Kosten- und Qualitätsoptimierung in Kommunikationsnetzen

Andreas Bley Arie M.C.A. Koster Alexander Kröller Roland Wessäly
Adrian Zymolka

2003

Zusammenfassung

Der scharfe Wettbewerb innerhalb der Telekommunikationsbranche zwingt die Netzbetreiber dazu, ihre Investitionen genau zu planen und immer wieder Einsparungsmaßnahmen durchzuführen. Gleichzeitig ist es jedoch wichtig, die Qualität der angebotenen Dienste zu verbessern, um neue Kunden zu gewinnen und langfristig an sich zu binden.

Die mathematische Optimierung bietet sich für viele solcher Aufgabenstellungen als hervorragend geeignetes Planungswerkzeug an. Ziel dieses Artikels ist es, ihre Methodik und ihre Anwendung speziell zur Kosten- und Qualitätsoptimierung in Kommunikationsnetzen vorzustellen. Anhand von vier konkreten Planungsaufgaben aus dem Bereich der Festnetzplanung wird aufgezeigt, wie sich komplexe Zusammenhänge in flexiblen mathematischen Modellen abbilden lassen und welche Verfahren zur automatisierten Bearbeitung der Probleme eingesetzt werden können. Die hier vorgestellten Methoden zeichnen sich insbesondere dadurch aus, dass sie neben hochwertigen Lösungen auch eine Qualitätsgarantie liefern, mit der sich die Lösungen fundiert bewerten lassen. Die dokumentierten Ergebnisse aus verschiedenen Industrieprojekten belegen die Eignung und Güte der mathematischen Optimierung für die Praxis.

Zuse-Institut Berlin, Takustr. 7, 14195 Berlin, {bley,koster,kroeller,wessaely,zymolka}@zib.de

Inhaltsverzeichnis

1 Einleitung: Effiziente Nutzung von Ressourcen **3**

2 Planungsaufgaben: Was soll optimiert werden? **4**

2.1 Netzstrukturplanung 6

2.2 Kostenoptimierung in SDH-Netzen 8

2.3 Design optischer Netze 10

2.4 Traffic Engineering in IP-Netzen 13

3 Modelle: Wie sieht's mathematisch aus? **15**

3.1 Modellierung mathematischer Optimierungsprobleme 15

3.2 Hardware und Netzkonfigurationen 16

3.3 Routing 21

3.4 Integrierte Optimierung 24

4 Algorithmen: Welche Methoden nutzen wir? **26**

4.1 Vollständige Enumeration 27

4.2 Branch&Bound 29

4.3 Branch&Bound für ganzzahlige lineare Programme 30

4.4 Branch&Cut 31

4.5 Branch&Cut&Price 33

5 Optimierung in Aktion: Was bringt's? **34**

5.1 Netzstrukturplanung 35

5.2 Kostenoptimierter Ausbau von SDH-Netzen 38

5.3 Ressourceneffiziente Ausfallsicherheit in optischen Netzen 41

5.4 Optimierung der Routinggewichte in IP-Netzen 45

6 Zusammenfassung, Summary, Résumé **47**

A Verzeichnis der Abkürzungen und Symbole **53**

1 Einleitung: Effiziente Nutzung von Ressourcen

Unternehmen, die langfristig erfolgreich am Markt agieren, zeichnen sich durch ihre Innovationskraft sowie durch eine effiziente Nutzung von Ressourcen aus. Fluggesellschaften optimieren ihre Flugpläne, Telekommunikationsdienstleister konkurrieren auf deregulierten Märkten, Hersteller müssen ihre Produktionsprozesse effizient gestalten und dabei Zulieferer und Abnehmer einbinden: Jedes Unternehmen muss die vorhandenen Optimierungsmöglichkeiten bestmöglich nutzen, um konkurrenzfähig zu sein.

Die technischen und organisatorischen Zusammenhänge bei der Optimierung der Ressourcennutzung sind normalerweise so komplex und umfassend, dass eine Planung von Hand entweder unmöglich oder nur mit unakzeptablem Zeitaufwand realisierbar ist. Mittlerweile kann durch den Einsatz automatisierter Verfahren oft eine hohe Planungsqualität und -sicherheit erreicht werden: Effiziente und intelligente Berechnungsverfahren können der Motor der Kosten- und Qualitätsoptimierung sein. Die Mathematik stellt als quantitative Wissenschaft ein geeignetes Hilfsmittel zur Lösung komplexer Aufgaben dar. Wesentlichen Anteil daran trägt der rasante Fortschritt der mathematischen Optimierung, die sich als hervorragend geeignetes Werkzeug zur realitätsgerechten Erfassung und erfolgreichen Bearbeitung schwieriger Planungsaufgaben etabliert hat. Motiviert durch anwendungsorientierte Fragestellungen hat sie ein umfangreiches Sortiment geeigneter Lösungsansätze und -methoden anzubieten, mit denen bereits heute viele aufwändige Aufgaben routinemäßig mit höchsten Qualitätsansprüchen gelöst werden.

Am Zuse-Institut Berlin (ZIB) werden seit vielen Jahren in Kooperationen mit der Industrie neuartige Optimierungsverfahren entwickelt und als Software für den täglichen Einsatz bei Projektpartnern implementiert, siehe [15] und vorangegangene Jahresberichte. Die praktischen Anwendungen kommen aus vielen verschiedenen Branchen: öffentlicher Nahverkehr, Logistik, Energieversorgung und nicht zuletzt Telekommunikation. Die Qualität und die praktische Relevanz dieser Entwicklungen spiegelt sich in praxisorientierten sowie wissenschaftlichen Auszeichnungen wider. Speziell aus dem Bereich der Telekommunikation sind jüngst der Vodafone Innovationspreis [25] und die Dissertationspreise der INFORMS Telecommunication Section [14] sowie der Gesellschaft für Operations Research [12] hervorzuheben.

Ziel dieses Artikels ist es, die mathematische Optimierung, ihre Methodik sowie ihre Anwendung bei der Kosten- und Qualitätsoptimierung in Telekommunikationsnetzen zu präsentieren. Dies geschieht anhand von vier konkreten Planungsproblemen aus der Festnetzplanung, für die wir in den vergangenen Jahren zusammen mit den jeweiligen Netzbetreibern Lösungsverfahren entwickelt haben. Wir orientieren uns dabei an einer sowohl pragmatischen als auch erfolgreichen Vorgehensweise bei der Lösung praktischer Planungsaufgaben: dem in Abbildung 1 illustrierten Problemlösungszyklus.

Die mathematische Optimierung beginnt mit einer möglichst genauen Beschreibung der betrachteten Aufgabenstellungen. Hierbei ist insbesondere auf die Spezifikation aller relevanten Rahmenbedingungen zu achten. Dazu beschreiben wir in Kapitel 2 zunächst eine Auswahl von Planungs- und Konfigurationsaufgaben aus der Festnetzplanung, mit denen wir uns im Rahmen unterschiedlicher Projekte beschäftigt haben. Darauf setzt der zentrale Schritt einer mathematisch geprägten Vorgehensweise auf: die Modellbildung, die wir in Kapitel 3 erläutern. Dabei wird die spezifizierte Aufgabe mit Hilfe von Variablen und Nebenbedingungen formal und abstrakt beschrieben und so einer algorithmischen Bearbeitung zugänglich gemacht. Das resultierende Modell ist die Schnittstelle zwischen der praktischen Aufgabe und ihrer methodischen Lösung.

Wie geeignete Lösungsverfahren entwickelt werden und worauf diese beruhen, stellen wir in Kapitel 4 vor. Dazu erläutern wir einige Grundlagen der Theorie moderner Optimierungsverfahren. Die beschrie-

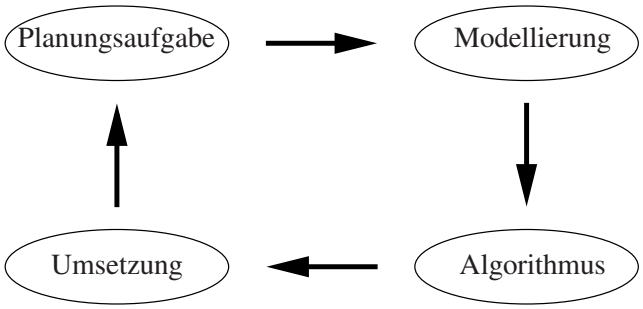


Abbildung 1: Problemlösungszyklus.

benen Techniken und (grundlegenden) Verfahren stellen natürlich nur einen kleinen Ausschnitt aus dem umfangreichen Instrumentarium der mathematischen Optimierung dar – eine vollständige Zusammenstellung kann und soll nicht Inhalt dieser Übersicht sein. Als Ausgangspunkt für interessierte Leser stellen wir am Ende von Kapitel 4 daher eine Liste einführender und vertiefender Literatur vor.

Um das Potential dieser Herangehensweise und der mathematischen Optimierung bei der Umsetzung in der Praxis zu demonstrieren, greifen wir abschließend in Kapitel 5 die eingangs vorgestellten Planungsaufgaben nochmals auf und präsentieren für realistische Szenarien die erzielten Ergebnisse.

Es ist angebracht, bereits an dieser Stelle auf einen wichtigen Aspekt aufmerksam zu machen: die konkrete Auffassung von Optimierung. In der umgangssprachlich verwendeten Form beschränkt sich dieser Begriff meist auf eine Verbesserung eines gegenwärtigen Zustandes. Ob es allerdings darüber hinaus noch weitere Verbesserungen gibt, bleibt oftmals unklar. Der mathematischen Optimierung liegt eine strengere Auffassung zugrunde: Es geht nicht nur um irgendeine Verbesserung, sondern um das Erreichen des tatsächlich bestmöglichen – optimalen – Zustandes unter allen für die gestellte Aufgabe möglichen Lösungen. Ein solches Optimum ist nicht weiter verbesserbar und hängt insbesondere nicht vom eingesetzten Verfahren ab. Die von uns beschriebenen Verfahren grenzen sich bezüglich dieser strengeren Auffassung in zweierlei Hinsicht ab. Erstens zielen sie stets auf die Ermittlung des Optimums ab. Da dies in vernünftiger Zeit allerdings nicht immer garantiert werden kann, wird zweitens stets eine Gütegarantie für die beste gefundene Lösung mitgeliefert, auf deren Basis eine fundierte Bewertung der Lösung vorgenommen werden kann. Das hebt die vorgestellten Methoden von Verfahren ab, die „nur“ eine Lösung erzeugen, aber keine Aussage darüber machen können, ob es noch bessere Lösungen geben kann und um wieviel diese höchstens besser sein könnten. In Kapitel 4 erläutern wir diese beiden Aspekte, die das charakteristische Merkmal der strengeren mathematischen Auffassung von Optimierung darstellen.

2 Planungsaufgaben: Was soll optimiert werden?

Die Struktur von Kommunikationsnetzen ergibt sich aus einer Vielzahl technischer und organisatorischer Überlegungen. Dabei finden nicht nur Kosten, sondern auch geografische Gegebenheiten, technische und funktionale Aspekte sowie die praktische Nutzbarkeit der Netze Berücksichtigung. Größere Netze werden üblicherweise partitioniert und in mehrere Hierarchieebenen gegliedert. Dadurch entstehen für die Planung und den Betrieb leichter handhabbare Teilnetze, beispielsweise Zugangs- und Kernnetz, wie in

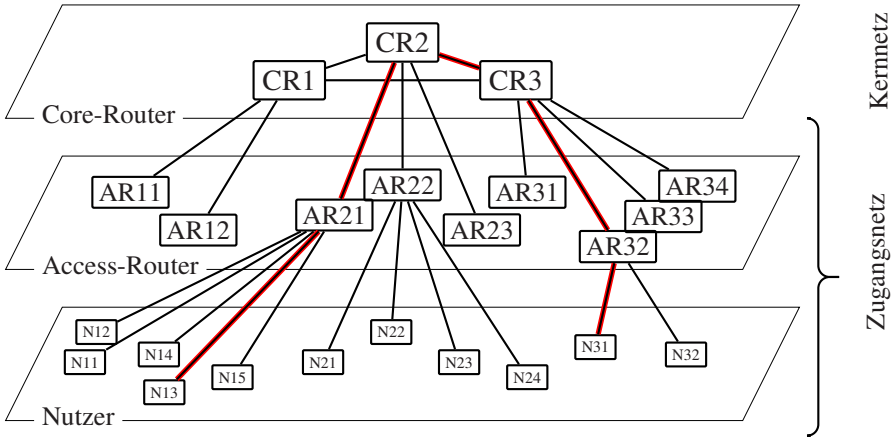


Abbildung 2: (IP-)Netz mit drei Hierarchieebenen: Netzstruktur und Routing.

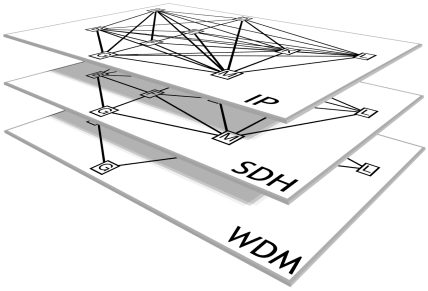


Abbildung 3: Schichtenmodell für IP über SDH über WDM.

Abbildung 2 dargestellt. Des weiteren sind Kommunikationsnetze auch technologisch in mehrere Netzschichten untergliedert. Die einzelnen Schichten bieten unterschiedliche Funktionalitäten und ermöglichen so eine bequeme Realisierung verschiedener Dienste in einem gemeinsamen physikalischen Netz. In Abbildung 3 ist ein mögliches Schichtenmodell dargestellt, bei dem ein IP- über einem SDH-Netz realisiert wird, das wiederum ein rein optisches WDM-Netz als Transportschicht für die SDH-Signale verwendet.

Die skizzierte Gliederung in verschiedene Teilnetze und Netzschichten überträgt sich auf den gesamten Planungsprozess für Kommunikationsnetze. Viele der praktischen Aufgaben aus dem Bereich der Festnetzplanung stehen in engem Zusammenhang mit dem technischen Aufbau und der organisatorischen Struktur der Netze. In diesem Kapitel stellen wir vier konkrete Planungsaufgaben vor, die sich mit charakteristischen Fragestellungen bei Planung, Aufbau und Betrieb von Kommunikationsnetzen beschäftigen, und die wir in den letzten Jahren in Kooperationen mit verschiedenen Netzbetreibern erfolgreich bearbeitet haben.

Bei der Netzstrukturplanung, die wir in Abschnitt 2.1 betrachten, ist eine geeignete Unterteilung in Kern- und Zugangsnetze festzulegen und jedes dieser Netze zu dimensionieren. Wegen des langen Planungshorizontes ist dies eine strategische Planungsaufgabe. Im Vordergrund steht dabei die Minimierung der

Gesamtnetzkosten, aber auch die Robustheit des Netzes gegenüber unsicheren Planungsdaten ist von Bedeutung. Ferner gehören Technologieentscheidungen und die Herstellerauswahl zu dieser Aufgabe.

Der kostenoptimale Auf- und Ausbau einzelner Netzschichten ist eine wichtige Konfigurationsaufgabe, die wir speziell für SDH-Netze in Abschnitt 2.2 sowie für optische Netze in Abschnitt 2.3 beschreiben. Diese Planungsaufgaben umfassen sowohl strategische Entscheidungen, wie die Hersteller- und Hardwareauswahl, als auch operative Entscheidungen, insbesondere die Konfiguration der Hardware und die Routenplanung. Primäres Ziel bei diesen Planungsaufgaben ist die Minimierung von Investitions- und Betriebskosten der jeweiligen Netzschicht.

Das Traffic Engineering für IP-Netze, das wir in Abschnitt 2.4 vorstellen, ist eine für die operative Netzplanung charakteristische Aufgabe. Für einen sehr kurzfristigen Planungshorizont sollen dabei die vorhandenen Ressourcen durch geeignetes Umrouten der Verkehrsströme besser genutzt werden. Hierdurch wird eine Verbesserung der Kommunikationsqualität erreicht.

Solche Planungsaufgaben lassen sich mit den in Kapitel 3 vorgestellten Methoden nicht nur ganzheitlich in jeweils einem integrierten Modell abbilden, sondern mit den in Kapitel 4 beschriebenen Techniken auch adäquat lösen.

2.1 Netzstrukturplanung

Der Auf- und Ausbau von Kommunikationsnetzen beinhaltet eine Vielzahl langfristiger und kostspieliger Entscheidungen: An welchen geografischen Standorten werden Netzknoten welcher Hierarchieebene eingerichtet, wie sind diese untereinander zu verbinden und wie sind die Verbindungen und die Knoten zu dimensionieren? Das Ziel der strategischen Netzstrukturplanung ist es, die Gesamtnetzkosten zu minimieren. Wir haben in verschiedenen Projekten die Entwurfs- und Ausbauplanung für das Gigabit-Wissenschaftsnetz G-WiN des DFN-Verein e. V., die Reduktion der benötigten Vollvermittlungsstellen im Sprachnetz der Telekom Austria sowie die MSC-Ausbauplanung für E-Plus untersucht [5, 6].

2.1.1 Standorte, Hardware und Kapazitäten

Ausgangspunkt der Netzstrukturplanung ist die Menge aller Standorte. Dazu zählen alle Nutzerstandorte, zwischen denen Verkehrsströme fließen sollen, sowie die Standorte aller existierenden oder möglicherweise aufzubauenden Knoten. Standortabhängig können Knoten bestimmter Hierarchieebenen eingerichtet werden. Kernnetzknoten dürfen zum Beispiel nur an Standorten aufgebaut werden, die über die nötige Infrastruktur sowie Wartungs- und Servicepersonal verfügen. Andererseits gibt es auch Standorte, an denen Knoten einer bestimmten Ebene eingerichtet werden müssen. Beispielsweise kann an Übergängen in andere Netze die Einrichtung eines Kernnetzknotens zwingend erforderlich sein. Schließlich kann die Anzahl der Knoten in den verschiedenen Teilnetzen auch beschränkt sein.

Jeder eingerichtete Knoten muss entsprechend den Anforderungen seiner Netzebene ausgebaut werden. Dazu gehören Infrastruktur (Gebäude, Stellflächen, Klima- und Stromversorgung usw.), Gerätetechnik (Switches, Router, Schnittstellenkarten, Geräte für Backup, Accounting oder Gebührenerfassung usw.) sowie Betriebs- und Administrationspersonal. Es ist zwar möglich, all diese Anforderungen im Detail zu berücksichtigen, in Anbetracht der häufig sehr ungenauen und unsicheren Datenlage ist dies jedoch in der Regel nicht sinnvoll. Der Planungshorizont ist oft zu lang für eine hinreichend präzise Prognose detaillierter Kostendaten. Bei der Planung für zukünftige Technologien lassen sich unter Umständen nicht

einmal die genauen technischen Eigenschaften der Geräte verlässlich vorhersagen. Meist beschränkt man sich daher bei der strategischen Netzstrukturplanung darauf, viele Details in aggregierter Form zu betrachten: Je Knoten stehen einige wenige Basiskonfigurationen und einige mögliche Erweiterungen zur Wahl, die den erwarteten Ausbaustufen des Knotens entsprechen.

Zwischen den Netzknoten können verschiedene Verbindungen eingerichtet werden, deren Art von den geografischen Standorten und den gewählten Netzebenen und Konfigurationen der Knoten abhängen kann. Zum Beispiel sind bestimmte Technologien und Kapazitätsstufen nur im Kern- oder nur im Zugangsnetz verwendbar. Ebenso können spezielle physikalisch getrennte Doppelverbindungen aus Sicherheitsgründen für die Verbindungen in der obersten Zugangsnetzebene zwingend gefordert sein. Zudem sind geografische Beschränkungen, wie maximale Entfernungen für Verbindungen, zu berücksichtigen. Ähnlich wie bei der Konfiguration der Knoten ist es auch hier sinnvoll, sich bei der Planung auf wenige Verbindungstypen zu beschränken, wenn technische Details nicht hinreichend genau prognostiziert werden können.

Das Zugangsnetz soll streng hierarchisch aufgebaut werden: Die Knoten einer Netzebene sind stern- oder baumförmig mit den Knoten der nächsthöheren Ebene zu verbinden. Als zentrales Element muss das Kernnetz allerdings erhöhten Sicherheitsanforderungen genügen. Es sollte daher mindestens so stark vermascht und so dimensioniert sein, dass es gegen Ausfälle einzelner Knoten oder Links abgesichert ist.

