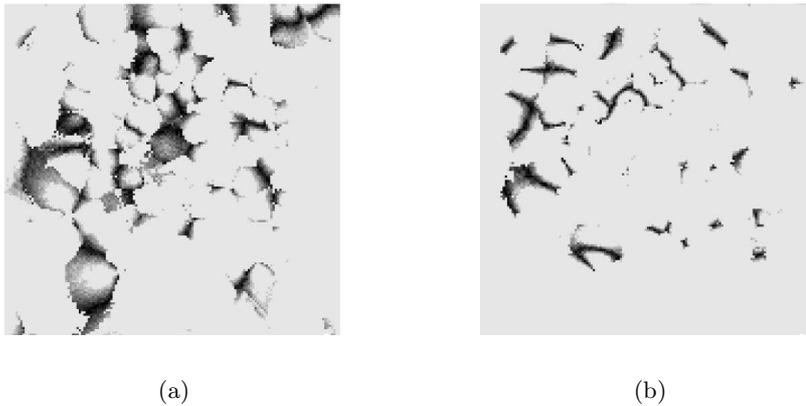
ne gegebene Zahl von Kanälen einen bestimmten positiven Interferenzwert aufweisen, können durch Tiltvariation nach unseren Bewertungskriterien so konfiguriert werden, dass keine nennenswerte Interferenz mehr auftritt.

Man beachte, dass die Zahl der Kanäle, die wir hier verwenden, im Verhältnis zur Zahl insgesamt zur Verfügung stehender Kanäle ziemlich klein ist. In GSM 900 stehen 124, in GSM 1800 stehen 374 Kanäle zur Verfügung (siehe [6]). Wir möchten die Zahl der verwendeten Kanäle trotzdem reduzieren, denn an unser Planungsgebiet grenzen möglicherweise andere Planungsgebiete an, wo dieselben Frequenzen wieder verwendet werden. Je weniger Frequenzen wir pro Planungseinheit benötigen, desto einfacher ist es hinterher, die Planung aneinandergrenzender Gebiete so zu koordinieren, dass in den Grenzgebieten keine Interferenzprobleme entstehen. Außerdem sei erwähnt, dass sich in Deutschland mehrere Anbieter die Frequenzbänder teilen. Es gibt durchaus Anbieter, denen nur so viele Kanäle zur Verfügung stehen, wie in unseren Analysen verwendet wurden.

Wir führten die Frequenzzuweisung ein zweites Mal mit weniger Kanälen durch. Das Resultat war, dass für die Ausgangsnetze in den meisten Fällen aufgrund der Separationsbedingungen keine gültigen Zuweisungen mehr ermittelt werden konnten. Die einzige Ausnahme bildet das Szenario BPNe, wo vier von zehn Zuweisungen gültig waren, jedoch inakzeptable Interferenzwerte aufwiesen. Die optimierten Netze weisen mit wenig Kanälen in allen Szenarien weniger Interferenz auf als die Ausgangsnetze mit vielen Kanälen.

Das Resultat der Optimierung ist also, dass durch Tiltvariation die Konfiguration eines Netzes so verbessert werden kann, dass bei einer kleineren Zahl verwendeter Kanäle weniger Interferenz auftritt.

Das Optimierungsergebnis ist in Abbildung 4.9 beispielhaft dargestellt. Wir sehen links das Ausgangsnetz des Szenarios BPN1 und rechts das optimierte Netz. Dargestellt ist das Verhältnis von Signal- zu Rauschleistung. Für jeden Pixel  $p$  wurde dazu der *best server* ermittelt und sein Pfadverlust. Dem *best server* sind eine Menge von Funkkanälen zugeordnet. Es wurden die Sender ermittelt, die einen dieser Funkkanäle benutzen, und ihr Pfadverlust zu Pixel  $p$  betrachtet. Je dunkler ein Pixel in den Bildern gefärbt ist, desto geringer ist die Differenz der Pfadverluste des *best servers* und des Störers. Sind die Pixel hellgrau gefärbt, so liegt das Signal- zu Rauschleistungsverhältnis über 9 dB. Damit sind die Signale nach dem in Abschnitt 4.1 erwähnten GSM-Standard [8] von den GSM-Empfängern entschlüsselbar. Auf den dunkler gefärbten Pixel ist das je nach Qualität der verwendeten Ausrüstung nur bedingt möglich. Es ist deutlich zu sehen, dass im optimierten Netz weit weniger Pixel Störungen aufweisen als im Ausgangsnetz. Durch die Optimierung konnte also die Abdeckung verbessert werden. Da die Abdeckung verbessert wurde, werden mehr Kunden erreicht. Damit versorgt das optimierte Netz mehr Kunden, womit sich seine Kapazität erhöht hat.



---

IST-2000-28088 Momentum, <http://momentum.zib.de>

Abbildung 4.9: Vergleich der Interferenz von Ausgangsnetz (a) und optimiertem Netz (b).

## 4.7 Zusammenfassung

In diesem Kapitel wurde ein erfolgreicher Ansatz zur Planung interferenzarmer GSM-Netze vorgestellt. Er beruht darauf, bei der Abdeckungsplanung Interferenzaspekte zu berücksichtigen. Das Interferenzpotenzial der verschiedenen Netzkonfigurationen wurde mit Kopplungsmatrizen beschrieben. Dass dieser Ansatz begründet ist, konnten wir zeigen.

Zunächst formulierten wir das Problem als nichtlineares Optimierungsprogramm. Zur Lösung transformierten wir es zunächst in ein pixelbasiertes, und dann in ein obermengenbasiertes ganzzahliges Programm. Beide Ansätze scheiterten an der Komplexität des Problems. Schließlich entwickelten wir heuristische Methoden, mit denen wir gute Lösungen erhielten. Die Heuristiken arbeiteten auf kleineren Probeinstanzen nachweislich optimal.

Wir zeigten, dass unsere Lösungen auf den von uns untersuchten Instanzen nach der Frequenzzuweisung deutlich weniger Interferenz aufwiesen als die Ausgangskonfigurationen.

Wir konnten nicht zeigen, dass die Heuristiken auch auf großen Instanzen optimal arbeiten. Hier wäre es interessant zu untersuchen, welche Ergebnisse mit enumerativen Algorithmen erreicht werden können. Die von uns beschriebene Startheuristik arbeitet so gut, dass man sie zur praktischen Arbeit weiterentwickeln sollte. Hier wäre interessant, ob man Fortschritte erzielen kann, wenn man die Zielfunktion variiert. Die Verfolgung des IP-Ansatzes halten wir für wenig vielversprechend. Zur besseren Abbildung der Kopplungsverhältnisse müsste die Formulierung der Beziehungsungleichungen deutlich straffer sein. Wir halten es für schwierig, hier Fortschritte zu erzielen.



## Kapitel 5

# Planung interferenzarmer UMTS-Netze

Nachdem wir in den Kapiteln 1 und 2 unseren allgemeinen Planungsansatz beschrieben haben, und ihn in Kapitel 4 auf die GSM-Technologie angewendet haben, stellen wir in diesem Kapitel unseren Planungsansatz für die UMTS-Technologie vor.

In Abschnitt 5.1 stellen wir ein mathematisches Modell vor, das Abdeckungsaspekte, Interferenz- und Kapazitätsaspekte von UMTS-Netzen abbildet. Dieses Modell wurde in [1] zur Netzbewertung von UMTS-Netzkonfigurationen entwickelt. Es verbindet Pfadverlust- und Verkehrsdaten und berücksichtigt Mehrwegeausbreitung, die Aktivität von Verbindungen und andere Daten. Wir beschreiben zunächst Eigenschaften der UMTS-Technologie, die zu so großen Unterschieden führen zwischen den Planungsaufgaben von UMTS- und GSM-Netzen. Daraufhin stellen wir das Planungsmodell vor. Es stellt eine Methode zur Verfügung, mittels des Verkehrsstatus im Netz den Interferenzzustand, und damit die Auslastung der Antennen zu ermitteln. Dies geschieht durch Lösen eines linearen Gleichungssystems. Die für dieses System hergeleitete Matrix bezeichnen wir als *Lastkopplungsmatrix*. Die Lösung des Systems gibt, sofern sie bestimmte Eigenschaften hat, eine Konfiguration von Leistungswerten der Antennen wieder. Der Planungsansatz, den daraufhin vorstellen, besteht darin, eine Netzkonfiguration zu ermitteln, deren Kopplungsmatrix kleine Matrixnormen hat. Wir formulieren die These, dass die Lösung des oben beschriebenen linearen Gleichungssystems umso besser ist, je kleiner die Matrixnormen sind. Da die Matrix nur nicht-negative Einträge hat, verwenden wir die Summe der Einträge als zu minimierende Matrixnorm.

In Abschnitt 5.2 beschreiben wir kurz, wie wir die in Abschnitt 5.1 formulierte Optimierungsaufgabe als IP formulieren. Das geschieht über ein nichtlineares 0/1-Programm, dessen nichtlineare Zielfunktion wir mit Hilfe von Beziehungsungleichungen linearisieren. Wir widmen uns dieser Formulierung nur kurz, denn es handelt sich um denselben Problemtyp, der

schon in Kapitel 4 ausführlich behandelt wurde. Auch hier zeigten unsere Untersuchungen, dass die Programme für große Instanzen nicht mit befriedigender Geschwindigkeit gelöst werden können. Wir verwendeten für die UMTS-Optimierung das heuristische Suchverfahren, das wir bei der GSM-Optimierung verwendeten.