2.1.2 Routing und Ausfallsicherheit

Verkehrsströme zwischen den Knoten des Netzes werden von ihrer Quelle zunächst durch das Zugangsnetz in der Hierarchie nach oben, dann möglicherweise einige Schritte durch das Kernnetz und schließlich wieder durch das Zugangsnetz in der Hierarchie nach unten zur Senke geroutet. Ein solcher Weg ist in Abbildung 2 durch stärkere Linien angedeutet. Mit der Entscheidung der Netzstruktur sind die Routingwege im Zugangsnetz aufgrund der Stern- oder Baumform festgelegt. Welche Wege im Kernnetz gewählt werden können und wie Ausfallsicherheit dort im Detail realisiert wird, hängt vom verwendeten Routingprotokoll ab. Die durch die Konfigurationen der Knoten und Verbindungen zur Verfügung gestellten Kapazitäten müssen jedoch ausreichen, um sowohl im Regelbetrieb als auch in Ausfallsituationen den prognostizierten Verkehr gemäß Routingprotokoll realisieren zu können.

2.1.3 Planungsaufgabe

Das Ziel der Netzstrukturplanung ist die Minimierung der Gesamtnetzkosten. Dabei sind die Einrichtung von Knoten und Verbindungen sowie einzelne dort installierbare Konfigurationen, zwischen denen auszuwählen ist, mit jeweils einem Kostenwert belegt, der bereits alle zugehörigen Kosten in aggregierter Form enthält. Die Aufgabe der Netzstrukturplanung lässt sich somit wie folgt kurz zusammenfassen:

PLANUNGSAUFGABE
Netzstrukturplanung

Planungsziel: Minimierung der Kosten

Entscheidungen:

- Auswahl der Netzknoten und Zuordnung zu Netzebenen
- Topologien des Kern- und des Zugangsnetzes
- Konfiguration der Netzknoten und Links
- Routing der Verkehrsbedarfe

Randbedingungen:

- Ausfallsicherheit im Kernnetz
- Stern- oder Baumstrukturen im Zugangsnetz
- Ausreichende Dimensionierung für Routing der Bedarfe

2.2 Kostenoptimierung in SDH-Netzen

Das SDH-Netz dient als gemeinsames Übertragungsmedium für andere Netzschichten, wie zum Beispiel für ATM, IP, SS7 oder VPNs. Bei seiner Planung sind zahlreiche, sich gegenseitig beeinflussende Entscheidungen zu treffen. Dies umfasst sowohl strategische Entscheidungen, wie die Auswahl von Standorten, die Topologie des Netzes oder die Gerätehersteller, als auch operative Entscheidungen über die Konfiguration einzelner Standorte, die Dimensionierung der Links oder die Routenplanung für die Bedarfsanforderungen. In langjährigen Projekten mit unseren Partnern E-Plus und Telekom Austria haben wir entsprechende Aufgabenstellungen mit dem Ziel untersucht, automatisierte Verfahren zur Planung von kosteneffizienten und ausfallsicheren SDH-Netzen zu entwickeln [1, 2, 26].

2.2.1 Standorte, Hardware und Kapazitäten

Die Standorte in SDH-Netzen sind durch die Quellen und Senken der Kommunikationsbedarfe fast vollständig definiert. Zur Disposition stehen normalerweise nur Verteilerstandorte, die weder Quelle noch Senke sind. An einem Standort befinden sich neben den Geräten anderer Netzschichten, über die Bedarfe eingespeist werden, die Geräte des SDH-Netzes. Als Netzelemente sind dies Add-Drop-Multiplexer (ADMs) oder Digital Cross-Connects (DXCs). Bei deren Konfiguration ist zu berücksichtigen, dass das sogenannte Koppelfeld als zentrale Einheit bei beiden Typen unter anderem durch die Schaltkapazität bestimmt ist, welche die an diesen ADM oder DXC anschließbaren Linkkapazitäten beschränkt. Die Netzelemente verschiedener Hersteller bzw. die unterschiedlichen Typen eines Herstellers unterscheiden sich ferner in der Anzahl und der Art der Steckplätze (Slots) für Schnittstellenkarten, wobei nicht jede Schnittstellenkarte in jeden Slot passt. Die Karten stellen einen oder mehrere STM-N-Ports zur Verfügung.

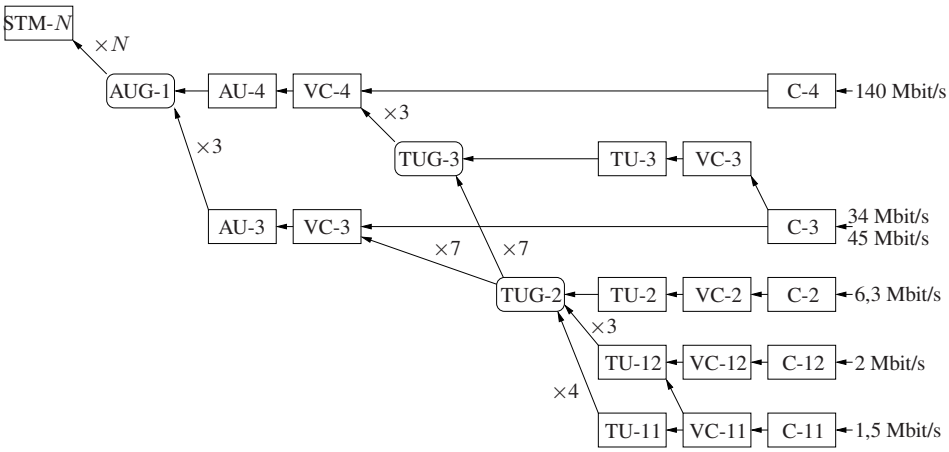


Abbildung 4: VC-N-Multiplexing-Hierarchie.

Die Netztopologie und die Linkkapazitäten entstehen durch den Anschluss von Fasern an den entsprechenden Port der Karten, wobei in der Praxis immer Faserpaare verwendet werden, die die bidirektionale Übertragung von STM-N-Signalen erlauben. Soll die Kapazität auf einem Link erhöht werden, so ist also entweder an beiden Faserenden eine Schnittstelle mit größerer Kapazität zu verwenden oder ein weiteres Faserpaar an eine weitere Schnittstelle anzuschließen.

Zudem ist festzulegen, welche Geräte von welchen Herstellern an welchen Standorten verwendet werden sollen, wie diese mit Schnittstellenkarten zu bestücken und welche Links mit welchen Kapazitäten einzurichten sind. Ferner wird nur selten auf der grünen Wiese geplant, und deshalb ist eine ggf. bereits vorhandene SDH-Infrastruktur in die Planung einzubeziehen. Speziell bei der Ausbauplanung ist die gegenwärtige Konfiguration der Hardware zu berücksichtigen, die häufig nur erweitert oder in sehr eingeschränkter Weise verändert werden darf.

2.2.2 Routing und Ausfallsicherheit

In der SDH-Technologie werden die Nutzsignale in standardisierten Containern fester Bandbreite (VC-11 ≅ 1,5 Mbit/s, VC-12 ≅ 2 Mbit/s usw.) übertragen, die nach einem vordefinierten Multiplexing-Schema in Signale fester Bandbreite (STM-1 ≅ 155 Mbit/s, STM-4 ≅ 622 Mbit/s usw.) eingebettet werden. Abbildung 4 illustriert diesen Zusammenhang. In der Planung ist dementsprechend ein statisches Routing der VC-N-Container festzulegen, deren Übertragungswege zwischen Paaren von Netzelementen im Netzwerkmanagement konfiguriert werden.

Über das SDH-Netz werden Nutzdaten aus heterogenen Bedarfsquellen übertragen. Diese Daten können unter anderem aus dem Sprach-, Signalisierungs-, ATM- oder auch IP-Netz kommen. Jede einzelne Bedarfsanforderung ist gemäß den individuellen Sicherheitsanforderungen auf einem oder mehreren Wegen zu realisieren. So werden die Bedarfe aus dem Signalisierungsnetz typischerweise mit sehr hohen Anforderungen an die Sicherheit geroutet, wohingegen bei Mietleitungskunden die Anforderungen je nach Tarifgruppe differenziert sein können. Wird Sub-Network Connection Protection (SNCP) als Schutzmechanismus eingesetzt, so ist für jede Bedarfsanforderung zu entscheiden, auf welchem Paar von unabhängigen Wegen diese geroutet werden soll. Alternativ können zur Realisierung von Multiplex Section Protection (MSP) als Schutzmechanismus einzelne Links durch die Festlegung und Konfiguration von

zugehörigen Ersatzwegen geschützt werden. In der operativen Planung steht die (Re-)Konfiguration einzelner VC-N im Vordergrund, während bei der strategischen Planung auch die Entscheidung zu treffen ist, welcher Schutzmechanismus verwendet werden soll.

2.2.3 Planungsaufgabe

Das Ziel ist es, ein kostengünstiges SDH-Netz so zu entwerfen, dass die zu realisierenden Bedarfe gemäß ihrer Routinganforderungen ausfallsicher übertragen werden können, wobei Kosten für SDH-Netzelemente, für Schnittstellenkarten sowie für Kabelrohre und Fasern anfallen. Eine wichtige Aufgabe der SDH-Netzplanung lässt sich folgendermaßen zusammenfassen:

PLANUNGSAUFGABE
SDH-Netze

Planungsziel: Minimierung der Kosten

Entscheidungen:

- Topologie
- Netzelemente inkl. Bestückung mit Schnittstellenkarten
- STM-N-Linkkapazitäten
- Routing in VC-N-Containern

Randbedingungen:

- Heterogene Bedarfsquellen
- Sicherheitsanforderungen
- Existierende Netzkonfiguration und VC-N-Beschaltung

2.3 Design optischer Netze

In modernen Telekommunikationsnetzen wird insbesondere im Kernnetz zunehmend optische Technologie eingesetzt, die neben hohen Übertragungsbandbreiten seit kurzem auch die notwendige Schaltflexibilität bereitstellt. Der folgende Abschnitt beschreibt dafür Planungsaufgaben, für die wir in mehrjährigen Kooperationen mit der Telekom Austria sowie der T-Systems Nova speziell zugeschnittene Lösungsverfahren zum kostenoptimalen Design optischer Netze entwickelt haben [18, 29].

2.3.1 Hardware und Kapazitäten

Ein optisches Netz setzt sich aus unterschiedlichen Hardware-Komponenten zusammen, deren Kombination die Übertragungskapazität auf den Links und die Schaltkapazität in den Knoten zur Verfügung stellt. In Abbildung 5 sind die einzelnen optischen Komponenten schematisch dargestellt.

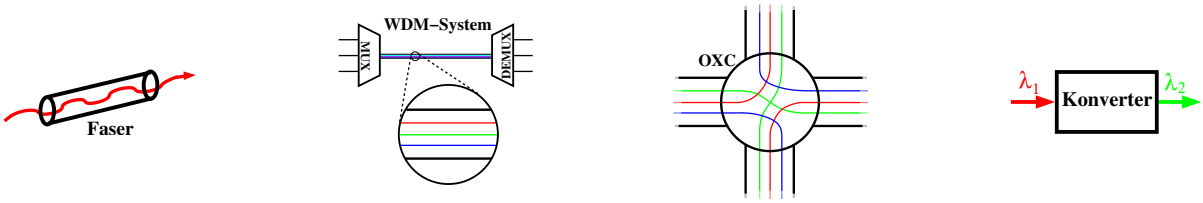


Abbildung 5: Optische Komponenten: Glasfaser, WDM-System, OXC, Konverter.

Den Grundbaustein optischer Netze stellen Glasfasern dar, auf denen Informationen in Form kurzer Lichtpulse übertragen werden. Zudem erlaubt das optische Medium das Wavelength Division Multiplexing (WDM), bei dem auf einer Glasfaser durch Verwendung verschiedener Wellenlängen mehrere Kanäle parallel genutzt werden können. Dazu ist auf der Faser ein entsprechendes WDM-System zu installieren. Jeder Link kann mehrere parallele Fasern mit jeweils einem eigenen WDM-System enthalten. Die Übertragungskapazität des Links ergibt sich dann aus der Summe der bereitgestellten Kanäle. Durch die stetige Weiterentwicklung der optischen Technologie stehen verschiedene Faser- und WDM-System-Typen mit unterschiedlichen Eigenschaften und Kosten zur Verfügung. Die WDM-Systeme variieren in der Anzahl und den möglichen Wellenlängen der kombinierbaren Kanäle. Fasertypen unterscheiden sich in optisch relevanten Eigenschaften wie ihrer Dispersion und Dämpfung, die sowohl das optische Signal beeinflussen als auch die Verwendbarkeit bestimmter WDM-Systeme bedingen. Infolgedessen sind nicht alle Faser- und WDM-System-Typen miteinander verträglich; dies muss in der Planung berücksichtigt werden.

Optische Cross-Connects (OXCs) und Optische ADMs sind die Schaltzentralen optischer Netze. Sie besitzen eine Reihe von Ports, die einander zuordenbar sind und so jeweils einen optischen Kanal beliebiger Wellenlänge ohne optisch-elektronische Wandlung direkt durchschalten können. Durch die Integration dieser Schalteinheiten entstehen durchgehend optische Verbindungen, die als charakteristisches Merkmal die eigentlichen optischen Netze kennzeichnen. OXCs gibt es ebenfalls in verschiedenen Varianten, die sich in der Anzahl vorhandener Ports bzw. schaltbarer Kanäle sowie in den Kosten unterscheiden. Das Schalten eines optischen Kanals ist unabhängig von seiner Wellenlänge. Allerdings entsteht ein Konflikt, wenn die verwendete Wellenlänge auf einer Faser nicht verfügbar oder bereits durch einen anderen Kanal blockiert ist. Solche Konflikte lassen sich mit Hilfe von Wellenlängenkonvertern vermeiden, die innerhalb eines Kanals das Wechseln auf eine beliebige, noch verfügbare Wellenlänge erlauben. Die Verwendung von Wellenlängenkonvertern ist allerdings mit zusätzlichen Kosten verbunden.

Die Kombination aller installierten Komponenten stellt die Hardware-Konfiguration des optischen Netzes dar. Ihre kosteneffiziente Gestaltung erfordert natürlich die Einbeziehung des bereits vorhandenen Equipments. Zusätzliche operative Restriktionen ergeben sich durch eingeschränkte Verfügbarkeit der Komponenten oder begrenzten Platz an einem Standort. Der Entwurf einer geeigneten Konfiguration muss sicherstellen, dass alle Kapazitäten ausreichend dimensioniert sind, um die geforderten Bedarfe innerhalb des Netzes routen zu können.

2.3.2 Routing und Ausfallsicherheit

In optischen Netzen wird der Verkehr auf sogenannten Lichtwegen geroutet. Ein Lichtweg ist eine durchgängig optische Verbindung zweier Knoten im Netz über einen oder mehrere Links. Die auf einem Lichtweg verfügbare Übertragungsbandbreite beträgt üblicherweise 2.5 oder 10 Gbit/s. Da in der

optischen Netzschicht nur ganze Lichtwege geschaltet werden können, werden die dort zu realisierenden Verbindungsbedarfe direkt in der Anzahl benötigter Lichtwege zwischen den Knoten angegeben.

Um alle Bedarfe im Netz zu erfüllen, ist eine geeignete Lichtweg-Konfiguration zu bestimmen. Dies umfasst sowohl das Routing der benötigten Lichtwege im physikalischen Netz, wobei die durch die Hardware-Konfiguration gegebenen Kapazitätsschranken zu berücksichtigen sind, als auch eine konfliktfreie Wellenlängenzuordnung, so dass je zwei Lichtwege auf der gleichen Faser unterschiedliche, dort verfügbare Wellenlängen verwenden. Sofern ein Lichtweg auf zwei aufeinander folgenden Links verschiedene Wellenlängen nutzt, benötigt er im dazwischenliegenden Knoten einen Wellenlängenkonverter. Aufgrund der hohen Konverterkosten werden dabei Lichtweg-Konfigurationen bevorzugt, die möglichst wenige Konvertierungen erfordern. Ferner ist bei der Planung der Lichtweg-Konfigurationen die durch Dispersion und Dämpfung begrenzte optische Übertragungslänge zu beachten. Wird diese überschritten, müssen zusätzliche Regeneratoren eingesetzt werden.

Schließlich spielt gerade im Kernnetz die Ausfallsicherheit eine zentrale Rolle. Da über Lichtwege enorme Datenmengen transportiert werden, wird der Ausfallschutz oft ebenfalls auf Ebene der Lichtwege realisiert und ist bei der Planung der Lichtweg-Konfigurationen einzubeziehen. Die Problematik der Wellenlängenzuordnung zu den Lichtwegen erschwert eine Nutzung gemeinsamer Reservekapazitäten, so dass Ausfallsicherheit in optischen Netzen meist einfacher durch die zusätzliche Bereitstellung von individuellen Ersatz-Lichtwegen für jedes Knotenpaar realisiert wird. In Abschnitt 5.3 werden wir dies mit einem neu entwickelten Schutzkonzept für optische Netze vergleichen.

2.3.3 Planungsaufgabe

Beim Design optischer Netze besteht das Ziel darin, zur Realisierung der gegebenen Verkehrsanforderungen eine geeignete Lichtweg-Konfiguration zu ermitteln, die sich mit einer möglichst kostengünstigen Hardware-Konfiguration umsetzen lässt. Dabei liegen den einzelnen Komponenten individuelle Kostenmodelle zugrunde. Die zahlreichen Entscheidungen und Bedingungen dieser Planungsaufgabe sind nachfolgend zusammengefasst:

PLANUNGSAUFGABE
Optische Netze

Planungsziel: Minimierung der Gesamtkosten des optischen Netzes

Entscheidungen:

- Physikalische Topologie
- Hardware-Konfiguration: Glasfasern, WDM, OXCs, Konverter
- Lichtweg-Konfiguration: Routing und Wellenlängenzuordnung

Randbedingungen:

- Einbeziehung vorhandenen Equipments
- Technische Restriktionen: z. B. unzulässige Faser-WDM-Kombinationen
- Operative Restriktionen: z. B. beschränkte Komponentenanzahlen
- Ausfallsicherheitsanforderungen

2.4 Traffic Engineering in IP-Netzen

Traffic Engineering ist eine Aufgabe aus der operativen Netzplanung, bei der das Routing der Verkehrsflüsse an das Übertragungsnetz angepasst wird. Das aktuelle Übertragungsnetz ist dabei gegeben, kann aber möglicherweise innerhalb gewisser Grenzen verändert werden. Das Ziel dieser Planung ist die Verbesserung der Übertragungsqualität im Netz. In diesem Abschnitt diskutieren wir als Anwendungsfall das Traffic Engineering in IP-Netzen, wie wir es in Projekten mit dem DFN-Verein e. V. für den Betrieb des Breitband- und des Gigabit-Wissenschaftsnetzes B-WiN bzw. G-WiN untersucht haben [4, 5, 6].

2.4.1 Hardware und Kapazitäten

Am Standort eines IP-Netzknotens können, je nach technischer Realisierung des Netzes, mehrere Geräte vorhanden sein. Das zentrale Netzelement in IP-Netzen ist der IP-Router. Dieser hat eine typabhängige maximale Routingkapazität und kann mit unterschiedlichen Routerkarten bestückt werden.

Zwischen den IP-Knoten sind Links mit individuellen Kapazitätsstufen installiert. Die möglichen Kapazitätsstufen ergeben sich häufig aus der verwendeten Technologie, wie etwa 155 Mbit/s, 622 Mbit/s usw. bei IP-over-SDH-Netzen. Prinzipiell sind aber auch beliebige feste Kapazitäten möglich, zum Beispiel reservierte VP-Kapazitäten in ATM-VPNs. Ist auf einem Link eine spezielle Kapazitätsstufe installiert, so müssen an den Endknoten passende Schnittstellen an den IP-Routerkarten vorhanden sein.

Die Konfiguration des Übertragungsnetzes ist beim Traffic Engineering bereits gegeben. Je nach praktischem Handlungsspielraum ist diese Konfiguration bei der Planung entweder gar nicht oder nur innerhalb bestimmter Grenzen veränderbar. Bei der Rekonfiguration sind dabei neben den üblichen technischen Randbedingungen in der Regel noch zusätzliche Budgets für Änderungskosten oder Investitionen in neue Hardware zu berücksichtigen. Da das Ein- oder Ausbauen von Routerkarten oder Verändern von Linkkonfigurationen in der Praxis mit erheblichem personellen Aufwand verbunden ist, beschränkt man häufig auch die mögliche Anzahl solcher Änderungen bei der Planung.

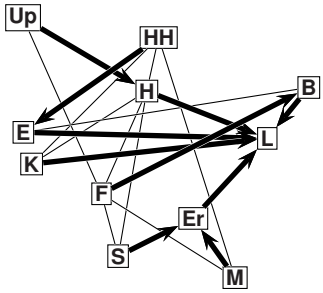
2.4.2 Routing und Ausfallsicherheit

Die Verkehrsanforderungen für das Traffic Engineering kommen für gewöhnlich direkt oder indirekt aus dem Accounting im aktuellen Netz. Sie werden mit dem OSPF-Routingprotokoll, dem derzeitigen IP-Standardprotokoll, durch das Netz geroutet. Im Betrieb des Netzes ist dieses Protokoll einfach: Zu vorgegebenen Routinggewichten für die Links werden gewichtsminimale Wege berechnet, entlang denen der Verkehr geroutet wird, siehe Abbildung 6. Die zentrale Aufgabe beim Traffic Engineering ist es, diese Routinggewichte so zu bestimmen, dass die sich aus den gewählten Gewichten ergebenden Routingwege und Verkehrsflüsse die Kapazitäten möglichst gleichmäßig auslasten. Insbesondere soll keiner der Links übermäßig stark ausgelastet werden, da dies bei den im Internet beobachtbaren Verkehrsschwankungen leicht zu einem Engpass führt.

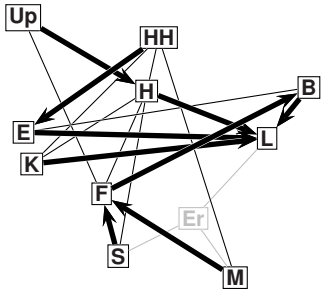
Bei der Wahl der Routinggewichte sind ferner alle möglichen Ausfälle von einzelnen Knoten oder Links zu berücksichtigen.

Die Routingwege sollen häufig weitere Anforderungen erfüllen. Beispielsweise soll das Routing aus administrativen Gründen symmetrisch sein, oder keiner der Routingwege darf mehr als eine bestimmte Anzahl von Links enthalten, um ausreichende Übertragungsqualität für Streaming-Anwendungen im Netz garantieren zu können. Wird im Netz die Basisvariante des OSPF-Protokolls verwendet, muss es außerdem für jedes Quelle-Senke-Paar genau einen eindeutigen kürzesten Weg bezüglich der gewählten Routinggewichte geben. Andernfalls ist nicht sichergestellt, dass die in der Planung bestimmten Routingwege auch tatsächlich verwendet werden.

Beim OSPF-Routing (**O**pen **S**hortest **P**ath **F**irst) wird jedem Link ein Gewicht zugeordnet. Jedes Datenpaket wird dann auf einem bezüglich dieser Gewichte kürzesten Weg zu seinem Ziel geroutet.



Die Routingwege nach Leipzig (L) bilden einen umgekehrten Wurzelbaum.



Bei Ausfall des Knotens Erlangen (Er) werden die kürzesten Wege aktualisiert. Vom Ausfall nicht betroffene Wege bleiben gleich.

Abbildung 6: OSPF-Routing.

2.4.3 Planungsaufgabe

Ziel der Planung ist es, Routinggewichte und ggf. eine Rekonfiguration der Hardware zu finden, so dass für die sich daraus ergebenden Routingwege die maximale Kapazitätsauslastung aller Links minimal ist. Empirische Untersuchungen haben gezeigt, dass die Kapazitätsauslastung der Links und der Router generell ein geeignetes Maß für die Übertragungsqualität im Netz ist. Häufig gibt es jedoch in den Netzen einige spezielle Links, die unabhängig von der Wahl der Routinggewichte immer extrem belastet sind, wie zum Beispiel Links zu zentralen Servern oder Gateways in andere Netze. Das hätte zur Folge, dass auf Kosten der Qualität im übrigen Netz nur noch die maximale Kapazitätsauslastung auf diesen Links minimiert wird. Um dies zu vermeiden, werden solche Links differenziert betrachtet. Das Traffic Engineering kann als Aufgabe wie folgt zusammengefasst werden:

PLANUNGSAUFGABE
Traffic Engineering in IP-Netzen

Planungsziel: Minimierung der maximalen Linkauslastung

Entscheidungen:

- Rekonfiguration der Knoten und Links
- OSPF-Routinggewichte

Randbedingungen:

- OSPF-Routing
- Ausreichende Dimensionierung für Routing der Bedarfe
- Kostenbudgets und Rekonfigurationsbeschränkungen
- Ausfallsicherheitsanforderungen

3 Modelle: Wie sieht's mathematisch aus?

Jede der zuvor beschriebenen Planungsaufgaben ist sehr komplex und „von Hand“ praktisch nicht lösbar. Schon bei relativ kleinen Netzen ist die Anzahl von Konfigurationen der Hardware oder von Wegen zum Routen der Bedarfe unüberschaubar groß. Die mathematische Optimierung hat sich in diesem Umfeld als eines der leistungsfähigsten Instrumente zur Berechnung qualitativ hochwertiger Lösungen erwiesen.

Das Fundament dieser Methoden sind die mathematischen Modelle, deren Grundprinzipien wir im folgenden Abschnitt beschreiben. Anschließend erläutern wir, wie diese dazu genutzt werden können, viele Planungs- und Optimierungsaufgaben im Bereich der Telekommunikation zu erfassen. Die Abschnitte 3.2 und 3.3 führen die zwei wesentlichen Modellbausteine ein, mit denen die Netzplanungsaufgaben aus Kapitel 2 mathematisch abgebildet werden können. Dies ist zum einen die Modellierung von Hardware-Konfigurationen und zum anderen die Modellierung der Routingaufgabe. Diese beiden Modellbausteine werden, wie in Abbildung 7 skizziert, über Linkkapazitäten miteinander verknüpft. Die Integration der Bausteine in ein Gesamtmodell, das die Berechnung von Lösungen bezüglich eines übergreifenden Optimierungszieles erlaubt, beschreiben wir in Abschnitt 3.4.

Abbildung 7: Modellbausteine in der Netzplanung.

3.1 Modellierung mathematischer Optimierungsprobleme

Im Rahmen dieses Kapitels fassen wir die skizzierten Planungsaufgaben verallgemeinernd als Optimierungsprobleme auf. Um zu ihrer Lösung Verfahren der mathematischen Optimierung einsetzen zu können, ist es notwendig, die Probleme zunächst abstrakt in Modellen darzustellen. Dies erfolgt durch Variablen, Parameter, Nebenbedingungen und eine Zielfunktion. Meist lässt sich ein praktisches Problem auf mehrere Arten in einem mathematischen Modell abbilden, die praktische Lösbarkeit hängt dann jedoch oft von den Eigenschaften des verwendeten Modells ab. Wir beschreiben und verwenden im Folgenden die in der Praxis erprobten gemischt-ganzzahligen linearen Modelle.

Bei der Modellierung eines Optimierungsproblems wird zwischen Parametern und Variablen unterschieden. Die Parameter eines Modells sind diejenigen Werte, die die Rahmenbedingungen und Eingabedaten des Optimierungsproblems beschreiben. Parameter sind gegebene bzw. planerisch vordefinierte Werte. Die noch zu treffenden quantitativen Entscheidungen werden durch Variablen abgebildet. Abhängig von der Art der Entscheidung sind zwei Typen von Variablen zu unterscheiden:

- Ganzzahlige Variablen dürfen nur ganze Zahlen als Wert annehmen und werden für Entscheidungen verwendet, die von diskreter Natur sind, beispielsweise die Anzahl von Netzelementen an einem Standort.
- Kontinuierliche Variablen dürfen jeden reellen Wert annehmen und werden für Entscheidungen verwendet, bei denen nicht nur ganzzahlige Vielfachheiten zulässig sind, sondern auch jeder beliebige Zwischenwert. Typisches Beispiel in der Netzplanung ist die relative Auslastung eines Links.

Beide Typen von Variablen sind normalerweise durch einen Minimum- und einen Maximumwert eingegrenzt. Bei ganzzahligen Variablen unterscheidet man in diesem Zusammenhang noch den Spezialfall, bei dem der Minimumwert gleich 0 und der Maximumwert gleich 1 ist. Diese Variablen werden auch binäre Variablen genannt, da sie nur die Werte 0 und 1 annehmen können. Binäre Variablen werden oft genutzt, um „Ja/Nein“-Entscheidungen abzubilden.

Die Beziehungen zwischen den Variablen und Parametern werden durch Gleichungen und Ungleichungen dargestellt, die auch Nebenbedingungen genannt werden. Eine zulässige Lösung der Optimierungsaufgabe entspricht einer Belegung der Variablen, bei der alle Nebenbedingungen erfüllt sind. Eine Bewertung jeder zulässigen Lösung wird schließlich mit Hilfe einer Zielfunktion bestimmt. Im Prinzip können die Bedingungen sowie die Zielfunktion beliebige mathematische Ausdrücke enthalten. Um die Modelle aber praktisch lösen zu können, versucht man häufig, sich auf eine lineare Darstellung der Beziehungen zu beschränken. Das auf diese Weise entstehende Modell wird gemischt-ganzzahliges lineares Programm genannt.

Der Prozess der Modellbildung dient in erster Linie dazu, eine verbal mehr oder weniger genau umrissene Aufgabenstellung in eindeutiger Weise so zu formulieren, dass automatisierte und effiziente Lösungsverfahren zur Berechnung optimaler Lösungen verwendet werden können. Sowohl mit der Modellbildung als auch mit der Entwicklung entsprechender Verfahren beschäftigt sich eine ganze mathematische Fachrichtung: die Mathematische Programmierung. In Kapitel 4 werden einige der verwendeten Methoden zur Lösung gemischt-ganzzahliger linearer Programme beschrieben. Softwarelösungen wie CPLEX [13], SOPLEX [28] oder XPRESS-MP [9] sind mittlerweile in der Lage, lineare Programme (d. h. ohne ganzzahlige Variablen) mit Millionen von Variablen in kürzester Zeit zu lösen. Große gemischt-ganzzahlige Programme (die sowohl kontinuierliche als auch ganzzahlige Variablen enthalten) lassen sich allerdings nur dann effizient lösen, wenn die speziellen Struktureigenschaften der Modelle ausgenutzt werden.

3.2 Hardware und Netzkonfigurationen

Eine mathematische Modellierung von Hardware- und Netzkonfigurationen muss in der Lage sein, technologieabhängige Zusammenhänge korrekt abzubilden, da sich beispielsweise die Spezifikation von Routern in IP-Netzen wesentlich von der für Add-Drop-Multiplexer in SDH-Netzen unterscheidet. Von einem ausreichend abstrakten Standpunkt aus betrachtet, zeichnen sich jedoch viele Bedingungen durch eine gemeinsame, technologie- und herstellerunabhängige Struktur aus. Alle benötigten Ressourcen müssen auch bereitgestellt werden. Auf Basis dieses Prinzips stellen wir im Folgenden ein allgemeines Modell vor, das als flexible Grundlage bei der Modellierung von Netzplanungsproblemen dienen kann. Insbesondere lassen sich damit alle in den Planungsproblemen skizzierten Technologien und ihre Eigenschaften erfassen. Dieser Ansatz ist detailliert in [19] beschrieben.

3.2.1 Komponenten und Ressourcen

Eine Möglichkeit, die verwendete Hardware und ihre technischen Details zu modellieren, ist die Verwendung von *Komponenten* und *Ressourcen*. Hierbei werden Dinge, die in festgelegten Vielfachheiten an bestimmten Orten innerhalb des Netzwerkes installiert werden können, durch Komponenten dargestellt. Es kann sich dabei sowohl um konkrete Gegenstände handeln, beispielsweise eine bestimmte Routerkarte, die in einem Netzelement verwendet wird, als auch um abstrakte Konzepte, wie die Bandbreite, die in fest definierten Einheiten auf einem Link reserviert wird. Ressourcen stellen begrenzte „Güter“ dar,

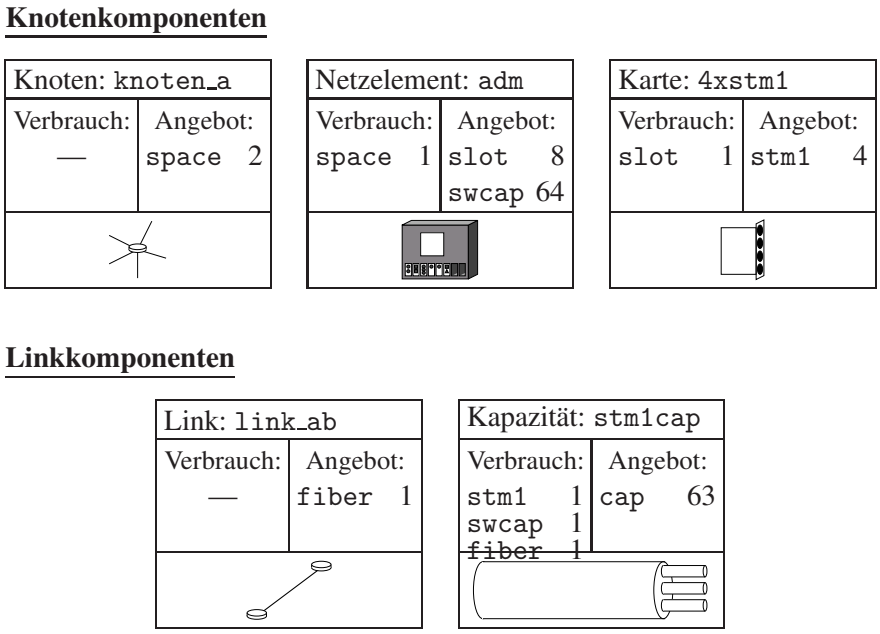
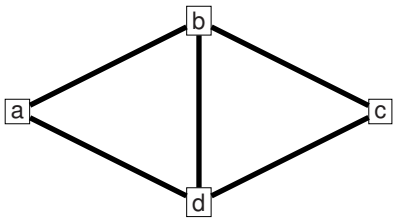


Abbildung 8: Beispiel: typische Komponenten und Ressourcen.

die von einigen Komponenten bereitgestellt und von anderen verbraucht werden. Dabei darf an keiner Stelle im Netz der Verbrauch das Angebot übersteigen. Ebenso wie Komponenten können Ressourcen sowohl physikalische Gegebenheiten, wie freie Schnittstellen an einer Routerkarte, als auch künstliche Begrenzungen abbilden, etwa ein vorgegebenes Kontingent eines bestimmten Gerätetypes.

Beispiel: In Abbildung 8 sind exemplarisch einige Komponenten mit zugehörigen Ressourcen dargestellt, die bei der Planung und Konfiguration von SDH-Netzen verwendet werden. Die Komponente `knoten_a` spezifiziert die mögliche Anzahl von Netzelementen in diesem Knoten, wobei hier durch die Ressource `space` der Platz für maximal zwei Netzelemente angeboten wird. Als Netzelement kommt hier etwa ein ADM in Frage, der vereinfacht aus einem Koppelfeld und einer Reihe von Steckplätzen besteht. Zu seiner Modellierung wird die Komponente `adm` eingeführt, deren Verwendung dem Aufstellen dieses Gerätes entspricht. Mit der Ressource `space` wird für diesen ADM die Anzahl der im Knoten benötigten Stellplätze beschrieben und mit der Ressource `swcap` die durch das Koppelfeld maximal zur Verfügung gestellte Schaltkapazität. Die Steckplätze bieten den Platz für Schnittstellenkarten, die ebenfalls durch Komponenten repräsentiert werden. Hier beschreibt die Komponente `4xstm1` eine solche Karte. Der ADM hat acht Steckplätze, von denen jeder eine Karte aufnehmen kann. Die Ressource `slot` wird verwendet, um das Angebot von 8 Steckplätzen durch das Netzelement `adm` und den Verbrauch je eines Steckplatzes durch die Schnittstellenkarte `4xstm1` zu spezifizieren. Die Schnittstellenkarten stellen Ports zum Anschluss von Glasfasern bereit. Eine `4xstm1`-Karte hat vier Ports vom Typ STM-1. Dies wird mit der Ressource `stm1` abgebildet. Die Komponente `link_ab` spezifiziert mit der Ressource `fiber` die prinzipielle Möglichkeit, die zwei Knoten a und b miteinander zu verbinden. Eine Faser, die zwei Knoten über je einen Port des Typs STM-1 verbindet, wird durch die Komponente `stm1cap` modelliert, die an jedem Endknoten eine Einheit `stm1` verbraucht und dafür die Ressource `cap` als Übertragungskapazität zum Routen von VC-12 bereitstellt.



Knoten		knoten_a	knoten_b	knoten_c	knoten_d
Angebot	space	2	1	2	1

Links		link_ab	link_bc	link_cd	link_ad	link_bd
Angebot	fiber	1	1	1	1	1

ADMs		adm1	adm2
Verbrauch	space	1	1
Angebot	slot	8	16
	swcap	64	128

Schnittstellenkarten		4xstm1	1xstm16
Verbrauch	slot	1	1
Angebot	stm1	4	0
	stm16	0	18

Kapazitäten		stm1cap	stm16cap
Verbrauch	fiber	1	1
	stm1	1	0
	stm16	0	1
	swcap	1	16
Angebot	cap	63	1008

Abbildung 9: Beispiel: Netzwerk mit allen Komponenten und Ressourcen.

Ein komplettes Beispiel ist in Abbildung 9 dargestellt. Für jeden Knoten und jeden Link des zugrunde liegenden Netzes wird eine Komponente eingeführt, die die Ressourcen `space` für je einen oder zwei ADMs bzw. `fiber` für je eine Linkkapazität zur Verfügung stellen. In den Knoten können zwei verschiedene ADM-Typen installiert werden, die den Komponenten `adm1` und `adm2` mit 8 bzw. 16 Steckplätzen und einer Schaltkapazität von 64 bzw. 128 STM-1 entsprechen. In beiden ADMs sind zwei Typen von Schnittstellenkarten verwendbar, die durch die Komponenten `4xstm1` und `1xstm16` repräsentiert werden. Außerdem stehen zwei Linkkapazitäten zur Verfügung: STM-1 und STM-16, für deren Installation jeweils eine freie Faser sowie eine passende Schnittstelle und hinreichende Schaltkapazität in den zu verbindenden Knoten nötig ist.

Durch das vorgestellte Konzept, die Modellierung einer Konfiguration anhand von Komponenten und Ressourcen vorzunehmen, können viele Nebenbedingungen an zulässige Konfigurationen einfach erfasst werden. Dadurch entsteht, ausgehend von Knoten und Links als elementaren Einheiten, schließlich die gesamte Konfiguration als Kaskade von bereitgestellten und verbrauchten Ressourcen, an deren Ende Kapazitäten zum Routen der Kommunikationsbedarfe zur Verfügung gestellt werden.

3.2.2 Mathematische Umsetzung

Die Eigenschaften der Komponenten und Ressourcen werden nun mit Variablen und Parametern in einem mathematischen (Teil-)Modell umgesetzt, dessen Lösungen den technisch realisierbaren Netzkonfigurationen entsprechen. Zur Beschreibung einer Komponenteninstallation werden ganzzahlige Variablen verwendet. Für eine Komponente k und einen Knoten oder Link g, an dem k verfügbar ist, gibt der Wert der Variablen $x_{k,g}$ an, wie oft k bei g zu installieren ist. Da Komponenten in ganzen Stückzahlen angeschafft werden, dürfen diese Variablen nur ganzzahlige positive Werte oder Null annehmen. Sowohl der Verbrauch als auch die Bereitstellung einer Ressource lässt sich mittels Parametern darstellen: Für jede Ressource r beschreibt $A(k,r)$, wieviele Einheiten der Ressource r von der Komponente k angeboten werden, und $V(k,r)$, wieviele von ihr verbraucht werden. Weiterhin bezeichnet die Menge K_g die in g installierbaren Komponenten.

Für eine Menge B von Knoten und Links im Netzwerk kann nun mittels

$$A_x(B,r) := \sum_{g \in B} \sum_{k \in K_g} A(k,r) \cdot x_{k,g}$$

ausgedrückt werden, wieviele Einheiten der Ressource r durch die Komponenten auf B in Summe bereitgestellt werden. Analog beschreibt

$$V_x(B,r) := \sum_{g \in B} \sum_{k \in K_g} V(k,r) \cdot x_{k,g}$$

den Gesamtbedarf der Ressource r aller Komponenten auf den Knoten und Links in B. Damit lässt sich nun durch die Ungleichung

$$A_x(B,r) = \sum_{g \in B} \sum_{k \in K_g} A(k,r) \cdot x_{k,g} \geq \sum_{g \in B} \sum_{k \in K_g} V(k,r) \cdot x_{k,g} = V_x(B,r) \tag{1}$$

einfach ausdrücken, dass von jeder Ressource mindestens soviel bereitgestellt werden muss, wie verbraucht wird. Dies ist eine lineare Ungleichung mit ganzzahligen Variablen und damit genau die Art von Bedingung, die in der ganzzahligen linearen Programmierung eingesetzt wird. Bei geeigneter Wahl der Mengen B entsteht durch Formulierung aller Ungleichungen vom Typ (1) für alle Ressourcen ein Ungleichungssystem, dessen Lösungen alle technisch realisierbaren Netzwerkinstallationen beschreiben.