In Abschnitt 5.3 stellen wir die Ergebnisse der UMTS-Optimierung vor. Zur Analyse verwenden wir Skalierungsfaktoren, mit denen wir die Zeilen der Kopplungsmatrix multiplizieren. Diese Faktoren erlauben uns, Leistungswerte für Netzkonfigurationen zu ermitteln, selbst wenn die Lösung des Gleichungssystems Werte aufweist, die größer sind als die maximalen Sendeleistungen oder kleiner sind als null. Die Ergebnisse unserer Optimierung sind nicht so eindeutig wie bei GSM. Wir führen das darauf zurück, dass die Skalierung mit den  $\lambda$ -Faktoren großen Einfluss auf die Analyse hat. Zwar werden diese Faktoren nach einem einheitlichen Verfahren ermittelt, aber es sind verschiedene Skalierungsverfahren denkbar, die vermutlich zu unterschiedlichen Ergebnissen führen. Unsere Methoden zeigen mit der in [7] vorgestellten Methode zur Ermittlung der Skalierungsvektoren in drei der vier betrachteten Netze Verbesserungen.

In Abschnitt 5.4 fassen wir dieses Kapitel zusammen.

## 5.1 Optimierungsansatz

Das Systemmodell, welches wir in diesem Abschnitt vorstellen, ist [1] entnommen. Es dient zur Bewertung des Systemzustands einer UMTS-Netzkonfiguration für eine gegebene Verkehrsverteilung. In 5.1.1 machen wir zunächst einige einleitende Bemerkungen zur UMTS-Technologie und den sich daraus ergebenden Planungsproblemen. In 5.1.2 führen wir einige Schreibweisen ein, erklären in das Modell eingehende Variablen und Parameter. Die Bezeichnungen wurden größtenteils aus [7] übernommen. In 5.1.3 konstruieren wir ein lineares Gleichungssystem, dessen Matrix wir als *Lastkopplungsmatrix* bezeichnen. Mit ihr können wir unter bestimmten Voraussetzungen Antennenleistungen so bestimmen, dass die CIR-Werte aller Empfänger mit Gleichheit erfüllt werden. Schließlich entwickeln wir in 5.1.4 Thesen zur Planung von UMTS-Netzen. Sie befassen sich damit, welche Struktur die Lastkopplungsmatrix haben sollte, damit sie zur Bestimmung günstiger Leistungskonfigurationen führt.

### 5.1.1 Eigenschaften der UMTS-Technologie

Wir machen in diesem Unterabschnitt einige Überlegungen zur UMTS-Technologie. Die darin beschriebenen Eigenschaften dieser Technologie werden sich in unserer mathematischen Modellierung wiederfinden.

In UMTS wird der *multiple access* durch Verwendung unterschiedlicher Verschlüsselungstechniken erreicht (*Wideband Code Division Multiple Access*, WCDMA) (siehe [19], [15]). Signale werden „gespreizt“, und damit

wird erreicht, dass sie das ganze in UMTS zur Verfügung stehende Frequenzband nutzen. Das hat zwei Effekte: Zum einen werden durch diese Technik hohe Datenraten erreicht, zum anderen können die Signale noch bei einem sehr niedrigen CIR-Wert entschlüsselt werden. Ein Sprachsignal kann durch dieses Verfahren aus einhundert und mehr gleichstarken Signalen herausgefiltert werden. Je größer die Datenrate eines Signals jedoch ist, desto anfälliger ist es für Interferenz. Das liegt daran, dass für Signale mit hohen Datenraten andere Spreizcodes verwendet werden müssen, und der daraus resultierende *spreading gain* geringer ist (siehe [19]). UMTS stellt verschiedene Dienste zur Verfügung, von denen einige sehr hohe Datenraten benötigen. Ein Beispiel hierfür ist Videotelefonie.

Interferenz entsteht in UMTS-Netzen sowohl innerhalb der eigenen Zelle (Intrazell-Interferenz) als auch durch Signale anderer Zellen (Interzell-Interferenz). Prinzipiell stören sich alle Sender im Netz, denn sie benutzen dieselbe Frequenz. Die Intrazell-Interferenz entsteht durch verschiedene Effekte. Einer dieser Effekte ist die Mehrwege-Ausbreitung, wo dasselbe Signal zeitverzögert mehrfach empfangen wird, weil es sich durch Reflexion an Gebäuden auf verschiedenen Wegen ausgebreitet hat.

Die Sendeleistung einer UMTS-Antenne ist variabel. Sie wird mit einer Frequenz von 1500 Hz an die augenblicklichen Verhältnisse angepasst. Dieser Vorgang heißt *power control*. Die Antenne erhält von jedem ihrer Empfänger Daten über die Qualität der empfangenen Signale. Wie hoch die Qualität der empfangenen Daten sein muss, hängt dabei vom Dienst ab. Ist die Qualität zu schlecht, so verwendet die Antenne auf diese Verbindung eine höhere Sendeleistung. Die *Auslastung* einer Antenne kann am Verhältnis zwischen ihrer momentanen Sendeleistung und ihrer maximalen Sendeleistung gemessen werden.

Kapazitätsplanung und Interferenzplanung sind in UMTS-Netzen eng miteinander verknüpft. Eine UMTS-Antenne muss einen Teil ihrer Sendeleistung dazu verwenden, die Intrazell-Interferenz zu überwinden. Je größer der momentane Verkehr ist, desto größer ist die Auslastung der Antenne. Je größer die Auslastung der Antenne ist, desto mehr Interzell-Interferenz erzeugt sie. Die Menge der Interzell-Interferenz im Netz hängt also vom Verkehr ab. Die Auslastung der Antennen hängt wiederum von der Menge der Interzell-Interferenz ab. Mit wie vielen Empfängern eine Antenne kommunizieren kann, hängt also von der Interferenzsituation ab. Ist das Netz sehr anfällig für Interferenz, so geht ein großer Teil der Kapazität verloren, weil die Antennen für ihre wenigen Verbindungen ihre gesamte verfügbare Leistung aufwenden müssen. Eine Motivation für Interferenzplanung in UMTS ist also die Erhöhung der Kapazität.

Abdeckungs- und Interferenzplanung sind ebenfalls eng miteinander verknüpft. UMTS-Antennen versenden ein *Pilotsignal*, damit Kunden Kontakt mit ihnen aufnehmen können. Ob die Empfänger das Pilotsignal entschlüsseln können, hängt von der Qualität des Pilotsignals ab, welche wie-

derum von der Interferenzsituation abhängt. Ist diese schlecht, so geht in Teilen des Netzes die Abdeckung verloren, weil Empfänger nicht in der Lage sind, die eingehenden Pilotsignale zu erkennen. Neben der  $E_c$ -Abdeckung – der Abdeckung, wo allein die Signalstärke bewertet wird – spielt also auch die  $E_c/I_0$ -Abdeckung eine Rolle, wo die *Qualität* der Abdeckung bewertet wird.

### 5.1.2 Planungsgrößen

Wir erläutern nun das in [1] eingeführte UMTS-Modell. Gegeben sei eine Menge  $I$  von Installationen, mit einer Partition  $\mathcal{S}$  in Sektoren. Im Planungsgebiet befindet sich eine Menge  $M$  von mobilen Nutzern (*mobile stations*). Den Dämpfungsfaktor zwischen einer Installation  $i \in I$  und einer *mobile station*  $m \in M$  bezeichnen wir mit  $\gamma_{i,m}^\downarrow$ .

Wir unterscheiden drei Arten von Sendesignalen: Das Pilotsignal, das zur Kontaktaufnahme dient, andere zur Steuerung der Kommunikation notwendige Signale und die Signale, die die eigentlichen Daten übertragen. Die Stärke des Pilotsignals bezeichnen wir mit  $\hat{p}^\downarrow$ , die Stärke der anderen Verkehrssignale mit  $\check{p}^\downarrow$ . Die Gesamtleistung einer Zelle  $i$  bezeichnen wir mit  $\bar{p}_i^\downarrow$ . Die maximale Sendeleistung von Zelle  $i$  heißt  $\Pi_i^{\max\downarrow}$ .

Jeder *mobile station*  $m \in M$  wird ein dienstabhangiger *Aktivitatsfaktor*  $0 \leq \alpha_m \leq 1$  zugeordnet. Dieser Faktor spiegelt den Anteil der Zeit wider, in dem die Zelle, die mit  $m$  verbunden ist, Signale ubermittelt. Eine Sprachverbindung hat in unseren Szenarien beispielsweise einen Aktivitatsfaktor von 0.5 zu, da der Empfanger in etwa der Halfte der Zeit zuhort. Eine Verbindung mit ununterbrochenem Datenstrom hat einen Aktivitatsfaktor von eins. Die Aktivitat einer Verbindung spielt deswegen eine Rolle, weil sie in Ruhephasen keinen nennenswerten Anteil an der Sendeleistung der Zelle in Anspruch nimmt.

Wir bezeichnen den Wert des *Rauschens*, dem eine *mobile station*  $m$  ausgesetzt ist, mit  $\eta_m$ . Dieser Wert setzt sich aus verschiedenen Werten, unter anderem dem des *thermischen* Rauschens zusammen.

Eine Antenne versieht jedes Signal mit einem eigenen Spreizcode, der orthogonal ist zu den Spreizcodes der anderen Signale (siehe [19]). Dadurch storen sich die Signale der gleichen Antenne theoretisch nicht gegenseitig. In der Praxis jedoch tragen viele Faktoren dazu bei, dass die Signale der Antenne sich gegenseitig beeinflussen (Intrazell-Interferenz). Dazu gehort zum Beispiel die Ausbreitung des Signals uber verschiedene Wege durch Reflexion an Gebauden. Diesem Fakt tragen wir dadurch Rechnung, dass wir den *mobile stations* einen Orthogonalitatswerte  $0 \leq \omega_m \leq 1$  zuweisen. Gelangen die Signale vom Sender *mobile station*  $m$ , ohne von ihrer Qualitat zu verlieren so setzen wir  $\omega_m = 1$ . Beeintrachtigen sich die Signale, so geben wir  $\omega_m$  je nach Ausma der Storung einen kleineren Wert als eins.