Mit jeder Netzkonfiguration ist schließlich auch die zum Routen der Kommunikationsbedarfe nutzbare Kapazität spezifiziert. Mit der speziellen Ressource `cap` lässt sich die insgesamt auf einem Link ℓ verfügbare Routingkapazität

$$C_\ell := \sum_{k \in K_\ell} A(k, \texttt{cap}) \cdot x_{k,\ell} \tag{2}$$

ausdrücken. Zur Veranschaulichung wird anhand des folgenden Beispieles ein solches Modell dargestellt.

Beispiel: Zur mathematischen Modellierung des im vorigen Abschnitt beschriebenen Beispieles werden nun für alle Knoten- und Linkkomponenten Variablen eingeführt. So repräsentiert $x_{\texttt{adm1},a}$ die Anzahl von Netzelementen vom Typ `adm1`, die am Knoten a installiert werden. Gleichermaßen bestimmt Variable $x_{\texttt{stm16cap},bc}$, ob die Linkkapazität STM-16 auf Link bc installiert wird. Die sich daraus ergebende mathematische Modellierung ist in Abbildung 10 dargestellt. Ungleichung (`space` für b) etwa beschränkt die Nutzung der Ressource `space` in Knoten b durch Platzierung von ADMs auf den vorhandenen Platz. Da dort nur ein Platz zur Verfügung steht, darf die Summe von $x_{\texttt{adm1},b}$ und $x_{\texttt{adm2},b}$ maximal eins betragen.

$$
\begin{array}{rrcll}
1x_{\texttt{adm1},a} & +1x_{\texttt{adm2},a} & \leq & 2 & (\texttt{space} \text{ für } a) \\
1x_{\texttt{4xstm1},a} & +1x_{\texttt{1xstm16},a} & \leq & 8x_{\texttt{adm1},a} + 16x_{\texttt{adm2},a} & (\texttt{slot} \text{ für } a) \\
1x_{\texttt{stm1cap},ab} & +16x_{\texttt{stm16cap},ab} & & & \\
+1x_{\texttt{stm1cap},ad} & +16x_{\texttt{stm16cap},ad} & \leq & 64x_{\texttt{adm1},a} + 128x_{\texttt{adm2},a} & (\texttt{swcap} \text{ für } a) \\
1x_{\texttt{stm1cap},ab} & +1x_{\texttt{stm1cap},ad} & \leq & 4x_{\texttt{4xstm1},a} & (\texttt{stm1} \text{ für } a) \\
1x_{\texttt{stm16cap},ab} & +1x_{\texttt{stm16cap},ad} & \leq & 1x_{\texttt{1xstm16},a} & (\texttt{stm16} \text{ für } a) \\
1x_{\texttt{adm1},b} & +1x_{\texttt{adm2},b} & \leq & 1 & (\texttt{space} \text{ für } b) \\
1x_{\texttt{4xstm1},b} & +1x_{\texttt{1xstm16},b} & \leq & 8x_{\texttt{adm1},b} + 16x_{\texttt{adm2},b} & (\texttt{slot} \text{ für } b) \\
1x_{\texttt{stm1cap},ab} & +16x_{\texttt{stm16cap},ab} & & & \\
+1x_{\texttt{stm1cap},bc} & +16x_{\texttt{stm16cap},bc} & & & \\
+1x_{\texttt{stm1cap},bd} & +16x_{\texttt{stm16cap},bd} & \leq & 64x_{\texttt{adm1},b} + 128x_{\texttt{adm2},b} & (\texttt{swcap} \text{ für } b) \\
1x_{\texttt{stm1cap},ab} & & & & \\
+1x_{\texttt{stm1cap},bc} & +1x_{\texttt{stm1cap},bd} & \leq & 4x_{\texttt{4xstm1},b} & (\texttt{stm1} \text{ für } b) \\
1x_{\texttt{stm16cap},ab} & & & & \\
+1x_{\texttt{stm16cap},bc} & +1x_{\texttt{stm16cap},bd} & \leq & 1x_{\texttt{1xstm16},b} & (\texttt{stm16} \text{ für } b) \\
1x_{\texttt{adm1},c} & +1x_{\texttt{adm2},c} & \leq & 2 & (\texttt{space} \text{ für } c) \\
1x_{\texttt{4xstm1},c} & +1x_{\texttt{1xstm16},c} & \leq & 8x_{\texttt{adm1},c} + 16x_{\texttt{adm2},c} & (\texttt{slot} \text{ für } c) \\
1x_{\texttt{stm1cap},bc} & +16x_{\texttt{stm16cap},bc} & & & \\
+1x_{\texttt{stm1cap},cd} & +16x_{\texttt{stm16cap},cd} & \leq & 64x_{\texttt{adm1},c} + 128x_{\texttt{adm2},c} & (\texttt{swcap} \text{ für } c) \\
1x_{\texttt{stm1cap},bc} & +1x_{\texttt{stm1cap},cd} & \leq & 4x_{\texttt{4xstm1},c} & (\texttt{stm1} \text{ für } c) \\
1x_{\texttt{stm16cap},bc} & +1x_{\texttt{stm16cap},cd} & \leq & 1x_{\texttt{1xstm16},c} & (\texttt{stm16} \text{ für } c) \\
\ldots & \ldots & & \ldots & (\ldots \text{ für } d) \\
1x_{\texttt{stm1cap},ab} & +1x_{\texttt{stm16cap},ab} & \leq & 1 & (\texttt{fiber} \text{ für } ab) \\
1x_{\texttt{stm1cap},bc} & +1x_{\texttt{stm16cap},bc} & \leq & 1 & (\texttt{fiber} \text{ für } bc) \\
1x_{\texttt{stm1cap},cd} & +1x_{\texttt{stm16cap},cd} & \leq & 1 & (\texttt{fiber} \text{ für } cd) \\
1x_{\texttt{stm1cap},ad} & +1x_{\texttt{stm16cap},ad} & \leq & 1 & (\texttt{fiber} \text{ für } ad) \\
1x_{\texttt{stm1cap},bd} & +1x_{\texttt{stm16cap},bd} & \leq & 1 & (\texttt{fiber} \text{ für } bd) \\
63x_{\texttt{stm1cap},ab} & +1008x_{\texttt{stm16cap},ab} & = & C_{ab} & (\texttt{cap} \text{ für } ab) \\
63x_{\texttt{stm1cap},bc} & +1008x_{\texttt{stm16cap},bc} & = & C_{bc} & (\texttt{cap} \text{ für } bc) \\
63x_{\texttt{stm1cap},cd} & +1008x_{\texttt{stm16cap},cd} & = & C_{cd} & (\texttt{cap} \text{ für } cd) \\
63x_{\texttt{stm1cap},ad} & +1008x_{\texttt{stm16cap},ad} & = & C_{ad} & (\texttt{cap} \text{ für } ad) \\
63x_{\texttt{stm1cap},bd} & +1008x_{\texttt{stm16cap},bd} & = & C_{bd} & (\texttt{cap} \text{ für } bd)
\end{array}
$$

Abbildung 10: Modell der Hardware- und Netzkonfiguration für das Beispiel.

3.3 Routing

Das Routing von Kommunikationsbedarfen hängt stark von den Übertragungstechnologien, den verwendeten Routingprotokollen und den umzusetzenden Ausfallsicherheitsanforderungen ab. Trotz der verschiedenen Ausprägungen solcher Routings kann wiederum eine gemeinsame Problemstruktur identifiziert werden. Im Folgenden wird daher zunächst die mathematische Modellierung eines grundlegenden Routingproblems vorgestellt, das gemeinsame Eigenschaften zahlreicher Routingmechanismen abbildet. Darüber hinaus ist dieses Modell flexibel genug, um zusätzliche Anforderungen an das Routing erfassen zu können, siehe Abschnitt 3.3.3.

3.3.1 Ein grundlegendes Routingproblem

Eine Gemeinsamkeit praktisch aller Routingmechanismen besteht darin, dass die Daten über einen oder mehrere Wege von einer Quelle zu einem Ziel durch das Netzwerk übertragen werden. Ausgehend von dieser abstrakten Sichtweise lässt sich ein grundlegendes Routingproblem wie folgt formulieren:

> *Gegeben sei eine Reihe von Bedarfsanforderungen, die zu jeweils einer Quelle und einem Ziel die zu übertragende Bandbreite spezifizieren. Es ist eine Auswahl von Übertragungswegen mit Bandbreitenzuordnung zu finden, so dass einerseits alle Bedarfe erfüllt und andererseits die verfügbaren Kapazitäten nicht überschritten werden.*

Zur mathematischen Abbildung dieser Problemstellung ist zunächst eine formale Beschreibung der Bedarfsanforderungen notwendig. Ein Bedarf q ist definiert durch eine Quelle s_q, eine Senke t_q und eine Bandbreitenanforderung d_q als ganzzahliges Vielfaches einer Basiseinheit b_q, die gleichzeitig die kleinste Granularität für eine Aufteilung auf mehrere Übertragungswege angibt. Die benötigte Gesamtbandbreite eines Bedarfes q ist also $d_q \cdot b_q$ und darf in Vielfachen von b_q auf s_q-t_q-Wegen im Netzwerk übertragen werden.

Die Wahl der Basiseinheit b_q hängt von der betrachteten Technologie und ihren spezifischen Eigenschaften ab. Beispielsweise lassen sich in optischen Netzen nur ganze Lichtwege schalten. Der Lichtweg stellt daher die geeignete Basiseinheit dar, und jeder Bedarf q wird als Anzahl benötigter Lichtwege zwischen s_q und t_q ausgedrückt. In SDH-Netzen hingegen stehen unterschiedliche VC-N-Containergrößen zur Verfügung (vgl. Abbildung 4 auf Seite 9). Diese fungieren als individuelle Basiseinheiten einzelner Bedarfe. Hier kann es vorkommen, dass zwei Bedarfe mit identischer Quelle und Senke, aber unterschiedlichen Basiseinheiten existieren. Insgesamt wird durch die Aufteilung der Bedarfsmengen in Vielfache einer Basiseinheit eine Abstraktion erzielt, mit der das Routing in unterschiedlichsten Netzen in einem gemeinsamen Modell erfasst werden kann.

3.3.2 Mathematische Umsetzung

Für einen Bedarf q bezeichnet $\mathcal{P}^q$ die Menge aller Wege zwischen der Quelle s_q und der Senke t_q, die für das Routing dieses Bedarfes genutzt werden dürfen. Sofern keine einschränkenden Bedingungen vorliegen, sind das alle Wege im Netzwerk zwischen s_q und t_q. Zu jedem solchen Weg $P \in \mathcal{P}^q$ wird eine ganzzahlige Variable f_P^q eingeführt, deren Wert die Anzahl der Basiseinheiten b_q des Bedarfes q angibt, die über den Weg P geroutet werden.

$$
\begin{array}{llllll}
f_{P_1}^{ac} & +f_{P_2}^{ac} & +f_{P_3}^{ac} & +f_{P_4}^{ac} & = & 4 & \text{(Bedarf } a-c\text{)} \\
f_{P_5}^{bd} & +f_{P_6}^{bd} & +f_{P_7}^{bd} & & = & 4 & \text{(Bedarf } b-d\text{)} \\
f_{P_1}^{ac} & +f_{P_3}^{ac} & +f_{P_6}^{bd} & & \leq & C_{ab} & \text{(Link } ab\text{)} \\
f_{P_1}^{ac} & +f_{P_4}^{ac} & +f_{P_7}^{bd} & & \leq & C_{bc} & \text{(Link } bc\text{)} \\
f_{P_2}^{ac} & +f_{P_3}^{ac} & +f_{P_7}^{bd} & & \leq & C_{cd} & \text{(Link } cd\text{)} \\
f_{P_2}^{ac} & +f_{P_4}^{ac} & +f_{P_6}^{bd} & & \leq & C_{ad} & \text{(Link } ad\text{)} \\
f_{P_3}^{ac} & +f_{P_4}^{ac} & +f_{P_5}^{bd} & & \leq & C_{bd} & \text{(Link } bd\text{)}
\end{array}
$$

Abbildung 11: Routing-Modell für das Beispielproblem.

Damit lassen sich die Routingbedingungen durch zwei (Un-)Gleichungsklassen beschreiben. Zum einen wurde gefordert, „*eine Auswahl von Übertragungswegen mit Bandbreitenzuordnung zu finden, so dass alle Bedarfe erfüllt sind*“. Anders ausgedrückt muss sich für jeden Bedarf die Anzahl der über die einzelnen Wege gerouteten Basiseinheiten insgesamt zur geforderten Bedarfsgröße aufsummieren:

$$\sum_{P \in \mathcal{P}^q} f_P^q \;=\; d_q \quad \text{für jeden Bedarf } q. \tag{3}$$

Gleichzeitig ist sicherzustellen, dass „*die verfügbaren Kapazitäten nicht überschritten werden*“. Mit Ungleichung (2) auf Seite 19 ist für jeden Link ℓ die verfügbare Kapazität durch die Hilfsvariable C_ℓ repräsentiert. Zu jedem Bedarf q bezeichnet ferner $\mathcal{P}_\ell^q \subseteq \mathcal{P}^q$ die Menge derjenigen Wege aus $\mathcal{P}^q$, die den betrachteten Link ℓ enthalten. Über all diese Wege darf zusammen für alle Bedarfe nicht mehr als die sich auf dem Link ergebende Kapazität C_ℓ geroutet werden:

$$\sum_q \sum_{P \in \mathcal{P}_\ell^q} b_q \cdot f_P^q \qquad \text{für jeden Link } \ell. \tag{4}$$

Jede Belegung der Variablen f_P^q mit nicht-negativen, ganzzahligen Werten, die alle angegebenen Bedingungen (3) und (4) erfüllt, repräsentiert ein zulässiges Routing aller Bedarfe und stellt somit eine mögliche Lösung des formulierten Routingproblems dar. Zur Veranschaulichung dieser Routingmodellierung dient das folgende Beispiel.

Beispiel: Betrachten wir noch einmal das in Abbildung 9 auf Seite 18 dargestellte Netzwerk mit vier Knoten und fünf Links. Seien in diesem Netzwerk zwei Bedarfe, von a nach c sowie von b nach d, der Größe $d_q = 4$ zu routen. Für Verbindungen von a nach c stehen dabei die vier Wege $P_1 = a \to b \to c$, $P_2 = a \to d \to c$, $P_3 = a \to b \to d \to c$ sowie $P_4 = a \to d \to b \to c$ zur Verfügung. Zwischen b und d existieren lediglich drei Wege $P_5 = b \to d$, $P_6 = b \to a \to d$ und $P_7 = b \to c \to d$. Das Routingproblem wird durch das in Abbildung 11 zusammengestellte lineare Programm modelliert. Nehmen wir nun an, dass sich mit Ungleichung (4) die Linkkapazitäten $C_{ab} = C_{bc} = C_{cd} = 3$ sowie $C_{bd} = 2$ und $C_{ad} = 4$ ergeben. Dann ist ein mögliches Routing durch die Variablenbelegung $f_{P_1}^{ac} = f_{P_2}^{ac} = 2$, $f_{P_3}^{ac} = f_{P_4}^{ac} = 0$, $f_{P_5}^{bd} = 2$ und $f_{P_6}^{bd} = f_{P_7}^{bd} = 1$ beschrieben.

Dieses einfache wegbasierte Modell hat bei direkter Umsetzung in die Praxis ein „Größenproblem“: Je mehr Wege es im Netzwerk gibt, desto mehr Variablen enthält das Modell. Schon in kleinen Netzen kann die Variablenanzahl so groß sein, dass heutige Computer selbst mit den besten Methoden zur direkten

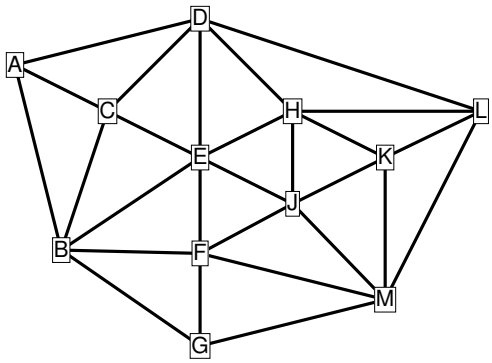


Abbildung 12: Netzwerk mit über 1400 Wegen zwischen A und M.

Behandlung eines solchen Modells wahrscheinlich keine Lösung in vertretbarer Zeit finden können. Die potentielle Anzahl unterschiedlicher Wege hängt dabei sehr stark von der topologischen Struktur des Netzes ab. In Ringen gibt es zwischen zwei Knoten genau zwei Wege, und folglich lassen sich Routings sehr einfach bestimmen. In vermaschten Netzen hingegen kann die Anzahl von Wegen explodieren.

Beispiel: In Abbildung 12 ist ein kleines Netzwerk mit 12 Knoten und 27 Links dargestellt. Zwischen den beiden Knoten A und M gibt es bereits über 1400 unterschiedliche Routingwege. Wenn zwischen je zwei Knoten ein Bedarf spezifiziert wird, müssen insgesamt 86 396 Wege berücksichtigt werden, und selbst bei Einschränkung der Lösungen auf nur einen Weg je Bedarf gibt es ca. 10^{205} verschiedene Routings (zum Vergleich: das Universum enthält „nur“ ca. 10^{80} Atome).

Dennoch kann dieses wegbasierte Modell erfolgreich für praktische Probleme eingesetzt werden. Auf geeignete Methoden zur Lösung solcher Modelle gehen wir in Kapitel 4 näher ein.

3.3.3 Zusätzliche Anforderungen

Das wegbasierte Modell ermöglicht auf einfache Weise, zusätzliche Anforderungen an das Routing zu formulieren. Im Folgenden wird dies anhand einiger Beispiele illustriert.

Weglängenbeschränkung: Aufgrund technischer Eigenschaften oder zusätzlicher Qualitätsanforderungen dürfen häufig nur Routingwege bis zu einer gewissen Maximallänge (in Kilometern oder Anzahl von Knoten) verwendet werden. Dies kann im Modell einfach berücksichtigt werden, indem $\mathcal{P}^q$ auf diejenigen Wege eingeschränkt wird, die dieser Längenbeschränkung genügen.

Schaltgranularität: Netzelemente erlauben oftmals nur das Schalten von Verbindungen ab einer bestimmten Containerkapazität. Infolgedessen können Bedarfe mit geringerer Basiskapazität in solchen Knoten nicht geschaltet werden. Bei gegebener Konfiguration der Netzelemente kann dies im Modell wiederum durch geeignete Einschränkung der Wegmengen $\mathcal{P}^q$ abgebildet werden.

Einzelweg-Routing: Eine spezielle Eigenschaft bestimmter Routingprotokolle besteht darin, dass jeder Bedarf nur einen einzigen Weg nutzt und nicht auf mehrere Wege verteilt geroutet werden darf. Diese Bedingung lässt sich wie folgt umsetzen: Da keine Bandbreitenaufteilung vorgenommen werden kann, wird für jeden Bedarf q die Basiseinheit auf die Gesamtbandbreite ($b_q := b_q \cdot d_q$)

gesetzt und die Vielfachheit entsprechend auf $d_q := 1$ normiert. Dadurch reduziert sich die Betrachtung auf binäre Variablen $f_P^q \in \{0, 1\}$, und Gleichung (3) vereinfacht sich zu

$$\sum_{P \in \mathcal{P}^q} f_P^q = 1 \quad \text{für jeden Bedarf } q \ .$$

Ausfallsicherheit: Eine Möglichkeit zur Erhöhung der Ausfallsicherheit im Netzwerk besteht darin, für jeden Bedarf q den maximalen Bedarfsanteil $0 < \delta_q \leq 1$ zu spezifizieren, der in einer Störsituation ausfallen darf. Dies kann leicht umgesetzt werden, indem der durch jeden Netzknoten (mit Ausnahme von s_q und t_q) fließende Anteil des Bedarfes durch

$$\sum_{P \in \mathcal{P}_u^q} f_P^q \quad \leq \quad \delta_q \cdot d_q \qquad \text{für jeden Knoten } u \text{ außer } s_q \text{ und } t_q$$

beschränkt wird, wobei $\mathcal{P}_u^q \subseteq \mathcal{P}^q$ analog zur Linknotation nur diejenigen Wege aus $\mathcal{P}^q$ umfasst, die u als inneren Knoten enthalten. Für einen eventuell vorhandenen direkten Link von s_q nach t_q muss analog eine zusätzliche Ungleichung formuliert werden. Durch diese Bedingungen wird sichergestellt, dass bei Ausfall von u (oder eines angrenzenden Links) mindestens der Bedarfsanteil $(1 - \delta_q) \cdot d_q \cdot b_q$ nicht betroffen ist. Mit dieser Methode kann zum Beispiel leicht 1+1-Protection realisiert werden, indem jeder Bedarfswert verdoppelt und $\delta_q = \frac{1}{2}$ gesetzt wird.

In ähnlicher Weise lassen sich zahlreiche weitere Bedingungen an das Routing modellieren und entsprechend berücksichtigen. Diese Erweiterungen übertragen sich dann auch auf das im Folgenden beschriebene Gesamtmodell zur integrierten Lösung von Netzplanungsproblemen.

3.4 Integrierte Optimierung

In den vorangegangenen Abschnitten sind abstrakte Modellierungen für zwei wesentliche Aspekte der Netzplanung vorgestellt worden, mit denen sich die resultierenden Teilprobleme weitestgehend unabhängig von der konkreten Technologie des zu planenden Netzes mathematisch abbilden lassen. Auf dieser Basis beschreiben wir nun, wie die im Einzelnen vorgestellten Bestandteile im Rahmen eines Gesamtmodells zusammenwirken. Abschließend erläutern wir Zielfunktionen, mit denen ein übergeordnetes Optimierungsziel abgebildet wird.

3.4.1 Zusammenstellung des Gesamtmodells

Das integrierte Gesamtmodell besteht aus den beiden Teilmodellen für die Hardware-Konfiguration und das Routing, die über Linkkapazitäten miteinander gekoppelt werden. Diesen Zusammenhang haben wir bereits in Abbildung 7 auf Seite 15 schematisiert. Für die Hardware-Konfiguration des Netzes wird das Komponenten-Ressourcen-Modell aus Abschnitt 3.2 eingesetzt, durch dessen Abstraktion alle notwendigen Randbedingungen für eine spezielle Technologie korrekt erfasst werden können, gleich um welche Technologie es sich auch handelt. Das Routingmodell aus Abschnitt 3.3 bietet die Möglichkeit, für verschiedene zugrunde liegende Routingprotokolle alle denkbaren Kommunikationsverbindungen zu erfassen. Fasst man beide Teilmodelle zusammen und legt die Parameter (Komponenten, Ressourcen, Bedarfe usw.) entsprechend der jeweiligen Technologie und den Planungsvorgaben fest, so entsteht ein konsistentes Gesamtmodell für die komplette Planungsaufgabe.

Den Kern dieser Integration bilden bei den hier vorgestellten Planungsaufgaben stets die Linkkapazitäten. Um ein mögliches Routing realisieren zu können, muss die Hardware die nötigen Kapazitäten bereitstellen. Das lässt sich sinnbildlich durch die folgenden Beziehungen veranschaulichen:

$$\begin{aligned} \text{Routing} \quad &\leq \quad \text{Linkkapazität}, \\ &\qquad\quad \text{Linkkapazität} \quad \leq \quad \text{Hardware-Konfiguration}. \end{aligned}$$

Zur Koppelung der Teilmodelle stehen die Hilfsvariablen C_ℓ zur Verfügung. Die Ungleichung (4) auf Seite 22 sagt aus, dass die den Bedarfen zugeordnete Bandbreite die Kapazität eines Links nicht übersteigen darf, während die Gleichung (2) auf Seite 19 die bereitgestellte Kapazität spezifiziert. Ersetzen wir die Hilfsvariable C_ℓ in (4) durch die Gleichung (2), so entspricht folgende Ungleichung der vorstehenden sinnbildlichen Beziehung:

$$\sum_q \sum_{P \in \mathcal{P}_\ell^q} b_q \cdot f_P^q \;\leq\; C_\ell = \sum_{k \in K_\ell} A(k, \texttt{cap}) \cdot x_{k,\ell} \;.$$

Die Aufgabe der Optimierung besteht nun darin, aus allen zulässigen Lösungen des Gesamtmodells und somit aus allen verfügbaren Hardware-Konfigurationen und Routings solche zu identifizieren, die bezüglich eines übergreifenden Optimierungszieles bestmöglich sind. Dieser Aspekt wird durch eine Zielfunktion abgebildet.

3.4.2 Zielfunktion

Eine Zielfunktion ordnet jeder zulässigen Lösung eine reelle Zahl zu, mit der die Lösung quantitativ bewertet wird. Als Optimierungsproblem bezeichnet man dann die Aufgabe, aus allen im Rahmen des Modells zulässigen Lösungen solche zu ermitteln, die bezüglich der Zielfunktion minimial (z. B. bei Kosten) oder maximal (z. B. für Gewinn) sind.

Bei der Beschreibung der Planungsaufgaben in Kapitel 2 wurden auch die jeweiligen Ziele formuliert, bezüglich derer das zu bestimmende Netz optimiert werden soll. Wie bei mehreren der dort vorgestellten Aufgaben sind häufig die Hardwarekosten zu minimieren. Es können aber auch andere quantifizierbare Netzeigenschaften wie etwa die höchste auftretende Linkauslastung beim Traffic Engineering als Zielfunktion modelliert werden.

Kostenminimierung: Bei der Kostenminimierung von Netzwerken drückt die Zielfunktion die notwendigen Investitionen aus. In Abschnitt 3.2 wurden die Variablen $x_{k,g}$ eingeführt, die angeben, wie oft eine Komponente k an einem Knoten oder Link g zu installieren ist. Bezeichnet man mit $c_{k,g}$ die Kosten einer solchen Installation pro Komponente, so drückt die Zielfunktion

$$\min \sum_{\text{Links und Knoten } g} \sum_{k \in K_g} c_{k,g} \cdot x_{k,g}$$

aus, dass die Gesamtkosten des zu planenden Netzes minimiert werden sollen. In Abbildung 13 ist eine mögliche Zielfunktion für das Beispiel in Abbildung 9 auf Seite 18 dargestellt.

$$
\begin{array}{llll}
\min & 20x_{\texttt{adm1},a} & +20x_{\texttt{adm1},b} & +20x_{\texttt{adm1},c} & +20x_{\texttt{adm1},d} \\
 & +30x_{\texttt{adm2},a} & +30x_{\texttt{adm2},b} & +30x_{\texttt{adm2},c} & +30x_{\texttt{adm2},d} \\
 & +4x_{\texttt{4xstm1},a} & +4x_{\texttt{4xstm1},b} & +4x_{\texttt{4xstm1},c} & +4x_{\texttt{4xstm1},d} \\
 & +5x_{\texttt{1xstm16},a} & +5x_{\texttt{1xstm16},b} & +5x_{\texttt{1xstm16},c} & +5x_{\texttt{1xstm16},d} \\
 & +8x_{\texttt{stm1cap},ab} & +8x_{\texttt{stm1cap},bc} & +8x_{\texttt{stm1cap},cd} & +8x_{\texttt{stm1cap},ad} \\
 & +9x_{\texttt{stm1cap},bd} & +32x_{\texttt{stm16cap},ab} & +32x_{\texttt{stm16cap},bc} & +32x_{\texttt{stm16cap},cd} \\
 & +32x_{\texttt{stm16cap},ad} & +36x_{\texttt{stm16cap},bd} & &
\end{array}
$$

Abbildung 13: Zielfunktion für das Beispiel der Hardware- und Netzkonfigurationen.

Auslastungsminimierung: Das zweite Beispiel befasst sich mit der Minimierung der Linklasten wie im Traffic Engineering aus Abschnitt 2.4. Als Linklast wird dabei das Verhältnis von genutzter zu vorhandener Kapazität verstanden. Für den vereinfachenden Fall, dass für alle Links ℓ die Kapazität C_ℓ ein fester Problemparameter und damit keine Variable ist, wird die Auslastung A_ℓ des Links ℓ durch

$$A_\ell = \frac{\text{genutzte Kapazität auf Link } \ell}{\text{Kapazität auf Link } \ell} = \frac{\sum_q \sum_{P \in \mathcal{P}_\ell^q} b_q \cdot f_P^q}{C_\ell}$$

gegeben. Um die maximale Linkauslastung im Netzwerk zu ermitteln, wird eine vom Link unabhängige Auslastung A eingeführt. Da diese mindestens so groß sein soll wie die Auslastung auf jedem einzelnen Link, gilt

$$A \geq A_\ell = \frac{\sum_q \sum_{P \in \mathcal{P}_\ell^q} b_q \cdot f_P^q}{C_\ell} \quad \text{für jeden Link } \ell\,. \tag{5}$$

Fügt man dem Modell die Ungleichungen (5) für alle Links ℓ sowie die neue Hilfsvariable A hinzu, so repräsentiert die Zielfunktion

$$\min\ A$$

das Optimierungsziel, die maximale relative Linkauslastung zu minimieren.

Durch das Hinzufügen der jeweiligen Zielfunktion wird das Gesamtmodell zu einer Planungsaufgabe vervollständigt. Für das durchgängig verwendete Beispiel dieses Kapitels besteht das Gesamtmodell aus der Zusammenfassung der Bilder 10, 11 und 13.

4 Algorithmen: Welche Methoden nutzen wir?

Mit einem mathematischen Gesamtmodell sind alle Randbedingungen einer Planungsaufgabe präzise formuliert, es ist ein eindeutiges Ziel definiert, und jede zulässige Lösung wird durch Auswertung der Zielfunktion quantitativ bewertet. Damit ist für je zwei zulässige Lösungen klar, welche die bessere ist, und es gibt insbesondere die eindeutige Teilmenge der *optimalen* Lösungen.

In den meisten kommerziellen Planungstools werden heuristische Verfahren verwendet. Diese zielen mit unterschiedlichsten Strategien auf das Berechnen von Lösungen ab. Zu den prominentesten Vertretern

heuristischer Verfahren gehören randomisierte Metaheuristiken wie Simulated Annealing, Genetische Algorithmen oder Tabu Search. Solche Heuristiken liefern in der Praxis zwar häufig zulässige Lösungen, es ist allerdings völlig unklar, wie gut eine Lösung ist bzw. wie weit diese von einer optimalen Lösung entfernt ist. Auf die Kostenoptimierung angewendet heißt dies Folgendes: Hat die Heuristik ein Netz mit Kosten 20 Mio. Euro berechnet, so bleibt vollkommen unklar, ob es zulässige Netze für 15 Mio. Euro oder gar 10 Mio. Euro gibt. Für den Netzbetreiber bleibt nur die Hoffnung, dass die beste gefundene Lösung nicht unnötig teuer ist.

Genau an diesem Schwachpunkt setzen moderne mathematische Optimierungsverfahren an. Über die letzten Jahrzehnte hinweg wurden praxistaugliche Lösungsverfahren entwickelt, die einerseits den gesamten Lösungsraum betrachten und damit sicher mit einer optimalen Lösung terminieren können und die andererseits zu jedem Zeitpunkt Rechenschaft darüber ablegen, wie weit die gegenwärtig beste Lösung maximal von einer optimalen Lösung entfernt ist. Erst dadurch ist eine sinnvolle Bewertung jeder Lösung und letztlich auch der Beweis von Optimalität möglich.

Dieses Kapitel führt schrittweise an solche Optimierungsverfahren heran. Ohne auf die Details der Verfahren einzugehen, werden die allgemeinen Prinzipien dargestellt. Ausgehend von dem für sich genommen „dummen“ Verfahren der vollständigen Enumeration werden nach und nach komplexer werdende Methoden präsentiert, die schließlich in einem Branch&Cut&Price-Verfahren münden.

4.1 Vollständige Enumeration

Bei der vollständigen Enumeration werden alle Lösungen eines Problems erzeugt und miteinander verglichen, um eine optimale Lösung zu finden. Ein allgemeiner und strukturierter Ansatz zur Bewältigung dieser Aufgabe ist die Divide&Conquer-Methode. Durch sukzessives Festlegen von Variablen werden dabei alle Lösungen erzeugt. Zur Illustration der getroffenen Entscheidungen dient ein Suchbaum, siehe Abbildung 14.

Jeder Knoten des Suchbaumes repräsentiert eine bestimmte Festlegung eines Teiles der Variablen. Am Startknoten sind alle Variablen frei, d. h., ihnen ist noch kein Wert zugewiesen. An jedem anderen Knoten wird eine noch freie Variable ausgewählt. Für jeden erlaubten Wert dieser Variablen wird dann ein neuer Kinderknoten erzeugt, der die Variablenfestlegungen des Elternknotens übernimmt und zusätzlich der ausgewählten Variable einen bestimmten Wert zuweist. An den Blättern des Baumes, also denjenigen Knoten, an denen alle Variablen zugewiesene Werte haben, sind keine weiteren Verzweigungen mehr möglich. Jedes Blatt repräsentiert damit eine potentielle Lösung des Optimierungsproblems, und umgekehrt gibt es für jede mögliche Lösung auch ein Blatt. Durch einfaches Überprüfen, ob die getroffenen Festlegungen der Variablen eine zulässige Lösung ergeben, lässt sich somit eine optimale Lösung finden.

Es gibt verschiedene Algorithmen zum gezielten Aufbauen und Durchsuchen („Traversieren“) des Baumes, wie beispielsweise die Tiefensuche. Der folgende Algorithmus erzeugt sukzessive alle zulässigen Lösungen, berechnet deren Zielfunktionswerte und liefert damit im Ergebnis eine garantiert optimale Lösung.

wiederhole

1. Wähle einen noch nicht betrachteten Knoten.
2. Wenn alle Variablen festgelegt sind, ist eine potentielle Lösung gefunden. Wenn sie zulässig und besser als alle bisher bekannten ist, speichere sie ab.

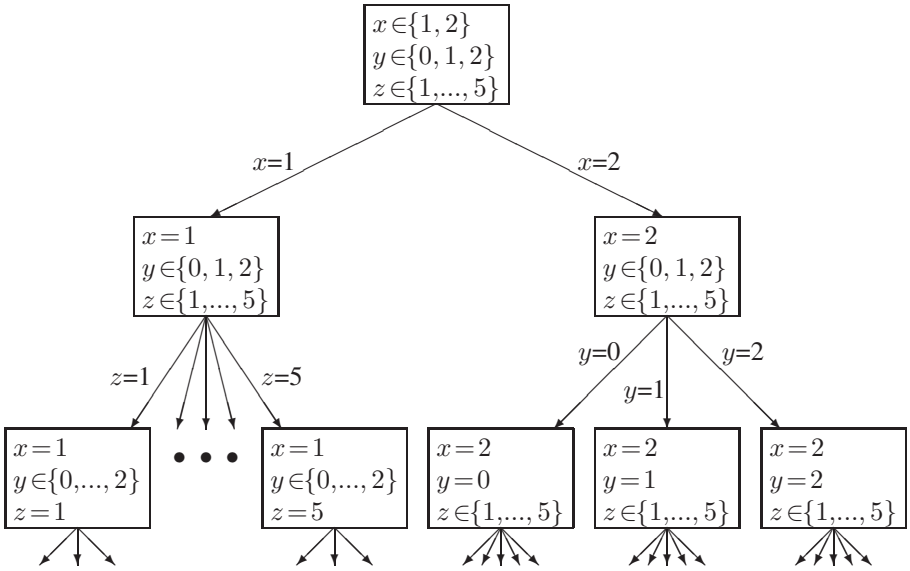


Abbildung 14: Ein Suchbaum.

3. Wenn noch nicht alle Variablen festgelegt sind, verzweige auf einer freien Variablen und erzeuge die zugehörigen Kinderknoten.

bis alle Knoten betrachtet sind.

Oft kann der Suchbaum bereits durch relativ einfache Techniken stark verkleinert und das Absuchen damit beschleunigt werden. Stellt man an einem der Knoten fest, dass es für die bisher getroffenen Variablenfestlegungen überhaupt keine zulässige Lösung geben kann, so müssen für die übrigen freien Variablen keine Festlegungen getroffen und Kinderknoten erzeugt und abgesucht zu werden – der gesamte Teilbaum unterhalb des aktuellen Knotens kann ignoriert werden. Man spricht dabei vom „Abschneiden“ oder „Pruning“ des Teilbaumes.

Beispiel: In Abschnitt 3.2 wurde ein Planungsbeispiel vorgestellt, bei dem Variablen für die Platzierung von Komponenten verwendet wurden. Dabei standen zwei ADMs zur Verfügung. Die Installation eines ADMs an Standort a entspricht dem Setzen der Variablen $x_{\texttt{adm1},a}$ bzw. $x_{\texttt{adm2},a}$ auf 1. Der Aufbau des Suchbaumes und das Pruning können anhand dieser Variablen verdeutlicht werden. Wenn die Kinder eines Knotens für die Variable $x_{\texttt{adm1},a}$ erzeugt werden, also ein Knoten mit $x_{\texttt{adm1},a} = 0$ und einer mit $x_{\texttt{adm1},a} = 1$, so übernehmen alle darunter liegenden Kinderknoten diese Festlegung. Alle Lösungen, bei denen der ADM `adm1` installiert wird, sind Kinder des Knotens mit $x_{\texttt{adm1},a} = 1$, und alle Lösungen, bei denen `adm1` nicht installiert wird, Kinder des Knotens mit $x_{\texttt{adm1},a} = 0$. Laut Planungsproblem ist es nicht erlaubt, die ADMs `adm1` und `adm2` gleichzeitig am selben Ort zu installieren. Wenn die Variablenfestlegung an einem Knoten also sowohl $x_{\texttt{adm1},a} = 1$ als auch $x_{\texttt{adm2},a} = 1$ fordert, so kann es an keinem Blatt unterhalb dieses Knotens eine zulässige Lösung geben. Der gesamte Teilbaum kann daher ignoriert werden.

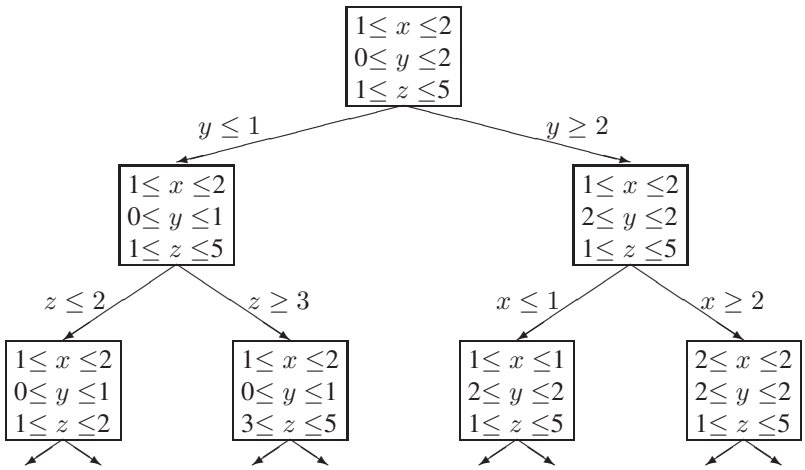


Abbildung 15: Ein Branch&Bound-Baum.

4.2 Branch&Bound

Die Branch&Bound-Methode basiert auf der vollständigen Enumeration, erweitert diese aber um einige wesentliche Aspekte.

Die erste Erweiterung betrifft das Pruning. Anders als bei der vollständigen Enumeration, bei der Teilbäume nur dann ignoriert werden, wenn sie keine zulässige Lösung enthalten, werden bei der Branch&-Bound-Methode auch solche Teilbäume ignoriert, die zwar noch zulässige, aber definitiv keine besseren Lösungen als die beste bereits gefundene enthalten. Für Kostenminimierungsprobleme wird dazu an jedem Knoten eine untere Kostenschranke bestimmt, eine „lower bound“. Das ist ein Kostenwert, den keine der im Teilbaum enthaltenen Lösungen unterschreiten kann. Wenn dieser Wert die Kosten einer bereits bekannten Lösung übersteigt, so kann der gesamte Teilbaum ignoriert werden. Diese Art der Verwendung von Schranken wird als „Bounding“ bezeichnet.