Eine *mobile station*, die mit einer Zelle verbunden ist, benotigt je nach Dienst eine bestimmte Signalqualitat, um die Eingangssignale entschlusseln

zu können. Die verlangte Qualität des Signals wird durch einen CIR-Zielwert  $\mu_m$  (*carrier-to-interference ration*) wiedergegeben. Wir möchten in unserem Modell eine Netzkonfiguration ermitteln, für die eine Einstellung der Sendeleistungen existiert, sodass das Verhältnis von Sendesignalen zu Rauschsignalen in jeder *mobile station* den Wert  $\mu$  annimmt:

$$\frac{\text{Empfangsleistung}}{\sum \text{Störleistung} + \text{Rauschen}} = \mu \quad (5.1)$$

### 5.1.3 Lastkopplungsmatrizen

Wir stellen nun die in [1] erstmals entwickelten Lastkopplungsmatrizen vor. Sie enthalten Lastkopplungsfaktoren, die den Einfluss von jeweils einer Zelle auf die andere reflektieren. Am Ende dieses Abschnittes erhalten wir ein Gleichungssystem, dessen Lösung Leistungswerte für die Zellen enthält. Versenden alle Antennen ihre Signale mit genau dieser Sendeleistung, so sind die CIR-Bedingungen 5.1 für alle *mobile stations* erfüllt. Mit diesen Antennenleistungen kann der prognostizierte Verkehr also vollständig bedient werden.

Wir leiten die Kopplungsmatrizen zunächst auf Basis von *snapshots* her. Das bedeutet, dass, wie oben beschrieben, eine Menge von *mobile stations* im Planungsgebiet verteilt ist. Auf der Basis ihrer Standorte und der von ihnen angeforderten Dienste werden die Kopplungsverhältnisse im Netz berechnet. Später stellen wir das Modell so um, dass es mit Verkehrsprognosen arbeitet, wie sie in Abschnitt 2 vorgestellt wurden. Für jeden Dienst beschreibt dann eine Flächenfunktionen die Intensitätswerte für jeden Pixel.

Man beachte, dass bei der Störleistung auch Störung vom eigenen Sender vertreten ist, die durch die Signale erzeugt wird, die für andere Empfänger innerhalb der Zelle bestimmt sind.

Formulieren wir die Gleichung 5.1 nun mit unseren Parametern. Sei Antenne  $i$  mit *mobile station*  $m$  verbunden. Die von  $i$  für  $m$  bestimmte abgestrahlte Leistung ist

$$p_{im}^\downarrow.$$

Die Empfangsleistung dieses Signals (Signalstärke) ist  $\gamma_{im}^\downarrow p_{im}^\downarrow$ . Die Stärke der bei *mobile station*  $m$  eintreffenden Störsignale ist

$$\gamma_{im}^\downarrow \bar{\omega}_m (\bar{p}_j^\downarrow - \alpha_m^\downarrow p_{im}^\downarrow) + \sum_{j \neq i} \gamma_{jm}^\downarrow \bar{p}_i^\downarrow + \eta_m,$$

wobei gilt  $\bar{\omega}_m := 1 - \omega_m$ . Der Summand

$$\gamma_{im}^\downarrow \bar{\omega}_m (\bar{p}_j^\downarrow - \alpha_m^\downarrow p_{im}^\downarrow)$$

gibt hierbei die Leistung wieder, die von Antenne  $i$  für andere *mobile stations* abgegeben wird und *mobile station*  $m$  erreicht. Im Idealfall sind die empfangenen Signale orthogonal zum eigenen Empfangssignal, und der Wert von

$M$		Menge der <i>mobile stations</i>
$M_i$	$\subset M$	Menge der von Installation $i$ bedienten <i>mobile stations</i>
$p_{im}^\downarrow$	$\in \mathbb{R}^+$	Für <i>mobile station</i> $m$ aufgewendete Leistung der Installation $i$
$\bar{p}_j^\downarrow$	$\in \mathbb{R}^+$	Sendeleistung Installation $i$
$\hat{p}_j^\downarrow$	$\in \mathbb{R}^+$	Pilotleistung Installation $i$
$\check{p}_j^\downarrow$	$\in \mathbb{R}^+$	Leistung für gemeinsame Kanäle Installation $i$
$\Pi_i^{\max\downarrow}$	$\in \mathbb{R}^+$	Maximale Sendeleistung Installation $i$
$\gamma_{im}^\downarrow$	$\in [0, 1]$	Pfadverlust von Installation $i$ zu <i>mobile station</i> $m$
$\omega_m$	$\in [0, 1]$	Orthogonalitätswert <i>mobile station</i> $m$
$\bar{\omega}_m$	$\in [0, 1]$	$1 - \omega_m$
$\alpha_m^\downarrow$	$\in [0, 1]$	Aktivitätsfaktor <i>mobile station</i> $m$
$\eta_m$	$\in \mathbb{R}^+$	Thermische Rauschleistung <i>mobile station</i> $m$
$\mu_m^\downarrow$	$\geq 0$	CIR-Zielwert <i>mobile station</i> $m$
$l_m^\downarrow$	$\in \mathbb{R}^+$	<i>Downlink</i> -Lastfaktor <i>mobile station</i> $m$
$C_i^\downarrow$	$\in \mathbb{R}^+$	Eigenkopplung Installation $i$
$C_i^{\downarrow j}$	$\in \mathbb{R}^+$	Kopplungswert von Installation $j$ für Installation $i$
$p_i^{\downarrow\eta}$	$\in \mathbb{R}^+$	Verkehrs-Rauschleistung Installation $i$

Abbildung 5.1: Abkürzungen.

$\bar{\omega}_m$  ist null. Dann haben die für Empfänger in der eigenen Zelle bestimmten Signale keinen störenden Einfluss.

Der Summand

$$\sum_{j \neq i} \gamma_{jm}^\downarrow \bar{p}_i^\downarrow$$

ist die von den anderen Antennen erzeugte Störleistung. Die Signale verschiedener Antennen sind nicht mit zueinander orthogonalen Spreizcodes versehen, daher verwenden wir für die von ihnen erzeugte Interferenz implizit Orthogonalitätswerte von eins. Der Wert  $\eta_m$  gibt die Stärke der übrigen Rauschsignale wieder.

Gleichung 5.1 stellt sich dann wie folgt dar:

$$\frac{\gamma_{im}^\downarrow p_{im}^\downarrow}{\gamma_{im}^\downarrow \bar{\omega}_m (\bar{p}_j^\downarrow - \alpha_m^\downarrow p_{im}^\downarrow) + \sum_{j \neq i} \gamma_{jm}^\downarrow \bar{p}_i^\downarrow + \eta_m} = \mu_m^\downarrow. \quad (5.2)$$

Stellen wir Gleichung 5.2 für jede Antenne und für jeden Pixel in ihrer Zelle auf, so erhalten wir ein Gleichungssystem, mit dem die Sendeleistungen ermittelt werden können. Wir werden dieses System nun vereinfachen. Das Resultat wird ein System sein, aus dem die  $p_{im}^\downarrow$ -Variablen eliminiert wurden.

Die Stärke der von Antenne  $i$  abgestrahlten Signale ist

$$\bar{p}_i^\downarrow = \sum_{m \in M_i} \alpha_m^\downarrow p_{im}^\downarrow + \hat{p}_j^\downarrow + \check{p}_j^\downarrow. \quad (5.3)$$

Durch Substitution der Gleichungen vom Typ 5.2 können wir die Variablen vom Typ  $p_{iq}^\downarrow$  eliminieren und erhalten ein System in den  $\bar{p}_i^\downarrow$ -Variablen. Hierzu stellen wir die Gleichungen vom Typ 5.2 zunächst um:

$$\alpha_m^\downarrow p_{im}^\downarrow = \frac{\alpha_m^\downarrow \mu_m^\downarrow}{1 + \bar{\omega}_m \alpha_m^\downarrow \mu_m^\downarrow} (\bar{\omega}_m \bar{p}_i^\downarrow + \sum_{j \neq i} \frac{\gamma_{jm}^\downarrow}{\gamma_{im}^\downarrow} \bar{p}_j^\downarrow + \frac{\eta_m}{\gamma_{im}^\downarrow}). \quad (5.4)$$

Der Faktor

$$l_m^\downarrow := \frac{\alpha_m^\downarrow \mu_m^\downarrow}{1 + \bar{\omega}_m \alpha_m^\downarrow \mu_m^\downarrow} \quad (5.5)$$

ist hierbei unabhängig von der Antenne, die *mobile station*  $m$  bedient. Wir bezeichnen ihn als *Downlink-Last* der mobile station  $m$ . Er ist ein Maß für die Last, die von *mobile station*  $m$  auf Antenne  $i$  erzeugt wird. Da der Wert  $\bar{\omega}_m \alpha_m^\downarrow \mu_m^\downarrow$  normalerweise deutlich kleiner als eins ist, hängt der Wert von  $l_m$  in erster Näherung nur vom Zähler, und damit von den Werten  $\mu_m$  und  $\alpha_m$  ab, die den verwendeten Dienst, also indirekt den Verkehr wiedergeben.