Die zweite Erweiterung der Branch&Bound-Methode gegenüber der vollständigen Enumeration betrifft die Erzeugung von Kinderknoten. Wurde bisher an jedem Knoten für jeden möglichen Wert einer freien Variable in einen Kinderknoten verzweigt, so erzeugt Branch&Bound an jedem Knoten stets nur zwei Kinder. Dabei wird der Wert der jeweiligen Variablen nicht vollständig festgelegt, sondern nur eingeschränkt. Diese veränderte Art der Erzeugung von Kinderknoten wird als „Branching“ bezeichnet und ist in Abbildung 15 veranschaulicht. Als praktischer Vorteil gegenüber der vollständigen Enumeration ergibt sich daraus, dass nicht mehr alle potentiellen Lösungen einzeln erzeugt und bewertet werden müssen, sondern alle Werte größer bzw. alle Werte kleiner der zum Branching gewählten Schranke gemeinsam bewertet werden. Oft lassen sich bereits diese eingeschränkten Teilprobleme durch Pruning bei Unzulässigkeit oder durch gute untere Kostenschranken bei der Suche nach der Optimallösung ausschließen, was jeweils zu einer Reduzierung der Suchbaumgröße führt.

Beispiel: Bei der vollständigen Enumeration werden zu einer Variablen x, die Werte zwischen 0 und 63 annehmen darf, 64 Kinder erzeugt. Wird durch diese Variable beispielsweise die Kapazität eines Links modelliert, und sind die vorhandenen Bedarfsanforderungen nur erfüllbar, wenn die installierte Kapazität mindestens 50 beträgt, so enthalten 50 der 64 erzeugten Teilbäume überhaupt keine zulässigen Lösungen, müssen aber dennoch aufwändig abgesucht werden. Bei der Branch&Bound-Methode könnte ein Knoten

für $0 \leq x \leq 31$ und einer für $32 \leq x \leq 63$ erzeugt werden. Erkennt man bereits hier die Unzulässigkeit des Knotens für $x \leq 31$, so kann dieser Teilbaum mit einem Schritt abgeschnitten werden.
Eine einfache Möglichkeit, eine untere Kostenschranke für den anderen Knoten zu berechnen, ergibt sich aus der Betrachtung der Variablenschranken. Für den Knoten mit $x \geq 32$ fallen mindestens die Kosten von 32 Einheiten der Kapazitätskomponente an. Wird im nächsten Schritt des Algorithmus dieser Knoten betrachtet und das Branching wieder auf der Variablen x durchgeführt, so erhält man die Knoten für $32 \leq x \leq 47$ und für $48 \leq x \leq 63$. Der Knoten $x \leq 47$ kann wieder wegen Unzulässigkeit ignoriert werden, für den Knoten $x \geq 48$ fallen mindestens die Kosten von 48 Einheiten der Kapazitätskomponente an.

Die Bestimmung guter unterer Kostenschranken ist für die Effizienz der Branch&Bound-Methode von zentraler Bedeutung. Das oben beschriebene einfache Auswerten der Variablenschranken ist in der Regel nicht ausreichend. Ein anderer Ansatz hat sich zum Lösen praktischer Probleme bewährt: anstelle des tatsächlichen Teilproblems löst man eine sogenannte „Relaxierung“. Eine Relaxierung ist eine Vereinfachung des Originalproblems, die diesem strukturell sehr ähnlich ist und insbesondere alle zulässigen Lösungen enthält. Daher sind alle Kostenschranken, die man für eine Relaxierung herleiten kann, auch Kostenschranken für das Originalproblem. Durch geeignete und leicht lösbare Relaxierungen lassen sich oft sehr gute untere Kostenschranken berechnen. Ist ferner die Optimallösung einer Relaxierung auch für das Originalproblem zulässig, so ist der entsprechende Knoten des Suchbaumes direkt optimal gelöst, und man braucht nicht weiter zu verzweigen. Dieses direkte Lösen der zu den Knoten des Suchbaumes gehörenden Teilprobleme ist die dritte wichtige Erweiterung der Branch&Bound-Methode.

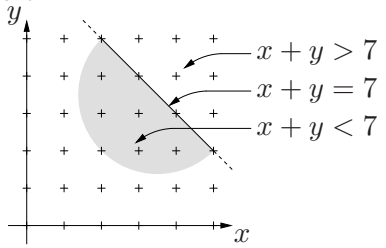
4.3 Branch&Bound für ganzzahlige lineare Programme

In Kapitel 3 wurden Optimierungsprobleme unter Verwendung von ganzzahligen Variablen und linearen Ungleichungen als ganzzahlige lineare Programme modelliert. Die Forderung nach Ganzzahligkeit hat zur Folge, dass diese Modelle im Allgemeinen nicht einfach und schnell lösbar sind. Lässt man bei einem ganzzahligen linearen Programm jedoch zu, dass die Variablen auch gebrochene Werte annehmen können, so erhält man die sogenannte „lineare Relaxierung“ des ganzzahligen linearen Programms: ein verwandtes, aber sehr viel einfacheres Optimierungsproblem, zu dessen Lösung mit der Simplex-Methode [7] und dem Innere-Punkte-Verfahren [24] praktisch effiziente Algorithmen existieren, die auch als (kommerzielle) Software [13, 9, 28] zur Verfügung stehen. Abbildung 16 stellt den Zusammenhang zwischen einem ganzzahligen linearen Programm und seiner linearen Relaxierung an einem Beispiel mit zwei Variablen dar. Da die Relaxierung auch alle Lösungen des Originalproblems enthält, kann es keine ganzzahlige Lösung geben, die günstiger als die Optimallösung der Relaxierung ist. Letztere liefert somit eine untere Kostenschranke für das ganzzahlige lineare Programm.

Beim Lösen ganzzahliger linearer Programme mittels Branch&Bound entspricht jeder Knoten des Suchbaumes einem ganzzahligen linearen Programm, das aus den Originalungleichungen und den Variablenschranken der bisherigen Einschränkungen besteht. Die zugehörige lineare Relaxierung ist schnell lösbar und liefert eine gute untere Kostenschranke. Diese Schranke ist üblicherweise sehr viel besser als die in Abschnitt 4.2 verwendete Schranke, die sich unmittelbar aus den Einschränkungen der Variablen ergibt. Mit diesem verbesserten Bounding wird meist die Größe des Suchbaumes nochmals deutlich reduziert.

Wie schon erläutert, verkleinert sich der Suchbaum noch weiter, wenn an einem seiner Knoten die Optimallösung der linearen Relaxierung allen Variablen nur ganzzahlige Werte zuweist, da der Teilbaum in diesem Fall direkt optimal gelöst ist. Aber auch wenn nicht alle Variablen ganzzahlige Werte annehmen, liefert die Optimallösung der linearen Relaxierung noch einige wertvolle Informationen zur Steuerung

Die Ungleichung $x + y \leq 7$ unterteilt den Variablenraum in drei Bereiche: die Gerade, auf der $x + y = 7$ gilt, sowie die Bereiche darüber und darunter. Die Lösungen der Ungleichung befinden sich alle auf und unter der Geraden.



Die Lösungen der linearen Relaxierung erfüllen alle Ungleichungen auf einmal, liegen also bei jeder der Geraden auf der „richtigen“ Seite.

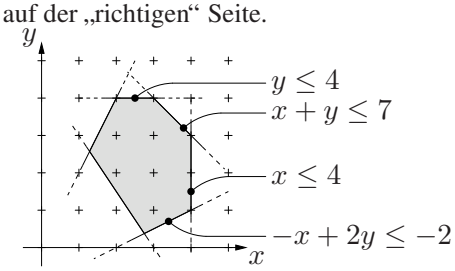


Abbildung 16: Geometrische Darstellung: Ungleichungen und lineare Relaxierung.

und Verbesserung des Branch&Bound-Verfahrens.

Häufig liegt die gebrochene Optimallösung der linearen Relaxierung sehr nah an einer ganzzahligen Lösung. Durch einfaches Runden kann dann aus der gebrochenen Lösung eine ganzzahlige Variablenbelegung gewonnen werden. Erfüllt diese gerundete Belegung auch alle Ungleichungen, so ist zumindest eine zulässige Lösung des ganzzahligen Programms gefunden. Auf diese Art erhält man normalerweise sehr schnell gute ganzzahlige Lösungen, die frühzeitig zum Bounding und Pruning herangezogen werden können.

Darüber hinaus kann die gebrochene Optimallösung auch dazu verwendet werden, den Branch&Bound-Algorithmus zu steuern. An jedem Knoten muss eine Variable für das Branching sowie Beschränkungen dieser Variable für die neuen Kinderknoten bestimmt werden. Es ist sinnvoll, diese so zu wählen, dass die gebrochenen Optimallösungen an den Kinderknoten besser werden, also näher an ganzzahligen Lösungen liegen. Typischerweise wählt man daher eine Variable mit einem gebrochenen Wert und legt für einen Kinderknoten den nächstgrößeren ganzzahligen Wert als untere Schranke für diese Variable und für den anderen Kinderknoten den nächstkleineren ganzzahligen Wert als obere Schranke fest.

Beispiel: Abbildung 17 illustriert das Branching auf einer Variablen, wobei die Kosten von Variable y gleich 1 und die Kosten von Variable x gleich 0 sind. Lösungen werden also mit kleiner werdendem y-Wert besser. Für die Relaxierung ist der Punkt bei x=2.75, y=0.375 optimal. Wenn mit $y \leq 2$ bzw. $y \geq 3$ verzweigt würde, so ist wenig gewonnen: Der eine Kinderknoten enthält nur sehr teure Lösungen, und die Optimallösung der Relaxierung des anderen Kinderknotens ist dieselbe Lösung wie zuvor. Wie erläutert, ist es also besser, mit $x \leq 2$ bzw. $x \geq 3$ zu verzweigen, da dann keiner der beiden Kinderknoten die bisherige gebrochene Lösung mit x=2.75 enthält.

4.4 Branch&Cut

Die Performance des gesamten Branch&Bound-Algorithmus hängt stark von der Qualität der Relaxierung ab, da deren Optimallösung sowohl untere Schranken für das Bounding liefert als auch für die Steuerung des Algorithmus verwendet wird. Im Idealfall ist die Optimallösung der linearen Relaxierung auch eine Lösung (und somit auch eine optimale Lösung) des ursprünglichen gemischt-ganzzahligen

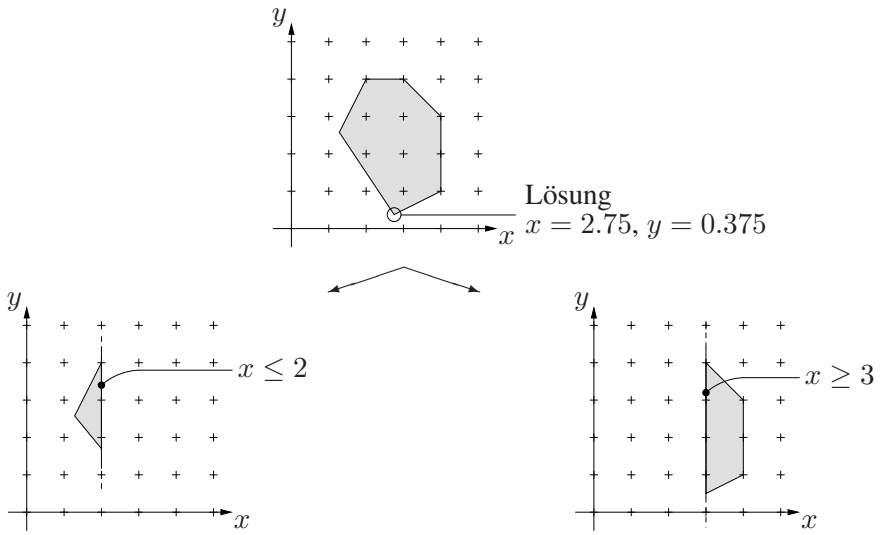


Abbildung 17: Branching auf einer Variablen.

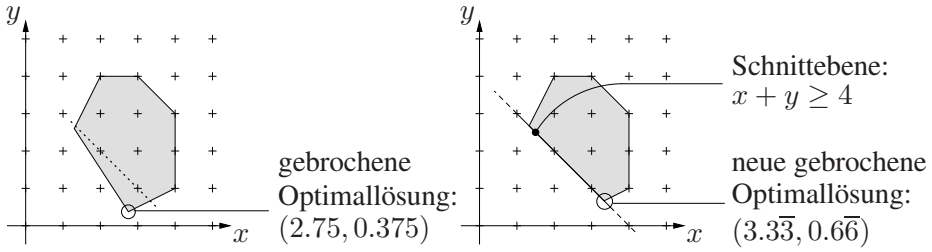


Abbildung 18: Separieren einer gebrochenen Lösung von der ganzzahligen Lösungsmenge.

Programms. In diesem Fall ist kein Branching mehr notwendig.

Um dem möglichst nahe zu kommen, wird im Branch&Cut-Algorithmus eine weitere Technik verwendet, das Separieren mit Hilfe von „Schnittebenen“. Es wurde im vorigen Abschnitt bereits erläutert, dass eine Ungleichung den Variablenraum so „zerschneidet“, dass alle zulässigen Punkte auf einer Seite liegen. Hat nun die gegenwärtige lineare Relaxierung eine gebrochene Optimallösung, so gibt es eine Ungleichung, die alle zulässigen ganzzahligen Punkte auf der einen und die gebrochene Lösung auf der anderen Seite hat. Diese sogenannte „Schnittebene“ (in höheren Dimensionen entsprechen Gleichungen nicht mehr Geraden, sondern (Hyper-)Ebenen) separiert die gebrochene Lösung von den gesuchten ganzzahligen Lösungen.

Durch solche zusätzlichen Ungleichungen lässt sich die lineare Relaxierung meist deutlich verbessern. Fügt man eine Schnittebenen-Ungleichung zu der linearen Relaxierung hinzu und löst diese erneut, so erhält man eine “bessere” Lösung, die näher an den zulässigen ganzzahligen Lösungspunkten liegt. Das Hinzufügen einer solchen Ungleichung wird in Abbildung 18 dargestellt.

Theoretisch ist es möglich, gemischt-ganzzahlige Programme ausschließlich durch das Hinzufügen solcher Ungleichungen zu lösen, also ohne Branching. Dies ist jedoch nicht praktikabel, da es kein univer-

selles und effizientes Verfahren gibt, um alle dazu notwendigen Schnittebenen zu identifizieren. Daher kommen beim Branch&Cut-Verfahren primär spezielle, problemspezifische Schnittebenen zum Einsatz, die man auch algorithmisch schnell finden kann.

Beispiel: Bei Netzplanungsproblemen kommen häufig Partitionsungleichungen zum Einsatz. Das Wort „Partition“ bezieht sich hierbei auf eine Menge von Links, die das Netzwerk in mehrere Teile zerlegen (partitionieren). Im Netzwerk aus Abbildung 9 auf Seite 18 zerteilen die Links ab, bd und cd das Netz in ein Teilnetz mit den Knoten b und c sowie einem zweiten Teilnetz mit den Knoten a und d. Besteht Kommunikationsbedarf zwischen diesen beiden Teilnetzen, so muss offensichtlich auf mindestens einem der Links ab, bd oder cd Kapazität installiert werden. Gibt es beispielsweise einen Gesamtbedarf von 7×STM-1 zwischen a und dem Teilnetz aus b und c, so kann es keine Lösung geben, die diesen Links nur eine Kapazität von STM-1 zuordnet. Eine Lösung kann nur zulässig sein, wenn auf mindestens einem der Links die höhere verfügbare Kapazität, STM-16, installiert wird. Die Ungleichung

$$x_{\texttt{stm16cap},ab} + x_{\texttt{stm16cap},bd} + x_{\texttt{stm16cap},cd} \geq 1$$

drückt genau diesen Sachverhalt aus.

Selbst für relativ kleine gemischt-ganzzahlige Programme kann es extrem viele zu berücksichtigende Schnittebenen geben. Würde man sie alle zu Beginn identifizieren und zur linearen Relaxierung hinzufügen, so würde dies das Verfahren allein wegen der Größe des zu lösenden linearen Programms erheblich verlangsamen. Um dies zu vermeiden, wird beim Branch&Cut-Verfahren das unten beschriebene Schnittebenenverfahren verwendet. Dabei werden die Ungleichungen erst „bei Bedarf“ erzeugt, und das lineare Programm bleibt klein und somit leichter lösbar.

wiederhole

1. Löse die lineare Relaxierung.
2. Suche eine gültige Ungleichung, die von der optimalen Lösung der Relaxierung verletzt wird.
3. Füge diese der Relaxierung hinzu.

bis keine verletzte Ungleichung mehr gefunden wird.

4.5 Branch&Cut&Price

Die Geschwindigkeit eines Algorithmus zur Lösung linearer Programme wird nicht nur durch die Anzahl von Ungleichungen beeinflusst, sondern auch durch die Anzahl von Variablen. In Abschnitt 3.3.2 wurde bereits erwähnt, dass beispielsweise eine wegbasierte Modellierung des Routingproblems eine enorme Menge von Variablen enthalten kann. Es ist jedoch nicht nötig, alle Variablen explizit zu betrachten. Ähnlich wie die Ungleichungen im Schnittebenenverfahren können auch Variablen bei der sogenannten „Spaltengenerierung“ („Pricing“) erst bei Bedarf erzeugt werden.

Bei diesem Ansatz wird initial nur eine vergleichsweise kleine Anzahl der Variablen betrachtet und für diese eine Lösung der linearen Relaxierung berechnet. Alle anderen Variablen werden implizit auf 0 gesetzt. Der Algorithmus zum Lösen des linearen Programms liefert neben der Lösung noch Informationen (u. a. eine Duallösung), mittels derer bestimmt werden kann, ob eine der bisher nicht berücksichtigten Variablen zu einer besseren Lösung führen kann. Wird solch eine Variable gefunden, so wird diese der

Relaxierung hinzugefügt und eine Reoptimierung vorgenommen. Da ein Ungleichungssystem normalerweise in Form einer Matrix abgespeichert wird und das Hinzufügen einer Variable eine neue Spalte in dieser Matrix erzeugt, spricht man von Spaltengenerierung. Wird das folgende Pricing-Verfahren in ein Branch&Cut-Verfahren integriert, so spricht man schließlich von Branch&Cut&Price.

Bestimme eine initiale Variablenmenge.

wiederhole

1. Löse die lineare Relaxierung.
2. Bestimme eine Variable, die die optimale Lösung der Relaxierung verbessern kann.
3. Füge diese der Relaxierung hinzu.

bis keine Verbesserungsvariable mehr gefunden wird.

Die implizite Betrachtung aller Wege bei einem Routingsproblem kann also folgendermaßen aussehen: Zuerst wird für jeden Bedarf ein einzelner, vielversprechend erscheinender Weg gewählt und die zugehörige Variable zur anfänglichen Variablenmenge hinzugefügt. Dabei kann es sich beispielsweise um den kürzesten Routingweg handeln. Mit diesen Variablen wird die lineare Relaxierung gelöst. Da nur ein Weg je Bedarf zur Verfügung steht, ist diese Relaxierung sehr einfach lösbar. Danach wird für jeden Bedarf ein weiterer Weg gesucht, dessen Variable die Relaxierung verbessern kann, und die Relaxierung wird erneut gelöst. Diese Prozedur wird so lange wiederholt, bis keine verbessernden Wege mehr gefunden werden. Auf diese Art werden nur sehr wenige der Variablen tatsächlich erzeugt, gleichzeitig wird aber garantiert, dass die gefundene Lösung bezüglich aller Variablen optimal ist.

Das Branch&Cut&Price-Verfahren stellt einen hochentwickelten Ansatz zur integrierten Bearbeitung komplexer Probleme dar. Neben der hier vorgestellten Methodenkaskade gibt es jedoch noch zahlreiche weitere Ansätze und Adaptionen, wie etwa dynamische Programmierung sowie verschiedene Relaxierungs- und Dekompositionstechniken. Die folgenden Referenzen sind grundlegende, einführende Werke in das Gebiet der mathematischen Optimierung. Für einen Einstieg in die Modellierung ganzzahliger linearer Programme eignet sich das Buch [3] sehr gut. Zur Lösung linearer Programme unterscheiden wir zwischen dem Simplex-Verfahren [7] und Inneren-Punkt-Methoden [24]. Eine ausführliche Beschreibung der Techniken für die Optimierung ganzzahliger linearer Programme findet sich in [17, 21, 27]. Deutschsprachige Fachbücher zu diesen Themen sind u. a. [10, 11].

5 Optimierung in Aktion: Was bringt's?

In den vorangehenden Kapiteln wurden flexible, vielseitig einsetzbare mathematische Modelle und Optimierungsverfahren zur Abbildung und Lösung verschiedener Netzplanungsaufgaben in der Telekommunikation vorgestellt. Die Modelle und die Methoden sind ein Ergebnis unserer Forschungsarbeit auf der Basis langjähriger Kooperationen mit verschiedenen Netzbetreibern. Die entwickelten Verfahren wurden bei Netzplanungsstudien eingesetzt und innerhalb unserer Netzplanungs- und -optimierungssoftware DISCNET umgesetzt, die in unterschiedlichen Ausprägungen bei den Betreibern verwendet wird.