Nach der Substitution von Gleichung 5.4 in 5.3 erhalten wir die Gleichung

$$\begin{aligned} \bar{p}_i^\downarrow &= \sum_{m \in M_i} l_m^\downarrow (\bar{\omega}_m \bar{p}_i^\downarrow + \sum_{j \neq i} \frac{\gamma_{jm}^\downarrow}{\gamma_{im}^\downarrow} \bar{p}_j^\downarrow + \frac{\eta_m}{\gamma_{im}^\downarrow}) + \hat{p}_j^\downarrow + \check{p}_j^\downarrow \quad (5.6) \\ &= \sum_{m \in M_i} l_m^\downarrow \bar{\omega}_m \bar{p}_i^\downarrow + \sum_{j \neq i} (\sum_{m \in M_i} l_m^\downarrow \frac{\gamma_{jm}^\downarrow}{\gamma_{im}^\downarrow}) \bar{p}_j^\downarrow + \sum_{m \in M_i} l_m^\downarrow \frac{\eta_m}{\gamma_{im}^\downarrow} + \hat{p}_j^\downarrow + \check{p}_j^\downarrow \quad (5.7) \end{aligned}$$

Nach Einführung der zwei *Lastwerte* – oder *Kopplungswerte* (*coupling factors*) –

$$C_i^\downarrow := \sum_{m \in M_i} l_m^\downarrow \bar{\omega}_m \quad \text{und} \quad C_i^{\downarrow j} := \sum_{m \in M_i} l_m^\downarrow \frac{\gamma_{jm}^\downarrow}{\gamma_{im}^\downarrow}$$

und der *Verkehrs-Rauschleistung*

$$p_i^{\downarrow \eta} := \sum_{m \in M_i} l_m^\downarrow \frac{\eta_m}{\gamma_{im}^\downarrow},$$

die Sendeleistung für die eigenen Nutzer darstellt, die zur Neutralisierung des Rauschen notwendig ist, erhält diese Gleichung die übersichtliche Form

$$\bar{p}_i^\downarrow = C_i^\downarrow \bar{p}_i^\downarrow + \sum_{j \neq i} C_i^{\downarrow j} \bar{p}_j^\downarrow + p_j^{\downarrow \eta} + \hat{p}_j^\downarrow + \check{p}_j^\downarrow. \quad (5.8)$$

Den Kopplungsfaktor  $C_i^\downarrow$  interpretieren wir als den Anteil der Sendeleistung, der von Antenne  $i$  zur Überwindung der Intrazell-Interferenz aufgebracht werden muss. Je größer der Kopplungsfaktor  $C_i^{\downarrow j}$  ist, desto mehr Sendeleistung muss zur Überwindung der von Antenne  $j$  erzeugten Störungen aufgebracht werden. Zur besseren Übersicht geben wir das lineare System in Punkt-Strich-Form an.

$$\begin{pmatrix} \bar{p}_1^\downarrow \\ \bar{p}_2^\downarrow \\ \vdots \\ \bar{p}_n^\downarrow \end{pmatrix} = \underbrace{\begin{pmatrix} C_1^\downarrow & C_1^{\downarrow 2} & \dots & C_1^{\downarrow n} \\ C_2^{\downarrow 1} & C_2^\downarrow & & \\ \vdots & & \ddots & \vdots \\ C_n^{\downarrow 1} & \dots & & C_n^\downarrow \end{pmatrix}}_C \cdot \begin{pmatrix} \bar{p}_1^\downarrow \\ \bar{p}_2^\downarrow \\ \vdots \\ \bar{p}_n^\downarrow \end{pmatrix} + \begin{pmatrix} p_1^{\downarrow \eta} \\ p_2^{\downarrow \eta} \\ \vdots \\ p_n^{\downarrow \eta} \end{pmatrix} + \begin{pmatrix} \hat{p}_1^\downarrow \\ \hat{p}_2^\downarrow \\ \vdots \\ \hat{p}_n^\downarrow \end{pmatrix} + \begin{pmatrix} \check{p}_1^\downarrow \\ \check{p}_2^\downarrow \\ \vdots \\ \check{p}_n^\downarrow \end{pmatrix}. \quad (5.9)$$

Die Matrix  $C$  nennen wir *Kopplungsmatrix*.

Wie man sehen kann, können mit der Kopplungsmatrix Leistungswerte für die Antennen berechnet werden, sodass alle *mobile stations* ihren CIR-Zielwert erreichen. Dieser Bewertungsansatz hat zwei Hauptnachteile, die in [7] bereits formuliert wurden: Ist das Netz in Überlast, können die Antennen die Menge der *mobile stations* also nicht versorgen, so sind die Ergebnisse der Leistungsberechnung wertlos. Außerdem müssen für sinnvolle Aussagen über die Netzqualität hunderte oder gar tausende von *snapshots* evaluiert werden. Aus diesen Gründen gehen wir hier über zu einem pixelbasierten Modell. Für jeden Pixel werden Lastwerte für die einzelnen Dienste definiert. Diese Werte vereinigen gewissermaßen das Wissen vieler *snapshots* durch Bildung von Durchschnittswerten.

Sei  $P$  das Planungsgebiet und  $P^i$  der *best server*-Bereich der Zelle  $i$ . Sei  $S$  die Menge der betrachteten Dienste. Die Funktion  $T_s : P \rightarrow \mathbb{R}^+$  ordnet jedem Pixel ein durchschnittliches Verkehrsaufkommen für Dienst  $s$  zu. Wir nehmen an, dass jeder Dienst konstante CIR-Zielwerte  $\mu_s^\downarrow$  und Aktivitätsfaktoren  $\alpha_s^\downarrow$  aufweist. Die  $\gamma$ - und  $\omega$ -Werte seien nun abhängig vom betrachteten Pixel, nicht von der betrachteten *mobile station*. Dann definieren wir die *Downlink*-Last des Pixels  $q$  als

$$l_q^\downarrow := \sum_{s \in S} \frac{\alpha_s^\downarrow \mu_s^\downarrow}{1 + \bar{\omega}_q \alpha_s^\downarrow \mu_s^\downarrow} T_s(q). \quad (5.10)$$

Die Lastkopplungsfaktoren sind definiert als

$$C_i^\downarrow := \sum_{q \in P_i} l_q^\downarrow \bar{\omega}_q \quad \text{und} \quad C_i^{\downarrow j} := \sum_{q \in P_i} l_q^\downarrow \frac{\gamma_{jq}^\downarrow}{\gamma_{iq}^\downarrow}. \quad (5.11)$$

Die Verkehrs-Rauschleistung schließlich ist definiert als

$$p_i^{\downarrow \eta} := \sum_{q \in P_i} l_q^\downarrow \frac{\eta_q}{\gamma_{iq}^\downarrow}.$$

Die pixelbasierte Version der Kopplungsmatrix hat den Vorteil, viele *snapshots* zu vereinigen und damit eine Netzbewertung zu erlauben, die verschiedene Verkehrssituationen berücksichtigt. In unseren Optimierungsverfahren verwendeten wir diese Version.

#### 5.1.4 Thesen zur Planung interferenzarmer UMTS-Netze

In diesem Abschnitt machen wir einige Beobachtungen über das im letzten Abschnitt hergeleitete Gleichungssystem. Aus diesen Beobachtungen leiten wir einen Ansatz für die Planung interferenzarmer UMTS-Netze ab. Wir beschreiben einen Fall, wann System 5.9 lösbar ist, und welche Eigenschaften die Lösung hat. Die in der Lösung enthaltenen Sendeleistungen sollen innerhalb des Leistungsspektrums der jeweiligen Antennen liegen. Je kleiner die Norm der Kopplungsmatrix ist, desto kleiner ist unsere Abschätzung für die Norm des Lösungsvektors. Unser Optimierungsansatz lautet, die Summe der Matrixeinträge zu minimieren, um damit die Norm der Matrix und letzten Endes die Norm des Lösungsvektors zu reduzieren.

Betrachten wir das Gleichungssystem 5.9 ein wenig genauer. Definieren wir

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} := \begin{pmatrix} p_1^{\downarrow\eta} \\ p_2^{\downarrow\eta} \\ \vdots \\ p_n^{\downarrow\eta} \end{pmatrix} + \begin{pmatrix} \hat{p}_1^{\downarrow} \\ \hat{p}_2^{\downarrow} \\ \vdots \\ \hat{p}_n^{\downarrow} \end{pmatrix} + \begin{pmatrix} \check{p}_1^{\downarrow} \\ \check{p}_2^{\downarrow} \\ \vdots \\ \check{p}_n^{\downarrow} \end{pmatrix} \quad (5.12)$$

so erhalten wir die Formulierung

$$(E_n - C) \cdot \bar{p}^{\downarrow} = b, \quad (5.13)$$

wobei  $E_n$  die  $n$ -dimensionale Einheitsmatrix ist. Sowohl  $C$  als auch  $b$  hängen in komplexer Weise von der Netzkonfiguration ab. Es ist also schwer zu sagen, welche Veränderungen an der Konfiguration vorgenommen werden müssen, damit System 5.13 eine gute Lösung hat. Gut bedeutet dabei, dass die  $\bar{p}^{\downarrow}$ -Werte im Bereich der möglichen Antennenleistungen liegen. Ist das der Fall, so werten wir das als Indiz dafür, dass das Netz für den gegebenen Verkehr ausreichende Kapazität aufweist. Man beachte jedoch, dass sich diese Sendeleistungen in der Realität vermutlich nicht so einstellen werden. Das würde eine zentrale Regulierung der Antennenleistungen voraussetzen. Die Antennenleistungen orientieren sich jedoch hauptsächlich an den Rückmeldungen der eigenen Empfänger. Außerdem wurde das Verkehrsaufkommen, dessen Intensität sich in Realität gegebenenfalls schnell verändert, durch die Aktivitätsfaktoren geglättet.

Einen Anhaltspunkt für die Optimierung liefert das folgende Resultat (siehe [16]):

**Satz 5.1.** Sei  $C \in \mathbb{R}^{n \times n}$ ,  $\|\cdot\|$  eine submultiplikative (siehe [18]) Norm und  $E_n \in \mathbb{R}^{n \times n}$  die Einheitsmatrix. Falls  $\|C\| < 1$ , so ist  $E_n - C$  nichtsingulär und es gilt

$$(E_n - C)^{-1} = \sum_{i=0}^{\infty} C^i \quad (5.14)$$

sowie

$$\|(E_n - C)^{-1}\| \leq \frac{1}{1 - \|C\|}. \quad (5.15)$$

Die Spalten- und die Zeilensummennormen sind konsistente Matrixnormen. Hat also jede Spaltensumme in  $C$  einen Wert von weniger als eins, so existiert ein eindeutiger Lösungsvektor für die Leistungen der Installationen im Netz. Da  $C$  keine negativen Werte enthält, hat die Matrix  $(E_n - C)^{-1}$  gemäß Gleichung 5.14 ausschließlich nicht-negative Werte, und damit der Lösungsvektor  $\bar{p}^\downarrow = (E_n - C)^{-1} \cdot b$ . Mit Hilfe des Vektors  $b$  erhalten wir unter Verwendung der Normen der Matrix eine Abschätzung über die Norm und damit die Elemente des Lösungsvektors  $p$ . Diese Erkenntnisse möchten wir noch einmal in den folgenden Thesen wiedergeben, auf die wir unseren Optimierungsansatz stützen.

**These 5.1.** Je kleiner die Summe der Nebendiagonaleinträge der Matrix  $C$  ist, desto kleiner sind in der Tendenz Spalten- und Zeilensummennorm.

Man beachte, dass wir hier absichtlich etwas vage bleiben. Selbstverständlich können diese Normen vergleichsweise groß sein, wenn zum Beispiel nur ein einziger Eintrag ungleich null ist. Wir gehen aber davon aus, dass Werte der Normen der Kopplungsmatrizen mit der Summe der Matrixeinträge einigermaßen zusammenhängt.

**These 5.2.** Die Norm des Vektors  $b$  verändert sich bei Änderung der Netzkonfiguration relativ wenig. Wir können sie als konstant ansehen.

**These 5.3.** Mit Hilfe der Normen von  $b$  und  $C$  gelangen wir zu einer Abschätzung der Norm von  $p$ . Gilt  $\|C\| < 1$ , so folgt

$$\|p\| \approx \frac{\|b\|}{1 - \|C\|}$$

Je kleiner dieser Schätzwert ist, desto besser.

Wir können diese Abschätzung treffen, weil Zeilen- und Spaltensummennorm *verträglich* (siehe [18]) sind. Die letzte These besagt, dass es besser ist, wenn die Abschätzung für  $\|p\|$  klein ist. In Abschnitt 5.3 werden wir die Sendeleistungen als Indikator für die Qualität der Netzkonfiguration einführen. Je kleiner die Norm des Lösungsvektors ist, desto kleiner sind tendenziell seine Einträge.

Unser Optimierungsansatz besteht darin, eine Netzkonfiguration zu ermitteln, wo die Normen der Matrix  $C$  möglichst klein sind. Dies erhoffen wir dadurch zu erreichen, dass wir die Summe der Matrixelemente von  $C$  minimieren. Eine Reduktion dieser Summe ist jedoch nur auf den Nebendiagonalen möglich. Denn wird jeder Pixel eindeutig einem *best server*-Bereich  $A_i$  zugeordnet, so ist die Spur der Matrix konstant.

$$\text{Spur}(C) = \sum_{i \in I} C_i^\downarrow = \sum_{i \in I} \sum_{q \in A_i} l_q^\downarrow \bar{\omega}_q = \sum_{q \in P} l_q^\downarrow \bar{\omega}_q$$

Eine Änderung der Netzkonfiguration bewirkt also nur eine Verschiebung der Lastfaktoren auf der Hauptdiagonalen von  $C$ . Die Summe der  $C_i^{\downarrow j}$ -Kopplungsfaktoren läßt sich jedoch reduzieren. Das Netz muss dabei so konfiguriert werden, dass die  $\gamma/\gamma$ -Quotienten (siehe Definition 5.11) möglichst klein werden. Das heißt, dass der *best server* auf möglichst vielen Pixeln ein deutlich stärkeres Signal senden soll als die Störer.

Unsere Betrachtungen haben uns zu einem Optimierungsansatz geführt, der dem der GSM-Planung sehr ähnlich ist. Dass die Kapazitäts- und die Interferenzplanung hier nicht von der Abdeckungsplanung entkoppelt ist, spiegelt sich dabei in der Kopplungsmatrix wider. Die Kapazitätsanforderungen gehen dadurch ein, dass in den Kopplungswerten die *Downlink*-Lastwerte  $l$  enthalten sind. In ihnen wird durch die  $T_s(q)$ -Werte der Verkehr wiedergegeben. Die Stärke der Interferenz geht durch das  $\gamma/\gamma$ -Verhältnis in den Kopplungsfaktoren ein. Bei GSM-Netzen interessierte uns nur, dass Störung auftrat. Diese wurde anschließend durch die Frequenzplanung reduziert. Bei UMTS interessiert uns außerdem die *Größe* der Störung. Hier geht sie direkt in die Beurteilung der Netzkonfiguration mit ein, ohne dass eine Frequenzplanung vorgenommen wird.

Unser Planungsansatz beruht auf der Annahme, dass es gut ist, wenn das gegebene Verkehrsaufkommen mit geringen Antennenleistungen bedient werden kann. Daher sind wir am Lösungsvektor von System 5.11 und infolgedessen nach Satz 5.1 an den Normen der Kopplungsmatrix interessiert. Unsere Optimierungsthese besagen, dass wir durch die Minimierung der Summe der Kopplungswerte Netzkonfigurationen mit guten Leistungswerten erhalten.

## 5.2 Formulierung als ganzzahliges Programm

Wir zeigen in diesem Abschnitt, wie man die UMTS-Optimierungsaufgabe aus Abschnitt 5.1 in ein IP transformiert. Zunächst formulieren wir sie als 0/1-Programm mit nichtlinearer Zielfunktion, die wir nachfolgend linearisieren. Das IP hat dieselbe Struktur wie das in Kapitel 4 vorgestellte Obermengenmodell. Wir können nach kleineren Anpassungen also die dort vorgestellten Lösungsalgorithmen verwenden. Die hier verwendete Notation ist teilweise aus Kapitel 4 übernommen.

Ziel der Optimierung ist die Minimierung der Summe der Nebendiagonalelemente der Matrix  $C$ . Man beachte, dass der Lastkopplungsfaktor

$$C_i^{\downarrow j} = \sum_{q \in A_i} l_q^{\downarrow} \frac{\gamma_{jq}^{\downarrow}}{\gamma_{iq}^{\downarrow}}.$$

gemäß Definition 5.11 auf allen Pixeln des *best server*-Bereichs  $A_i$  ermittelt wird, also kein zu Kapitel 4 analoger  $\delta$ -Schwellwert existiert. Aus praktischen und numerischen Gründen rechnen wir in unseren Modellen allerdings mit einem Schwellwert von 25 dB, was einem Verhältnis der Dämpfungsfaktoren von 316 entspricht. Dadurch erhalten wir auch hier  $Z^u$ -Mengen, analog zu denen in Kapitel 4. Welchen Einfluss die Einführung eines Schwellwertes auf die Lösung des linearen Gleichungssystems 5.9 hat, kann man durch Bestimmung der Konditionszahl (siehe [18]) der Kopplungsmatrix abschätzen. Wir untersuchten diese Auswirkungen nicht, zumal die Werte der Dämpfungsfaktoren vom präferierten Wellenausbreitungsmodell abhängen.

Wir definieren

$$A(i, j, Z) := \{p \in P^i \mid Z^o(i, p) = Z, j \in Z^u(i, p)\} \quad (5.16)$$

als die Menge der Pixel, wo Installation  $i$  die Obermenge  $Z$  hat und von Installation  $j$  gestört wird. Der Wert

$$c_{i,j}^Z := \sum_{q \in A(i,j,Z)} l_q^{\downarrow} \frac{\gamma_{jq}^{\downarrow}}{\gamma_{iq}^{\downarrow}} \quad (5.17)$$

gibt den Anteil der Kopplung wieder, der zwischen Installation  $i$  und Installation  $j$  auf der Menge  $A(i, j, Z)$  induziert wird, falls  $i$  und  $j$  beide benutzt werden, und aus  $Z$  keine Installation gewählt wurde. Man beachte, dass die Definition derjenigen in Gleichung 4.18 sehr ähnlich ist. Für die GSM-Planung wurde nur für jeden Pixel ein gleicher Beitrag von 1 zur Kopplung angenommen, hier hängt dieser Beitrag vom Verhältnis der Sendeleistungen der beiden Sender und der *Downlink*-Last auf  $q$  ab. Für die Kopplung zweier Installationen gilt

$$\begin{aligned} C_i^{\downarrow j}(N) &= \sum_{N \cap Z = \emptyset} c_{i,j}^Z \cdot z_i \cdot z_j \\ &= \sum_{Z \in Z(i)} c_{i,j}^Z \cdot z_i \cdot z_j \cdot \prod_{k \in Z} (1 - z_k) \end{aligned} \quad (5.18)$$

Das nichtlineare UMTS-Optimierungsproblem lautet:

$$\min \sum_{i \in I} \sum_{Z \in Z(i)} \sum_{j \in I \setminus i} c_{i,j}^Z \cdot z_i \cdot z_j \cdot \prod_{k \in Z} (1 - z_k)$$

sodass

$$\begin{aligned}
\sum_{i \in \bar{I}_p} z_i &\geq 1 & \forall p \in \bar{P} \\
\sum_{i \in S} z_i &= 1 & \forall S \in \mathcal{S} \\
z_i &\in \{0, 1\}
\end{aligned} \tag{5.19}$$