Das Ziel der folgenden Abschnitte ist es, an den konkreten praktischen Netzplanungsaufgaben aus Kapitel 2 die Anwendung der vorgestellten Techniken zu veranschaulichen und dabei das Potential des Einsatzes mathematischer Optimierungsmethoden in der Praxis aufzuzeigen.

Um dem Leser das Nachvollziehen unserer Berechnungen zu ermöglichen, sind die Datensätze und die berechneten Lösungen aller im Folgenden vorgestellten Planungsaufgaben im Internet unter `www.zib.de/telekomaktuell2003` öffentlich zugänglich. Aus Gründen der Vertraulichkeit sind die ursprünglichen Daten allerdings anonymisiert und leicht verändert worden.

5.1 Netzstrukturplanung

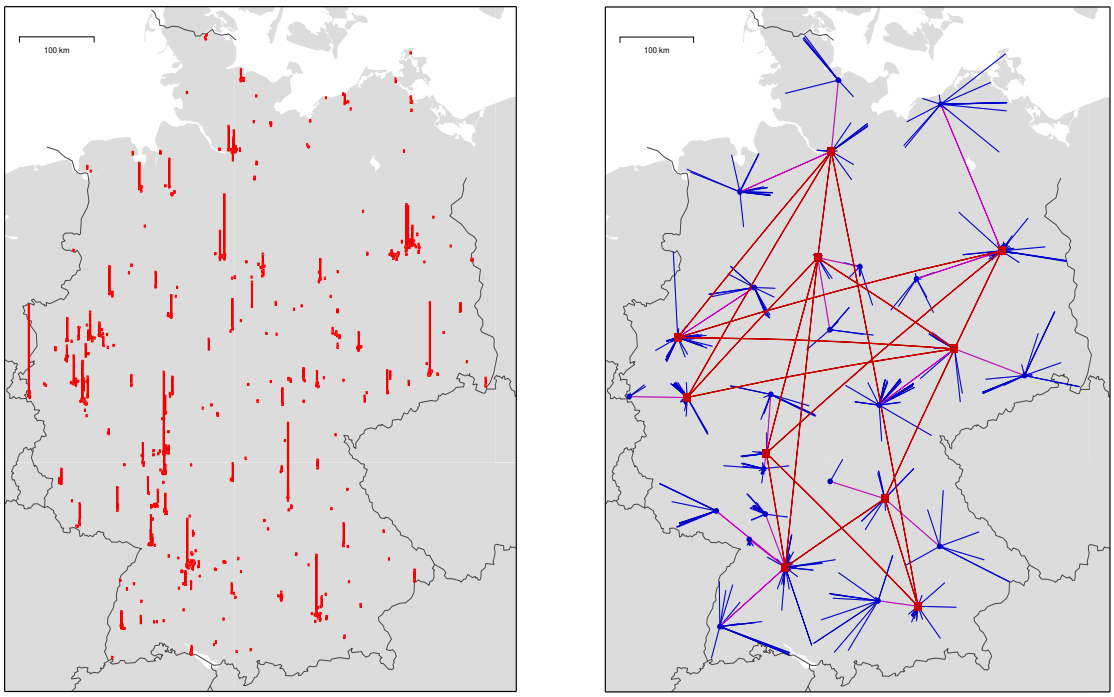
In Abschnitt 2.1 haben wir die Netzstrukturplanung beschrieben, die beim langfristigen Auf- oder Ausbau großer Netze mit mehreren Hierarchieebenen durchgeführt wird. In den letzten Jahren haben wir in verschiedenen Projekten solche Planungsaufgaben für unterschiedliche Kommunikationsnetze bearbeitet. In diesem Abschnitt stellen wir nun exemplarisch Resultate für zwei konkrete Szenarien aus der Entwurfsplanung des Gigabit-Wissenschaftsnetzes (G-WiN) vor.

5.1.1 Entwurfsplanung des G-WiN

Das vom DFN-Verein betriebene Deutsche Forschungsnetz ist das von der Wissenschaft selbst verwaltete Hochleistungsnetz für Lehre und Forschung in Deutschland. Es verbindet zahlreiche Hochschulen und Forschungseinrichtungen miteinander und unterstützt die Entwicklung und Erprobung neuer Anwendungen. Der nationale Kern des Forschungsnetzes ist das Gigabit-Wissenschaftsnetz G-WiN, das im Sommer des Jahres 2000 das damalige Breitband-Wissenschaftsnetz (B-WiN) ersetzte, um mit dem beständig wachsenden Datenvolumen Schritt halten zu können und eine auch zukünftigen Anforderungen gerechte Infrastruktur zu schaffen. Das G-WiN verbindet mehr als 750 Einrichtungen miteinander und sollte bereits in seiner ersten Ausbaustufe im Sommer 2000 ein Datenvolumen von ca. 220 Terabyte pro Monat bewältigen können.

Der Entwurf eines solchen Netzes stellt allein wegen seiner Größe bereits eine Herausforderung dar. Zunächst musste eine geeignete Struktur des Netzes ermittelt werden. Dabei sollte ein relativ kleines, aber hochkapazitatives Kernnetz aus wenigen Standorten gebildet werden, an das alle anderen Einrichtungen – ggf. in weitere Hierarchieebenen unterteilt – anzuschließen waren. Wegen seiner zentralen Rolle war das Kernnetz der höchsten Ebene ausfallsicher zu planen. Demgegenüber sollten die unteren Hierarchieebenen des Netzes in sternförmig gegliederten Access-Bereichen nur einfach mit dem Kernnetz verbunden werden. Neben dieser typischen Netzstrukturplanungsaufgabe, die im Folgenden näher betrachtet wird, muss nach dem Aufbau des Netzes für den tatsächlichen Verkehr ein geeignetes Traffic Engineering durchgeführt werden. Auf diese eher operative Planungsaufgabe gehen wir in Abschnitt 5.4 genauer ein.

Die Netzstrukturplanung für das G-WiN umfasste unter anderem die Entscheidungen, an welchen geografischen Standorten Netzknoten welcher Hierarchieebene eingerichtet werden, wie diese untereinander zu verbinden und wie die Verbindungen und die Knoten zu dimensionieren sind. Das primäre Ziel dieser strategischen Planungsaufgabe war die Minimierung der Gesamtnetzkosten. Da es zudem praktisch unvertretbar ist, in kürzeren Zeitabständen größere Änderungen an der Struktur des Netzes vorzunehmen, wie etwa Kernnetzknoten zu verlegen, sollte ein entsprechendes Verkehrswachstum bereits bei der



(a) Quellbedarfe

(b) Netzentwurf

Abbildung 19: G-WiN: Bedarfsstruktur und Planungsergebnis für die erste Ausbaustufe.

Planung der ersten Ausbaustufe des G-WiN berücksichtigt werden. Um eine sichere zukunftsorientierte Planung zu gewährleisten, wurde von einem jährlichen Verkehrszuwachs um den Faktor 2.2 ausgegangen. Dazu sind entsprechend weit in die Zukunft extrapolierte Verkehrsdaten aus dem Vorgängernetzwerk B-WiN zugrunde gelegt worden. Im Dezember 2002 lag der monatliche Transfer erstmals über einem Petabyte und bestätigte damit die angenommene Zuwachsrate.

5.1.2 Problemdimensionen und Planungsszenarien

Das G-WiN umfasst insgesamt 759 Einrichtungen mit den in Abbildung 19(a) angedeuteten Bedarfen. Das erste hier vorgestellte Planungsszenario zielt auf eine äußerst einfache Netzstruktur ab. Es beinhaltet lediglich zwei Hierarchieebenen L1 und L2. Die L1-Knoten bilden das Kernnetz und sind untereinander mindestens zweifach zusammenhängend zu verbinden. Die L2-Knoten sind jeweils direkt mit einem der L1-Knoten zu verbinden. Prinzipiell können an den 30 größten der 759 Standorte L1-Knoten eingerichtet werden; aus betrieblichen Gründen ist ihre Anzahl aber auf höchstens 20 Knoten begrenzt. Alle L1- und alle L2-Knoten werden jeweils identisch ausgebaut, für die Verbindungen stehen verschiedene Linkkapazitäten zwischen 128 kbit/s und 2,4 Gbits/s zur Verfügung. Kosten entstehen für das Einrichten eines jeden L1-Knotens unabhängig von seinem Standort sowie für jeden gewählten Link in Abhängigkeit von seiner Länge und seiner Kapazität. Die Kosten für das Einrichten von L2-Knoten sind für die Optimie-

rung irrelevant, da die entsprechende Gerätetechnik für den Access-Bereich auch an allen L1-Knoten und somit an allen Standorten bereitgestellt werden muss.

Das zweite Szenario ermöglicht eine stärker untergliederte Netzstruktur mit drei Ebenen L1, L2 und L3. Für die L1- und L2-Knoten gelten die gleichen Bedingungen wie beim ersten Szenario, und jeder L3-Knoten wird direkt an einen der L2- oder L1-Knoten angebunden (siehe auch Abbildung 19(b)). In diesem Szenario können an 261 der 759 Standorte Knoten der höheren Hierarchieebenen L1 oder L2 eingerichtet werden, da dort die prinzipiell nötige Infrastruktur vorhanden ist. Die übrigen Standorte können nur der unteren Hierarchieebene L3 zugeordnet werden. Als Randbedingungen sind allerdings vorgegeben, dass insgesamt 10 Knoten in Ebene L1 und 20 Knoten in Ebene L2 liegen und dass an jedem L1-Knoten genau zwei L2-Knoten angebunden werden sollen. Wie im ersten Szenario werden auch hier alle L1-, L2- und L3-Knoten jeweils identisch ausgebaut, und es stehen Linkkapazitäten zwischen 128 kbit/s und 2,4 Gbits/s zur Verfügung. Aus Sicherheitsgründen müssen für die Links zwischen den L1- und L2-Knoten jedoch Doppelverbindungen aus jeweils zwei physikalisch getrennten Leitungen gleicher Kapazität verwendet werden. Kosten entstehen auch hier durch das Einrichten der L1- und L2-Knoten sowie der Links.

5.1.3 Die erzielten Ergebnisse

Zur Lösung dieser sehr komplexen Planungsaufgaben wurden verschiedene Ansätze und Techniken verwendet und durch einige speziell für diese Problemklasse entwickelten Verfahren erweitert. Dabei kamen neben den in Kapitel 4 beschriebenen Methoden der gemischt-ganzzahligen Programmierung auch Kombinationen aus geschickt operierenden Heuristiken und alternative Relaxierungs- und Dekompositionstechniken zum Einsatz.

Für das erste Planungsszenario konnten mit einem integrierten Lösungsverfahren außerordentlich gute Lösungen berechnet werden – und das in kurzer Zeit! Auf einem handelsüblichen PC wurden in weniger als 15 Minuten Lösungen berechnet, die von der gleichzeitig ermittelten unteren Kostenschranke um weniger als 1% abweichen und daher nachweislich fast optimal sind.

Bei der Lösung des zweiten Planungsszenarios wurde auf einen Dekompositionsansatz zurückgegriffen, der das Gesamtproblem in zwei Teile zerlegt, diese unabhängig voneinander löst und die Teillösungen schließlich wieder zu einer Gesamtlösung kombiniert (vgl. [5, 6]). Das erste Teilproblem umfasst die Auswahl der L1- und L2-Knoten sowie die Anbindungen der L2- an L1-Knoten sowie der L3- an L1- und L2-Knoten. Sowohl dieses sogenannte Clustering als auch das zweite Teilproblem, die L1-Kernnetzplanung, konnten mit den in Kapitel 4 beschriebenen Techniken optimal gelöst werden. Der Dekompositionsansatz erleichterte es den Planern, den Einfluss verschiedener Parameter auf das Zugangs- und das Kernnetz einzeln zu analysieren und die Randbedingungen und Parameter so schrittweise gemäß ihren Vorstellungen zu präzisieren. Das resultierende Netzwerk, die erste Ausbaustufe des G-WiN, ist in Abbildung 19(b) dargestellt.

Probleme dieser Größenordung können nur durch mathematisch fundierte Verfahren optimal oder nahezu optimal gelöst werden. So ist das integrierte Lösen des ersten Planungsszenarios überhaupt erst dadurch möglich, dass durch die Verwendung spezieller Relaxierungstechniken (siehe [23]) die meisten Variablen und Nebenbedingungen implizit betrachtet werden können. Während eines Durchlaufes des Optimierungsverfahrens werden für dieses Planungsszenario insgesamt „nur“ ca. 16.000 Variablen und 40.000 Bedingungen tatsächlich erzeugt. Würde man dagegen versuchen, das gemischt-ganzzahlige lineare Programm komplett aufzuschreiben, so hätte man allein $8.9 \cdot 10^{30}$ verschiedene Routingweg-Variablen zu

berücksichtigen. Dies liegt weit jenseits aller Speichermöglichkeiten, und allein das Aufzählen aller Wege würde mit heutigen Computern die geschätzte Lebensdauer unseres Sonnensystems überschreiten.

5.2 Kostenoptimierter Ausbau von SDH-Netzen

Auf Basis der in Abschnitt 2.2 beschriebenen Planungsrandbedingungen für SDH-Netze betrachten wir im Folgenden eine Aufgabe, die von den Netzbetreibern periodisch zu lösen ist. Ausgangspunkt ist das existierende Netz, bestehend aus den Standorten mit konfigurierten Netzelementen, die über Glasfasern, Mietleitungen oder Richtfunkstrecken miteinander verbunden sind. Die Planungs- und Optimierungsaufgabe ist einfach zu beschreiben: Für mittelfristig prognostizierte Bedarfsanforderungen ist eine kostenoptimierte Ausbauplanung zu entwickeln.

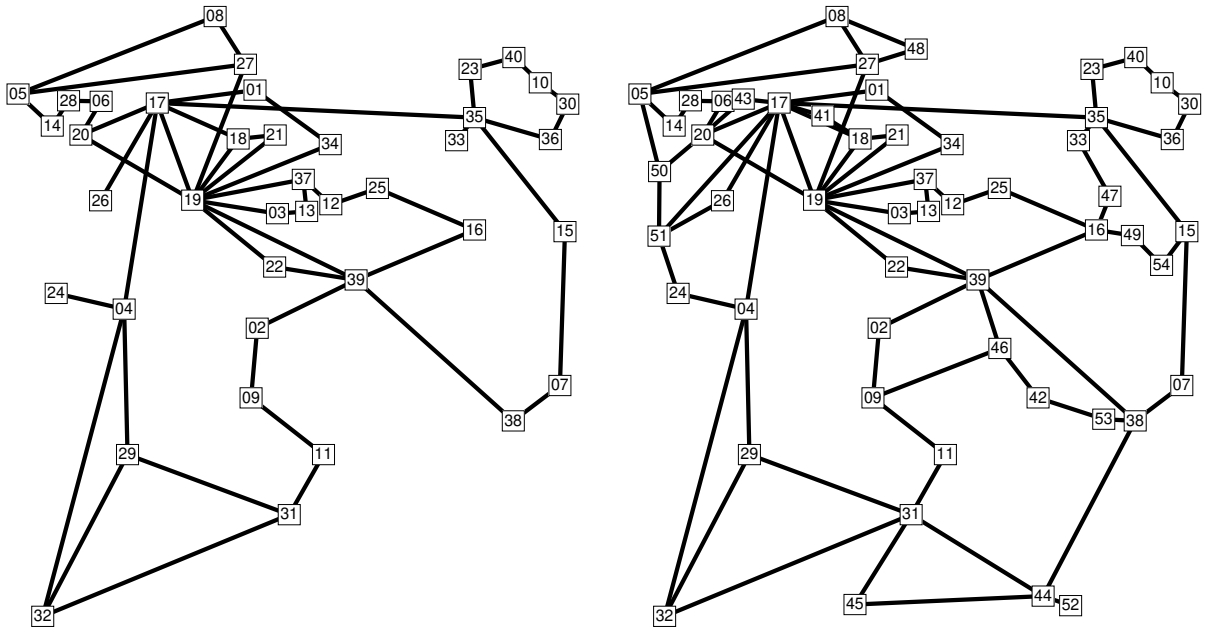
Wir haben dieses Planungsszenario mit den in Kapitel 3 beschriebenen Methoden modelliert und zur Lösung unser Planungswerkzeug DISCNET eingesetzt, das im Kern auf einem Branch&Cut-Verfahren basiert, wie es in Kapitel 4 beschrieben ist. Mit diesem Szenario unterstreichen wir die Vorteile eines Branch&Cut-Verfahrens gegenüber herkömmlichen heuristischen Lösungsverfahren. Es werden hier nicht nur konkrete Vorschläge zum Aus- und Umbau des existierenden Netzes berechnet. Gleichzeitig (inhärent im Verfahren) werden Informationen zur quantitativen Bewertung dieser Lösungen bereitgestellt. Damit wird eine Abschätzung der Güte von Lösungen ermöglicht und im Idealfall die Optimalität der berechneten Netzvorschläge garantiert.

5.2.1 Das konkrete Planungsszenario

Das existierende und dem folgenden Szenario zugrunde liegende SDH-Netz hat die in Abbildung 20(a) dargestellten 40 Standorte, die durch 42 Glasfasern miteinander verbunden sind. Die Linkkapazitäten sind STM-4 und STM-16, und es sind 503 VC-12-Bedarfe geroutet.

Für die Ausbauplanung ist das Netz um bis zu 14 Standorte zu erweitern. Ferner gibt es 80 Glasfaserverbindungen, d. h., zusätzlich zu den 42 existierenden stehen zur Planung weitere 38 Glasfaserverbindungen als Erweiterungsalternativen zur Verfügung. Diese dürfen in der Zieltopologie verwendet werden, es sollen aber möglichst wenige nach Bedarfs- und Kostengesichtspunkten ausgewählt werden. Die daraus resultierende potentielle Topologie für den Ausbau ist in Abbildung 20(b) dargestellt.

Hardware und Kapazitäten An den Standorten des existierenden Netzes sowie an den 14 neuen Standorten ist es möglich, einen von zwei Typen von Netzelementen (NE1 und NE2) auszuwählen und zu konfigurieren. Für Standorte, die weder Quelle noch Senke von Kommunikationsbedarfen sind, muss nicht zwingend ein Netzelement ausgewählt werden. Das NE1 hat 2 Aggregatsteckplätze für STM-4-Schnittstellenkarten, eine Gesamtschaltkapazität von 4032·VC-12 und kostet 8500 (fiktive Einheiten). Das NE2 hat 2 Aggregatsteckplätze für STM-16-Karten, und zusätzlich können bis zu 8 Steckplätze mit Schnittstellenkarten vom Typ STM-4 oder STM-16 bestückt werden. Die Gesamtschaltkapazität von NE2 ist daher 16128·VC-12, und ein solches Element kostet 18000. Beide Typen von Netzelementen können VC-12 verschalten. Alle Karten haben einen einzigen Port entsprechender Kapazität und kosten zwischen 1400 für eine STM-4-Karte und 3800 für eine STM-16-Karte. Topologieerweiterungen sind in begrenztem Maße erlaubt. Als neue Links kommen dabei die oben beschriebenen 38 potentiellen Glasfaserverbindungen in Frage. Auf jedem der Links kann als Übertragungskapazität eine Kombination der



(a) Ausgangsnetz

(b) Potentielle Ausbautopologie

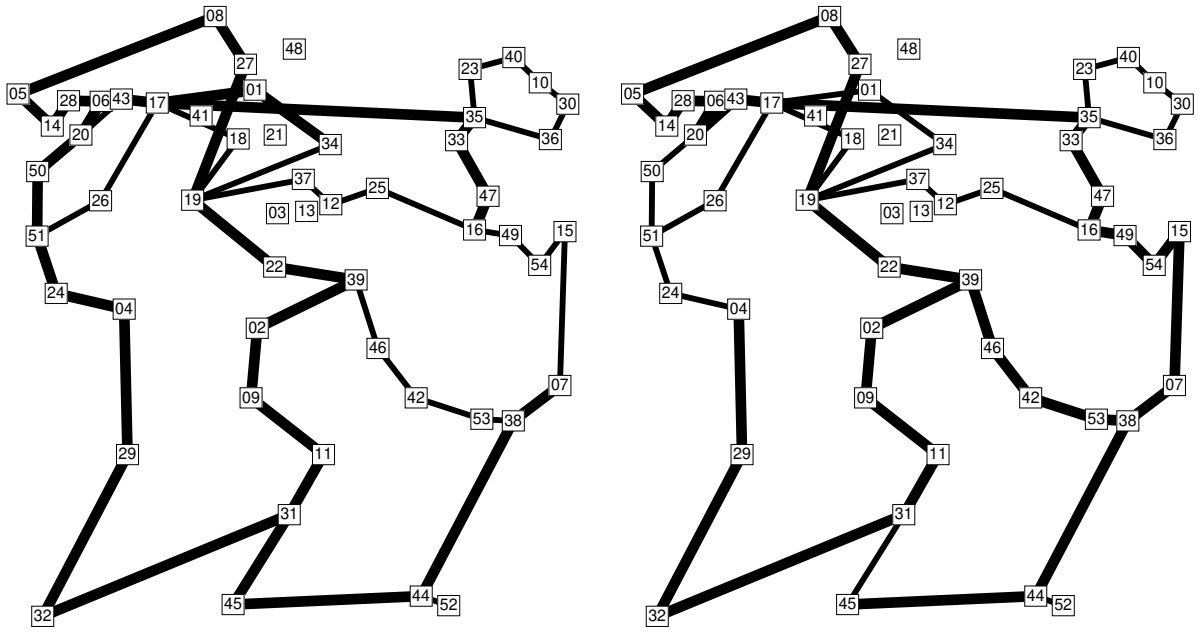
Abbildung 20: Netzstrukturen im SDH-Szenario.