Die daraus nach Fortet abgeleitete lineare Formulierung lautet

$$\min \sum_{i \in I} \sum_{Z \in Z(i)} \sum_{j \in I \setminus i} c_{i,j}^Z \cdot y(i, j, Z)$$

sodass

$$\begin{aligned}
z_i + z_j - \sum_{k \in Z} z_k - 1 &\leq y(i, j, Z) & \forall i \neq j \in I, Z \in Z(i) \\
0 &\leq y(i, j, Z) & \forall i \neq j \in I, Z \in Z(i) \\
\sum_{i \in \bar{I}_p} z_i &\geq 1 & \forall p \in \bar{P} \\
\sum_{i \in S} z_i &= 1 & \forall S \in \mathcal{S} \\
z_i &\in \{0, 1\} \\
y(i, j, Z) &\in \{0, 1\} & \forall i \neq j \in I, Z \in Z(i)
\end{aligned} \tag{5.20}$$

Ein ähnliches IP wurde bereits in [7] vorgestellt, wobei die Zielfunktion hier anders gewählt ist und die Beziehungsungleichungen anders formuliert sind. Außerdem ist das Modell in [7] flächenbasiert, während unseres pixelbasiert ist.

Die Größe der Gleichungssysteme ist mit denen im GSM-Ansatz vergleichbar. Daher ist auch hier die Lösung mit einem IP-Löser auf großen Szenarien schwierig. Mathematisch gesehen ist das IP das gleiche, weswegen dieselben Schnittebenen definiert werden können. Zur Optimierung der UMTS-Netze verwendeten wir einen lokalen Suchalgorithmus, der dem in Kapitel 4 beschriebenen ähnelt. Die Unterschiede bestehen darin, dass die Bewertung der Netze während des Algorithmus anders implementiert ist, und dass wir in der Regel andere Parameter verwenden.

Wegen der großen mathematischen Ähnlichkeiten verzichten wir an dieser Stelle auf eine detaillierte Analyse der Größe der ganzzahligen Programme und eine Beschreibung des lokalen Suchalgorithmus und verweisen auf Kapitel 4.

### 5.3 Ergebnisse

In diesem Abschnitt stellen wir die Ergebnisse vor, die wir bei der Optimierung von UMTS-Netze erzielt haben. Die hier vorgestellte UMTS-Analyse ist etwas komplexer als unsere Analyse von GSM-Netzen. Haben wir den Optimierungserfolg dort letzten Endes an einer Zahl festgemacht – nämlich dem aus der Frequenzzuweisung resultierenden Interferenzwert –, so ziehen wir bei der Analyse von UMTS-Netzen verschiedene Indikatoren zurate.

Die Analyse von UMTS-Netzen orientiert sich hauptsächlich am Lösungsvektor von System 5.9 aus Abschnitt 5.1, welcher mögliche Sendeleistungen für die Antennen repräsentiert. Wir untersuchen, wie viele Antennen an ihrer Leistungsgrenze sind, welche durchschnittlichen Sendeleistungen vorliegen und welcher Anteil der Sendeleistungen aufgewendet wird, um die Störungen durch andere Antennen zu neutralisieren.

Zunächst leiten wir eine Methode her, wie wir auch dann zu einer Netzbewertung gelangen, wenn die Norm der Lastkopplungsmatrix nicht kleiner als eins ist. Dazu führen wir Skalierungsvektoren für die Zeilen ein. Diese Skalierungsvektoren können wir ebenfalls zur Beurteilung der Netzqualität heranziehen. In Anschluss stellen wir unsere Ergebnisse vor. Wir sehen hier unsere Thesen aus Abschnitt 5.1 bestätigt, dass die Minimierung der Summe der Kopplungswerte zur Planung interferenzarmer UMTS-Netze geeignet ist.

Wir bezeichnen den Wert

$$\frac{\bar{p}_i^\downarrow}{\Pi_i^{\max\downarrow}}$$

als *Auslastung* einer Antenne. Wir möchten diesen Wert zunächst am *snapshot*-basierten Modell erläutern. An Gleichung 5.8 erkennt man, dass die Auslastung einer Antenne davon abhängt, wie viele *mobile stations* die Antenne versorgen muss, welche Last die einzelnen *mobile stations* erzeugen, und wie die Interferenzbedingungen in der Zelle sind. Es ist möglich, dass mehr Kunden mit einer Antenne Kontakt aufnehmen wollen, als diese gleichzeitig bedienen kann. In unserem Modell drückt sich das darin aus, dass die Auslastung einen Wert größer als eins annimmt. Wir sagen dann, die Antenne sei in *Überlast*, oder *überlastet*. Die nachteiligen Folgen der Überlastung sind, dass unter Umständen die möglichen Datenraten der Verbindungen reduziert werden müssen, Kunden abgewiesen werden beim Versuch, eine Verbindung herzustellen oder – im schlimmsten Fall – bestehende Verbindungen unterbrochen werden. Letzteres ist zum Beispiel möglich, wenn eine Kunde bei bestehender Verbindung in eine überlastete Zelle wechselt. Im flächenbasierten Modell deuten wir die Überlastung einer Antenne so, dass sie den Verkehr, der auf allen von ihr versorgten Pixel entsteht, nicht bedienen kann.

Ein Lösungsvektor  $\bar{p}^\downarrow = (E_n - C)^{-1} \cdot b$  unseres Gleichungssystems repräsentiert Sendeleistungen für Antennen, bei denen alle  $\mu$ -Zielwerte mit

Gleichheit erfüllt sind. Hat die Matrix  $C$  eine Norm, die größer ist als eins, so fällt es uns schwer, Aussagen über  $\bar{p}$  zu machen. Unter Umständen ist System 5.8 nicht eindeutig lösbar, eventuell enthält  $\bar{p}$  negative Werte. Dann ist es nicht sinnvoll, den Vektor  $\bar{p}$  als eine Konfiguration möglicher Sendeleistungen zu interpretieren. In [7] wurde eine Methode vorgestellt, wie auch in den oben beschriebenen Fällen sinnvolle Informationen mit Hilfe von System 3.3 erhalten werden können.

Bei der Analyse wird für das Netz  $N$  ein Vektor  $\lambda \in [0, 1]^{|N|}$  ermittelt, der Skalierungswerte enthält. Er wird einer Diagonalmatrix  $\Lambda$  einbeschrieben. Diese Matrix hat die Eigenschaft, dass der Vektor

$$\bar{p}^{\lambda \downarrow} = (I - \Lambda C)^{-1} \cdot b \quad (5.21)$$

Werte enthält, die im Leistungsspektrum der Antennen liegen. Das heißt es gilt

$$\hat{p}_i^{\downarrow} + \check{p}_i^{\downarrow} \leq \bar{p}^{\lambda \downarrow} \leq \Pi_i^{\max \downarrow} \quad (5.22)$$

für alle Antennen im Netz.

Der Vektor  $\lambda$  kann als Skalierung des Verkehrsaufkommens in den Zellen gedeutet werden. Gilt  $\lambda_i < 1$ , so kann Installation  $i$  den Verkehr in ihrer Zelle nicht vollständig versorgen. Die durch die Matrix  $\Lambda$  erfolgende Multiplikation der  $i$ -ten Zeile der Matrix  $C$  mit  $\lambda_i$  kann als Senkung der  $T_s(p)$ -Verkehrswerte in den Pixeln der Zelle  $i$  gedeutet werden. Man vergleiche hierzu die Definition der *Downlink*-Last in Gleichung 5.10.

Bei der Bewertung der Netzqualität kann die Zahl der überlasteten Zellen, also die Zahl Einträge ungleich eins im  $\lambda$ -Vektor, als Indikator verwendet werden. Allerdings muss dieser Indikator mit Vorsicht behandelt werden: In der Regel existieren natürlich viele  $\lambda$ -Vektoren, die Bedingungen 5.21 und 5.22 erfüllen. Möchte man zwei  $\lambda$ -Vektoren vergleichen, die für verschiedene Netze ermittelt wurden, so sollte zumindest das Verfahren ihrer Ermittlung gleich sein. Das in [7] vorgestellte Verfahren erfüllt diese Bedingung.

Mit Hilfe der ermittelten  $\lambda$ -Vektoren kann man verschiedene Indikatoren für die Netzqualität bestimmen:

- Anzahl überlasteter Antennen,
- Prozentualer Anteil des Planungsgebietes, in dem das vorliegende Pilotsignal keine ausreichende Qualität aufweist,
- Anteil der Antennenleistungen, der zur Überwindung der Interferenz verwendet wird.

In Tabelle 5.1 sind unsere Ergebnisse der UMTS-Planung für vier Instanzen dargestellt. Aus ihr kann man die ersten beiden Indikatoren der obigen Auflistung entnehmen. Jede Zeile repräsentiert ein Netz. Die Ausgangsnetze sind mit ihren Szenario-Name bezeichnet, die optimierten Netze einfach

mit „OPT“ und einer Ziffer. Unter jedem Ausgangsnetz stehen ein oder zwei optimierte Konfigurationen.

Die zweite Spalte gibt die Summe der Nebendiagonalelemente der Kopplungsmatrix wieder. In der dritten Spalte ist der prozentuale Anteil des Planungsgebiets festgehalten, wo Pilotsignale ausreichender Stärke, nicht jedoch ausreichender Qualität empfangen werden können. Die vierte Spalte führt die Anzahl der überlasteten Zellen im Netz auf. Es handelt sich hierbei um diejenigen Zellen, für die der Eintrag im oben beschriebenen  $\lambda$ -Vektor einen Wert von kleiner als eins hat. In den letzten beiden Spalten sind Leistungswerte notiert. Die vorletzte enthält die durchschnittliche Auslastung, die letzte die im Netz abgestrahlte Gesamtleistung (*total power*) in dBm.

Name	$\Sigma$	$E_c/I_0$ %	Überlastete Zellen	durchschn. Auslastung %	TP dBm
UBPe	36,09	10,28	16	26,44	1018,00
OPT1	25,67	6,16	13	23,57	907,62
OPT2	24,69	5,85	14	23,37	899,90
UPB	49,31	22,67	24	30,33	895,56
OPT1	28,69	18,24	21	27,62	815,59
UHG1	29,17	2,59	8	40,01	287,37
OPT1	10,46	1,56	7	37,33	268,14
UHG3	15,08	34,00	1	18,68	223,58
OPT1	8,68	0	1	19,00	227,45

Tabelle 5.1: Ergebnisse der UMTS-Planung.

Es ist zu erkennen, dass die Kopplungswerte in jedem der vier betrachteten Szenarios deutlich reduziert werden konnten. Die Reduktion liegt zwischen 32% bei UBPe und 64% bei UHG1. Die  $E_c/I_0$ -Spalte zeigt, dass sich durch die Optimierung deutliche Verbesserungen der Netzabdeckung erzielt wurden. Besonders hervorzuheben ist das Ergebnis im Szenario UHG3, wo zuvor 34 Prozent der Fläche nicht bedient werden konnten. Hier ist nun überall Empfang möglich, was unter Zulassung einer leichten Erhöhung der Sendeleistungen erreicht wurde. Die geringsten Fortschritte bei der  $E_c/I_0$ -Abdeckung wurden in UPB gemacht.

Die Zahl der Überlasteten Antennen konnte in allen Szenarien außer UHG3 leicht gesenkt werden. Wie oben beschrieben, ist dieser Indikator nur bedingt verwendbar. Es fällt aber auf, dass die Zahl überlasteter Antennen nicht im selben Maße sinkt wie die Kopplungswerte in der zweiten Spalte.

Die Netzlast ist nur in Verbindung mit der Anzahl skaliertes Zellen als Indikator verwendbar. Nur wenn beide Zahlen bei der Optimierung, können wir das optimierte Netz in dieser Hinsicht als besser einstufen. Ansonsten

sind die Einsparungen bei den Sendeleistungen unter Umständen nur erreicht worden, weil viele Zellen ihren Verkehr nicht bedienen. In UHG3 sehen wir, dass die abgestrahlte Leistung des optimierten Netzes größer ist als die des Ausgangsnetzes. Die Gewinne bei der  $E_c/I_0$ -Abdeckung gleichen diesen Nachteil nach unserer Auffassung allerdings aus. Die Senkungen der durchschnittlichen Antennenleistungen sind in allen Szenarien relativ gering. Die größten Einsparungen von knapp 12 Prozent erzielen wir im Szenario UBPe mit der Konfiguration OPT2. Auch hier könnte man ähnlich wie bei der Anzahl der überlasteten Zellen aufgrund der deutlicheren Reduzierungen der Kopplungswerte größere Einsparungen vermuten.

Kommen wir nun zur Analyse des letzten Indikators, des Anteils der Antennenleistung, der aufgewendet werden muss, um die Interferenz durch andere Antennen zu überwinden. Wir nennen diesen Wert  $p_i^*$  und möchten ihn nun genau definieren: Aus Gleichung 5.8 kann man ablesen, wieviel Leistung Antenne  $i$  zur Versorgung des eigenen Verkehrs aufwenden müsste, wenn keine Interferenz durch andere Antennen vorläge. Dieser Wert heiße  $p_i^\downarrow$ . Aus

$$p_i^\downarrow = C_i^\downarrow p_i^\downarrow + p_j^{\downarrow\eta} + \hat{p}_j^\downarrow + \check{p}_j^\downarrow \quad (5.23)$$

folgt

$$p_i^\downarrow = \frac{p_j^{\downarrow\eta} + \hat{p}_j^\downarrow + \check{p}_j^\downarrow}{1 - C_i^\downarrow}. \quad (5.24)$$

Die tatsächliche Antennenleistung unter Verwendung des  $\lambda$ -Vektors ist  $\bar{p}_i^\downarrow$ . Es sei

$$p_i^* := \frac{\bar{p}_i^\downarrow - p_i^\downarrow}{\Pi_i^{\max\downarrow}}. \quad (5.25)$$

In Tabelle 5.2 sind die Ergebnisse der Optimierung dargestellt. Zu jedem Szenario gehören zwei Spalten, eine für das Ausgangsnetz (A) und eine für das optimierte Netz (O). In den Zeilen ist die jeweilige Anzahl der Zellen aufgeführt, wo der  $p^*$ -Wert einen gegebenen Schwellwert überschreitet. Die erste Zeile führt dabei die Zahl der Antennen auf, die mehr als zehn Prozent ihrer Gesamtleistung zur Neutralisierung der Interferenz aufwenden. In der zweiten Zeile sind es zwanzig Prozent usw. In der letzten Zeile ist noch einmal die jeweilige Zellenanzahl der Netze aufgeführt.

In allen Szenarien außer in UPB sind die Zahlen der optimierten Netze mit Ausnahme von  $p^* = 20\%$  für UHG1 kleiner als die Zahlen der Ausgangsnetze. Ihre Antennen verwenden also weniger Energie darauf, die Interferenz durch die anderen Zellen zu neutralisieren.

## 5.4 Zusammenfassung

In diesem Kapitel stellten wir ein Verfahren zur Planung interferenzarmer UMTS-Netze vor. Dieses Verfahren stellt eine Realisierung unseres allgemeinen Ansatzes der Funknetzplanung dar. Dieser Ansatz besteht darin,

$p^*$ - Schwell- wert %	UPBe		UPB		UHG1		UHG3	
	A	O	A	O	A	O	A	O
10	71	46	34	27	28	24	19	11
20	41	23	22	17	17	18	8	7
30	24	18	14	14	14	12	3	2
40	18	14	10	12	12	8	1	1
50	6	2	4	9	6	5	1	0
	193		148		36		60	

Tabelle 5.2: Vergleich der  $p^*$ -Werte von Ausgangs- und optimierten Netzen.

unter gegebenen Abdeckungsbedingungen das Interferenzpotenzial zu minimieren. Dieses Interferenzpotenzial drückten wir in unserem UMTS-Ansatz durch die Lastkopplungsfaktoren

$$C_i^{\downarrow j} := \sum_{q \in P_i} l_q^{\downarrow} \frac{\gamma_{jq}^{\downarrow}}{\gamma_{iq}^{\downarrow}}$$

aus. Die Kopplung zweier Zellen ist umso größer, je näher ihre Pfadverlustwerte auf den Pixeln beieinanderliegen. Sie ist allerdings auch abhängig von dem prognostizierten Verkehr in den *best server*-Bereichen. Je größer der Verkehr in einem solchen Bereich ist, desto größer ist die Kopplung der betroffenen Zelle mit anderen Zellen.

Die Lastkopplungsfaktoren sind Elemente eines linearen Gleichungssystems. Die Lösung dieses Systems können wir bei Erfüllung bestimmter Lösbarkeitsbedingungen als Werte für die Sendeleistungen der Antennen interpretieren. Wir stellten drei Thesen auf, wie die Netzkonfiguration im Hinblick auf die Kopplungsmatrix zu verändern ist, damit das lineare Gleichungssystem für uns wertvolle Lösungen hat.

Sind die Lösbarkeitsbedingungen nicht erfüllt, so können wir das System mit den eingeführten  $\lambda$ -Vektoren so manipulieren, dass es lösbar ist, und dass die Lösungswerte innerhalb der möglichen Antennenleistungen liegen. Zur Netzbewertung zogen wir sowohl die resultierenden Leistungswerte als auch die  $\lambda$ -Vektoren heran.

Unsere Analysen belegen, dass die optimierten Netze nach unseren Kriterien in der Regel bessere Eigenschaften haben als die Ausgangsnetze. Die Verbesserungen lassen sich allerdings nicht in befriedigender Weise quantifizieren, weil die von uns verwendeten Indikatoren nur unter Zuhilfenahme der anderen Indikatoren und bei Betrachtung der Netzstruktur ausgewertet werden können. So zeigt sich bei UHG3 eine deutliche Verbesserung der Abdeckung bei leichter Verschlechterung der Sendeleistungen. Bei UHG1 zeigt sich nur eine Verringerung der Zahl überlasteter Antennen von 8 auf

7, obwohl die Werte der Kopplungsmatrix deutlich geringer sind. Bei der Konfiguration OPT1 für UBPe ist die Zahl der überlasteten Antennen geringer als bei OPT2, obwohl alle anderen Indikatoren anzeigen, dass OPT2 die bessere Konfiguration ist. Die Unterschiede sind allerdings sehr gering.

Wir konnten unsere Thesen, auf denen der Optimierungsansatz basiert, nicht bestätigen, weil die Matrizen in sämtlichen von uns untersuchten Szenarien die Lösbarkeitsbedingungen nicht erfüllten. Die maximalen Zeilensummen waren immer größer als eins, was zur Folge hatte, dass immer überlastete Zellen existierten.

Wir ziehen aus unseren Ergebnissen zwei Schlüsse: Wir glauben, dass der Ansatz, die Summe der Kopplungsfaktoren zu minimieren, sinnvoll ist. Dieser Sinn erschließt sich schon, wenn man sich die Bedeutung dieser Faktoren anhand der Definition klarmacht. Wir glauben aber auch, dass die Methode der Zeilenskalierung mit den  $\lambda$ -Faktoren auf die Analyse einen zu großen Einfluss hat. Wir vermuten, dass ein schlechtes Netz durch Wahl eines günstigen  $\lambda$ -Vektors „verbessert“ werden kann, und ein gutes Netz durch Wahl eines ungünstigen  $\lambda$ -Vektors „verschlechtert“ werden kann. Aus unseren Ergebnissen geht nicht hervor, ob solche Effekte auf unseren Instanzen auftreten. Unsere Optimierungsthesen lassen sich nur anhand von Netzen belegen, die sowieso schon gut konfiguriert sind, also so kleine Kopplungswerte aufweisen, dass die Zeilensummen kleiner als eins sind.

Es gibt allerdings auch eindeutige Faktoren, die für die Relevanz unserer Resultate sprechen: Die Skalierungsvektoren wurden hier für alle Instanzen mit demselben Verfahren ermittelt. Es ist also wenig plausibel, dass es auf den optimierten Netzen gute und auf den Ausgangsnetzen schlechte Resultate produziert. Außerdem konnten wir deutliche Fortschritte bei der  $E_c/I_0$ -Abdeckung erreichen, die keine Verbindung mit den Skalierungsvektoren aufweist. Außerdem möchten wir die Summe der Lastkopplungsfaktoren an sich schon als Qualitätsmaß heranziehen. Kleinere Kopplungswerte besagen eindeutig, dass die betreffenden Antennen auf den relevanten Pixeln große Pfadverlustdifferenzen aufweisen und sich gegenseitig weniger stören. Dieses Ergebnis ist unabhängig von der Betrachtung von Sendeleistungen ein Erfolg. Nichtsdestotrotz wäre es wünschenswert, ein Analyseverfahren zu entwickeln, das die Tauglichkeit unseres Planungsansatzes auch für schlechte Ausgangskonfigurationen nachweist.



## Kapitel 6

# Zusammenfassung

In dieser Arbeit stellten wir einen technologieunabhängigen Ansatz zur Planung der Funkschnittstelle digitaler Funknetzen vor. Dabei wurden Abdeckung, Kapazität und Interferenz als Hauptaspekte der Planung betrachtet. Da die letzten Endes auftretende Interferenz oft schwer zu ermitteln ist – in GSM ist hier beispielsweise eine komplexe Frequenzplanung vonnöten, in UMTS eine mitunter schwierige Abschätzung der Sendeleistungen – entwickelten wir Methoden, das Interferenzpotenzial abzuschätzen. Wir belegten am Beispiel der GSM- und der UMTS-Technologie, dass die Minimierung des Interferenzpotenzials zu deutlichen Verbesserungen beim Interferenzverhalten führt.

Um die Netzabdeckung zu gewährleisten, wählten wir einen verallgemeinerten *set covering*-Ansatz. Die hier enthaltenen Nebenbedingungen besagen, dass der Pfadverlust des *best servers* im Planungsgebiet nirgendwo einen bestimmten Wert überschreiten darf. Als Zielfunktion wählten wir die Minimierung des angesprochenen Interferenzpotenzials. Die Definition dieses Potenzials wird den Gegebenheiten jeder Technologie angepasst. Durch die resultierende Reduktion der Interferenz erreichen wir eine Steigerung der Kapazität des Netzes. Sie wird in jeder Technologie anders errechnet, doch ist das zugrundeliegende Prinzip dasselbe: Die Ausgangskonfiguration weist zuviel Interferenz auf, was bedeutet, dass sie mit dem gegebenen Verkehr nicht fertig wird. Durch die Optimierung erhalten wir eine Konfiguration, die weniger Interferenz aufweist. Sind die Interferenzwerte gering genug, so legen wir das so aus, dass das Netz für den gegebenen Verkehr ausreichend Kapazität aufweist. Es wurde also eine Kapazitätssteigerung erreicht.

In der GSM-Technologie betrachteten wir die Überlappungen der Zellen. Je größer sie waren, desto größer war das Interferenzpotenzial. Allerdings spielten die Überlappungen nur in Bereichen eine Rolle, die in vielen Netzkonfigurationen im *best server*-Bereich einer der beiden Zellen lagen. Die letztendliche Beurteilung einer Netzkonfiguration konnte erst nach vollzogener Frequenzzuweisung erfolgen. Unsere GSM-Ergebnisse sind sehr befriedigend, denn die Interferenzwerte konnten in den betrachteten Szenarien

deutlich gesenkt werden.

In der UMTS-Planung spielten die Überlappungen der Zellen ebenfalls eine Rolle. Allerdings war es von entscheidender Bedeutung, wie groß die Differenz der Pfadverluste in den Überlappungsbereichen waren. Wurde in GSM nur ein qualitatives Urteil gestört/nicht gestört gefällt, so lag hier grundsätzlich eine Störung vor, die wir nur vernachlässigten, wenn sie sehr klein war. Die UMTS-Analyse erfolgte auf Basis einer möglichen Konfiguration der Antennenleistungen, bei der ein Großteil des Verkehrs bedient werden konnte. Als problematisch erwies sich, dass die Verwendung von Skalierungsvektoren einen für uns schwer einschätzbaren Einfluss auf das Resultat hatte. Die Ergebnisse zeigten in drei von vier Fällen Verbesserungen bezüglich der Auslastung. Es wäre jedoch wünschenswert, ein Qualitätsmaß zu entwickeln, an dem man wie bei GSM die Netzqualität eindeutig beurteilen kann.

Nach unserer Einschätzung ist es lohnenswert, den Lösungsalgorithmus so zu verändern, dass untere Schranken für die Summe der Kopplungswerte angegeben werden können. Das IP-Modell ist nach unserem Eindruck nicht dafür geeignet. Wie gesehen führten unsere Schnittungleichungen nur zu unwesentlichen Verbesserungen, und wir halten es für fraglich, ob durch das Finden weiterer Klassen deutliche Verbesserungen erzielt werden können.

Wir halten unseren Planungsansatz für einen Erfolg. Sowohl in GSM als auch in UMTS konnten wir deutliche Erfolge bei der Entkopplung der Zellen aufweisen, was wir an sich schon als Qualitätskriterium betrachten. In GSM konnten wir das mit den Ergebnissen der Frequenzzuweisung belegen. In UMTS fehlt uns noch ein eindeutiges Qualitätsmaß, wenn man die deutlichen Verbesserungen der  $E_c/I_0$ -Abdeckung und die Reduktion der für die Überwindung von Interferenz benötigten Sendeleistungen nicht schon als Erfolg wertet.

# Literaturverzeichnis

- [1] Mathematical methods for automatic optimization of umts radio networks. Technical report, Zuse Institute Berlin, 2003.
- [2] E. Balas and J.B. Mazzola. Nonlinear 0-1 programming; I. linearization techniques. *Mathematical Programming*, 30(1):1–30, 1984.
- [3] R. E. Bixby. Implementing the simplex method: The initial basis. Technical report, Rice University, 1990.
- [4] R. Borndörfer. *Aspect of Set Packing, Partitioning, and Covering*. PhD thesis, Technische Universität Berlin, 1998.
- [5] T.H. Cormen, C.E. Leiserson, and R.L. Rivest. *Introduction to Algorithms*. MIT Press, 1990.
- [6] A. Eisenblätter. *Frequency Assignment in GSM Networks: Models, Heuristics and Lower Bounds*. Cuvillier Verlag Göttingen, 2001.
- [7] A. Eisenblätter, H.-F. Geerdes, T. Koch, A. Martin, and R. Wessäly. UMTS Radio Network Evaluation and Optimization beyond Snapshots, 2004.
- [8] ETSI. *GSM Digital cellular telecommunications system (Phase2+); Radio transmission and reception (GSM 05.05)*, July 1996.
- [9] R. Fortet. L’algèbre de boole et ses applications en recherche opérationnelle. *Cahiers du Centre d’Études de Recherche Opérationnelle*, 1959.
- [10] F. Glover and E. Woolsey. Converting the 0-1 polynomial programming problem to a 0-1 linear program. *Operations Research*, 22:180–182, 1974.
- [11] F. Granot and P. L. Hammer. On the role of generalized covering problems. *Cahiers du Centre d’Études de Recherche Opérationnelle*, 1975.
- [12] P. Hansen, B. Jaumard, and V. Mathon. Constrained nonlinear 0-1 programming. *ORSA Journal on Computing*, 5(2):97–119, 1993.

- [13] S. Hurley. Planning effective cellular mobile radio networks. *IEEE Transactions on Vehicular Technology*, 51(2):243–253, 2002.
- [14] A. Jedidi, A. Caminada, G. Finke, and T. Dony. Geometrical constraint handling and evolutionary algorithm for cellular network design. In *INOC 2003*.
- [15] J. Laiho, A. Wacker, and T. Novosad, editors. *Radio Network Planning and Optimization for UMTS*. John Wiley & Sons, Ltd, 2002.
- [16] V. Mehrmann. Vorlesungsskript Praktische Mathematik I, 2001.
- [17] MOMENTUM. Ist-2000-28088 momentum, models and simulation for networkplanning and control of umts. <http://momentum.zib.de>.
- [18] R. Plato. *Numerische Mathematik kompakt*. Vieweg, 2000.
- [19] B. Walke, M.P. Althoff, and P. Seidenberg. *UMTS - Ein Kurs, Zweite Auflage*. J. Schlembach Fachverlag, Weil der Stadt, 2002.
- [20] M.D. Yacoub. *Foundations of Mobile Radio Engineering*. CRC Press, Inc., 1993.