Kapazitäten STM-4 und STM-16 installiert werden. Die Kosten für die Übertragungskapazitäten setzen sich aus den Kosten für die jeweils notwendigen Karten und kilometerabhängigen Glasfaserkosten zusammen. Jeder Kilometer Glasfaser kostet 10000.

Bedarfe, Routing und Ausfallsicherheit Die Kommunikationsbedarfe stammen aus unterschiedlichen Quellen: dem Sprachnetz, dem ATM-Netz und dem Signalisierungsnetz. Spezifiziert ist jeder Bedarf in einer Anzahl von VC-12. In der bereits existierenden Beschaltung sind 503 VC-12 vorkonfiguriert. Zusätzlich sollen für eine mittelfristige Prognose 432 weitere VC-12 konfiguriert werden. Eine ausgewählte Teilmenge von Bedarfen ist mit SNCP (siehe Abschnitt 2.2) zu schützen. Der Schutz durch SNCP soll bei Ausfall eines einzelnen Standortes sowie beim Ausfall einer einzelnen Verbindung gewährleistet sein.

Ziel der Planung In diesem Planungsszenario ist in einem Planungsschritt integriert zu entscheiden,

- ob Standorte ohne Bedarfsanforderungen eingerichtet werden sollen,
- wie eingerichtete Standorte zu konfigurieren sind,
- mit welchen Links die Topologie erweitert werden soll,
- welcher Link mit welcher Kapazität beschaltet wird und



(a) Beste Lösung (11784400)

(b) Alternativlösung (11816800)

Abbildung 21: Lösungen für das SDH-Netzplanungsszenario.

- wie jeder einzelne VC-12-Bedarf entsprechend seinen individuellen Ausfallsicherheitsanforderungen zu konfigurieren ist.

Als übergreifendes Ziel sollen dabei die Kosten für die Erweiterungen des existierenden Netzes minimiert werden.

5.2.2 Die erzielten Ergebnisse

Unter den hier beschriebenen Rahmenbedingungen stellt sich dem Planer eine komplexe Aufgabe, bei der eine Reihe von sich gegenseitig beeinflussenden Entscheidungen zu treffen sind. So impliziert das Routing der VC-N-Bedarfe die Kapazitätsanforderungen der einzelnen Links. Diese Kapazitäten sind technologisch bedingt jedoch nur in den entsprechenden STM-N-Granularitäten verfügbar, und auch die Netzelemente sind unter anderem bezüglich der Schaltkapazität und der verfügbaren Steckplätze beschränkt. Erfordert nun das Routing auf einem bestimmten Link nur einen VC-12 mehr als die 63 verfügbaren VC-12 eines STM-1, so wird sofort ein weiterer STM-1 oder ein STM-4 notwendig und das Netz dadurch teurer. Dementsprechend ermöglicht nur die Integration aller genannten Entscheidungen in einen einzelnen Planungsschritt das vorhandene Optimierungspotential voll auszuschöpfen. Eine sequentielle Planung, die beispielsweise erst eine Netztopologie bestimmt, darauf zulässige Routings berechnet, um anschließend die Linkkapazitäten und Netzelemente zu konfigurieren, kann kaum in der Lage sein, Entsprechendes zu leisten.

In den Abbildungen 21(a) und 21(b) sind zwei ausgewählte Lösungen dargestellt. Dazu stellt Abbil-

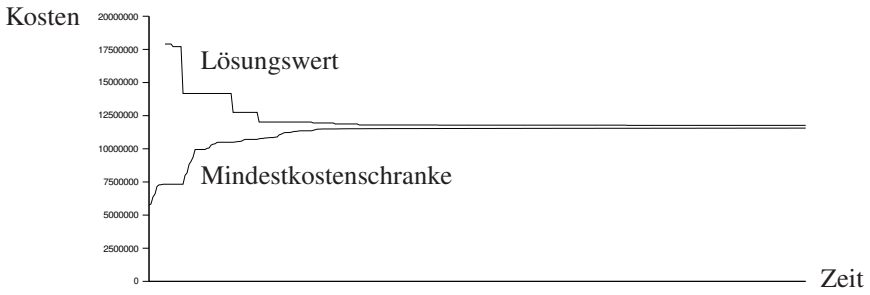


Abbildung 22: Zeitlicher Verlauf des Branch&Cut-Verfahrens.

dung 22 den Verlauf des verwendeten Branch&Cut-Verfahrens (siehe auch Abschnitt 4.4) für das beschriebene Planungsszenario dar.

Zum einen sieht man, dass im Laufe der Zeit immer wieder neue Lösungen berechnet werden und damit schrittweise eine entsprechende Verbesserung der zuvor gefundenen Lösungen erzielt wird. Die untere Kurve repräsentiert die inhärent berechnete Information über Mindestkosten des zu planenden Netzes. Diese Mindestkosten steigen im Laufe des Verfahrens und nähern sich dem Wert der besten gefundenen Lösung an. Das für die Praxis Entscheidende an dieser zusätzlichen Information ist: der Planer kann nun erstmals mathematisch fundiert eine kostenmäßige Bewertung einer Netzerweiterung vornehmen. Am Beispiel dieses Szenarios wird es dadurch möglich in folgender Weise zu argumentieren:

> Unter den bekannten Rahmenbedingungen ist es nicht möglich, eine Erweiterung des Netzes vorzunehmen, die weniger als 11526222 Einheiten kostet. Die beste berechnete Netzerweiterung kostet 11798700 (= 11526222+272488). Im günstigsten Fall könnte also ein anderes Verfahren 272488 einsparen, d. h. nicht mehr als maximal 2 Prozent.

Im Gegensatz dazu kann der Planer bei den üblicherweise implementierten heuristischen Verfahren nur die absoluten Kostenwerte von ermittelten Lösungsalternativen miteinander vergleichen. Dabei kann aber nicht eingegrenzt werden, um wieviel günstiger eine andere (vor allem die optimale) Lösung höchstens sein könnte.

5.3 Ressourceneffiziente Ausfallsicherheit in optischen Netzen

Bei der in Abschnitt 2.3 beschriebenen Planung optischer Netze besteht das Ziel darin, ein den gestellten Bedarfsanforderungen gerechtes Netzwerk zu möglichst geringen Kosten zu entwerfen. Aufgrund der typischerweise sehr hohen Verkehrsvolumina kommt dabei dem Schutz vor Störsituationen besondere Bedeutung zu. In diesem Abschnitt stellen wir ein Ausfallsicherheitskonzept vor, das wir speziell für optische Netze entwickelt haben, und das erst durch einen integrierten Lösungsansatz für das umfassende Planungsproblem ermöglicht wird. Um die gewünschte Sicherheit zu gewährleisten, erfordert dieses neuartige Konzept deutlich weniger Netzwerkressourcen als der gegenwärtig am meisten umgesetzte Mechanismus, dessen Ergebnisse als Vergleich herangezogen werden.

5.3.1 Ausfallsichere Planung optischer Netze

In optischen Kernnetzen wird die Kommunikation zwischen Knoten durch die Einrichtung von Lichtwegen realisiert, vgl. Abschnitt 2.3.2. Um diese gegen Ausfälle zu schützen, sind entsprechende Ersatzverbindungen vorzusehen. Das einfachste und in der Praxis am häufigsten realisierte Konzept dafür stellt 1+1-Protection dar, bei dem jedem abzusichernden Lichtweg ein Ersatz-Lichtweg zugeordnet ist, der auf einem unabhängigen Weg geroutet wird. Unabhängig bedeutet dabei, dass beide Lichtwege durch keine gemeinsame Netzkomponente (außer ihren Endknoten) verlaufen und somit nicht gleichzeitig ausfallen können. Aufgrund der vorab eingerichteten Ersatzwege ermöglicht 1+1-Protection eine äußerst schnelle Wiederherstellung ausfallender Verbindungen, bindet jedoch gleichzeitig enorm hohe Reservekapazitäten.

Auf eine effizientere Nutzung der Ressourcen zielen sowohl Varianten von 1+1-Protection (wie etwa 1:1- oder M:N-Protection) als auch gänzlich andere Konzepte ab, zum Beispiel 'Shared Protection'-Ansätze (vgl. [16]). Diese Mechanismen führen allerdings entweder zu komplexen und damit langsamen Verfahren zur (zeitkritischen) Wiederherstellung von Verbindungen oder bieten nur eingeschränkten Schutz. Um solche Nachteile zu vermeiden und dennoch die geforderte Ausfallsicherheit kapazitätsschonend zu realisieren, wurde in Zusammenarbeit mit unseren Kooperationspartnern speziell für optische Netze das Konzept 'Demand-wise Shared Protection' (DSP) entwickelt (vgl. [18]). Sein Vorteil basiert auf einer effizienteren Ausnutzung der Netzstruktur, so dass die Anzahl notwendiger Ersatzwege reduziert werden kann. Eine zentrale Rolle spielt dabei der Netzzusammenhang jedes Knotenpaares, der als maximale Anzahl paarweise unabhängiger Wege zwischen den betrachteten Knoten definiert ist.

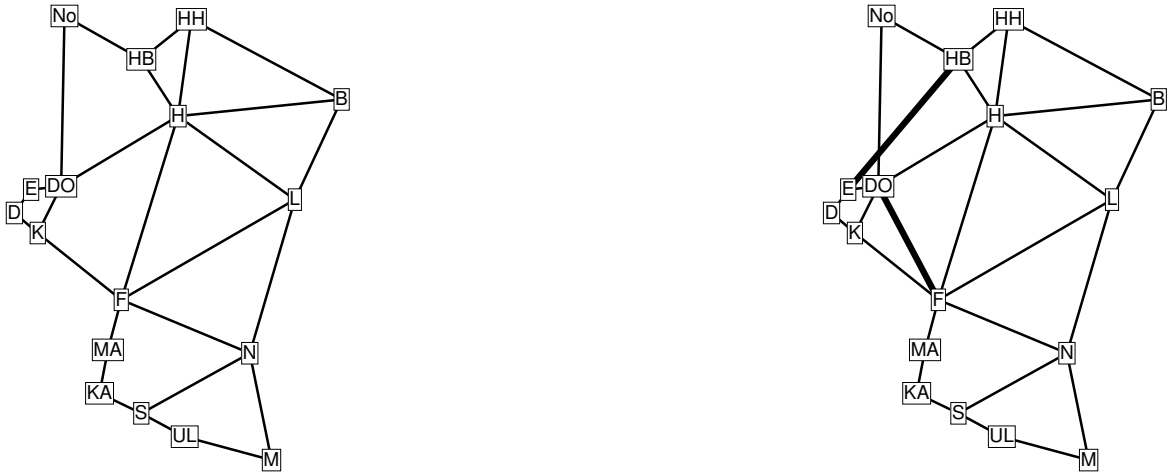
Anstatt wie bei 1+1-Protection jeden Lichtweg einzelnen mit seinem Ersatzweg zu berechnen, wird mit dem Konzept DSP das gesamte Routing eines Demands in einem Schritt ermittelt. Im Rahmen eines integrierten Ansatzes wird es dadurch möglich, den Netzzusammenhang optimal auszunutzen und die zu routenden Verbindungen eines Bedarfes auf eine maximale Anzahl unabhängiger Wege zu verteilen. Da so weniger Verbindungen durch die gleiche Komponente laufen, sind bei ihrem Ausfall auch nur weniger Ersatzschaltungen auf alternativen Wegen nötig. In vermaschten Netzen lassen sich so in der Regel zahlreiche Ersatz-Lichtwege und damit die dafür notwendigen Kapazitäten einsparen.

Zur Verdeutlichung des Konzeptes dient das folgende Beispiel: Für einen Bedarf mit fünf Lichtwegen sollen drei gegen Ausfall geschützt werden. Mit 1+1-Protection wären daher insgesamt acht Lichtwege zu routen: drei Paare von jeweils einem Lichtweg mit zugeordnetem (unabhängigen) Ersatzweg sowie zwei ungeschützte Lichtwege. Sofern das Netz für dieses Knotenpaar dreifach zusammenhängt, erfordert DSP lediglich fünf Lichtwege bei gleichem Schutz: dazu werden auf drei unabhängigen Wegen jeweils zwei, zwei und ein Lichtweg geroutet. In jeder Ausfallsituation bleiben damit stets drei Lichtwege erhalten. Selbst bei nur zweifachem Zusammenhang sind ingesamt nur sechs Lichtwege nötig: jeweils drei auf zwei unabhängigen Wegen.

In Tabelle 1 sind die benötigten Ersatz-Lichtwege für 1+1-Protection sowie für das Konzept DSP aufgeführt. Als Grundlage dienen Bedarfshöhen d zwischen einem und 50 Lichtwegen, von denen (aufgerundet) zwei Drittel oder alle Lichtwege gegen Ausfall eines beliebigen Links oder Knotens geschützt werden sollen. Die Werte für das Konzept DSP sind ferner nach dem zur Verfügung stehenden Netzzusammenhang aufgeschlüsselt. Diese (netzunabhängige) Übersicht verdeutlicht, dass das Konzept DSP stets weniger zusätzliche Lichtwege erfordert als 1+1-Protection und zudem insbesondere für geringen Schutzbedarf sowie bei steigendem Netzzusammenhang vorteilhaft ist. Die im Folgenden dokumentierten Rechnungsergebnisse belegen das Potential der durch DSP erzielbaren Kosteneinsparungen und zei-

	d	1	2	3	5	7	10	15	20	35	50
$\frac{2}{3}$-Schutz des Bedarfes											
1+1-Protection		1	2	2	4	5	7	10	14	24	34
DSP	$k = 2$	1	2	1	3	3	4	5	8	13	18
	$k = 3$	1	1	0	1	1	1	0	1	1	1
	$k = 4$	1	1	0	1	0	0	0	0	0	0
Vollständiger Schutz des Bedarfes											
1+1-Protection		1	2	3	5	7	10	15	20	35	50
DSP	$k = 2$	1	2	3	5	7	10	15	20	35	50
	$k = 3$	1	1	2	3	4	5	8	10	18	25
	$k = 4$	1	1	1	2	3	4	5	7	12	17

Tabelle 1: Vergleich der Anzahl zusätzlich benötigter Lichtwege für 1+1-Protection und *DSP* , um Ausfallschutz bei verschiedenen Bedarfshöhen d sowie unterschiedlichem Schutzgrad und Netzzusammenhang k zu gewährleisten.



(a) Netztopologie

(b) Erweiterte Netztopologie

Abbildung 23: Netztopologie (17 Knoten, 26 Links) und Erweiterung um zwei Links.

gen zudem weitere Analysemöglichkeiten auf.

5.3.2 Vergleichsszenarien

Für die Vergleichsrechnungen hat unser Projektpartner das in Abbildung 23(a) dargestellte Netzwerk zur Verfügung gestellt. Die vermaschte Topologie besteht aus 17 Knoten und 26 Links und wird auch als

Szenario	Gesamtbedarf (# Lichtwege)	davon zu schützen	Gesamtanzahl Lichtwege zu routen 1+1-Protection	DSP	Reduktion
ungeschützt	686	0	686	686	0.0 %
$\frac{1}{3}$-Schutz	686	254	940	699	25.6 %
$\frac{2}{3}$-Schutz	686	478	1164	835	28.3 %
voller Schutz	686	686	1372	1191	8.7 %

Tabelle 2: Resultierende Anzahl der zu routenden Lichtwege für die beiden Ausfallsicherheitskonzepte 1+1-Protection und DSP.

Referenznetz innerhalb des BMBF-Projektes 'MultiTeraNet' verwendet (vgl. [20]). Zur Berechnung der Ergebnisse für 1+1-Protection wurde das Programm 'WDM Guru' der Firma OPNET [22] verwendet, während die Resultate für DSP mit unserem Softwarepaket 'OND' (Optical Network Design, in [29] detailliert vorgestellt) erzeugt wurden.

Zur Hervorhebung des Einflusses der beiden Ausfallsicherheitskonzepte verwenden wir ein gegenüber [29] vereinfachtes Komponentenmodell, das sich auf einen Fasertyp, einen WDM-System-Typ mit 40 bereitgestellten Wellenlängen, einen Wellenlängenkonverter-Typ und zwei OXC-Typen mit 64 bzw. 128 bidirektionalen Ports beschränkt. Die einheitlich skalierten Kosten der Komponenten betragen 193.2 für den 64-Port-OXC, 312.4 für den 128-Port-OXC, 24 für ein WDM-System sowie 2 pro Wellenlängenkonverter. Die Faserkosten setzen sich aus zwei Bestandteilen zusammen: einem Anteil von 0.05 pro Kilometer sowie Regenerationskosten von 6 je angefangenem 70-km-Abschnitt eines Links.

Der zu realisierende Verbindungsbedarf wurde von einem Populationsmodell abgeleitet und umfasst insgesamt 686 Lichtwege zwischen 57 verschiedenen Knotenpaaren, von denen 38 zweifach-, 18 dreifach- und ein Paar vierfach-zusammenhängend ist. Hinsichtlich der Ausfallsicherheit unterscheiden wir vier Anforderungsszenarien, die nach dem Anteil der zu schützenden Lichtweg-Verbindungen eines jeden Bedarfes benannt sind. Der Anteil variiert von 1, also vollem Schutz aller Lichtwege, über $\frac{2}{3}$- bzw. $\frac{1}{3}$-Schutz (wobei Brüche stets aufgerundet werden) bis hin zu 0, d. h. vollständig ungeschütztem Verkehr. Das letztgenannte Szenario weist natürlich keine konzeptionellen Unterschiede aus, sondern dient lediglich als Bezugspunkt zum Vergleich der beiden eingesetzten Verfahren, mit denen die Lösungen für beide Sicherheitskonzepte ermittelt wurden.

5.3.3 Die erzielten Ergebnisse

In Tabelle 2 wird zunächst für die betrachteten Szenarien aufgeführt, wieviele Lichtwege bei Anwendung der verschiedenen Ausfallsicherheitskonzepte tatsächlich insgesamt zu routen sind. Daraus ist bereits ersichtlich, dass durch Ausnutzung der Netzstruktur mit DSP insgesamt deutlich weniger Ersatz-Lichtwege erforderlich sind als für 1+1-Protection. Allerdings ist anzumerken, dass die breitere Streuung der Verbindungen im Durchschnitt zu längeren Lichtwegen führt. Dadurch erhöht sich zwar der Gesamtbedarf an Ressourcen für das Routing, jedoch wird dies durch die eingesparten Lichtwege und der dafür benötigten Kapazitäten bei weitem überkompensiert.

Anhand der Kostenergebnisse lässt sich diese Beobachtung nachhaltig belegen. Tabelle 3 listet die resultierenden Gesamtnetzkosten auf, die sich für die einzelnen Szenarien ergeben. Für die mit OND er-

Szenario	1+1-Protection Knoten-kosten	Link-kosten	DSP Knoten-kosten	Link-kosten	Güte (in %)	Gesamt-ersparnis (in %)
ungeschützt	6186	1465	6186	1436	11.0	0.0
$\frac{1}{3}$-Schutz	8225	2369	6663	1717	13.2	20.9
$\frac{2}{3}$-Schutz	10918	2963	8895	2623	6.5	17.0
voller Schutz	12093	3644	11855	3477	6.4	2.6

Tabelle 3: Gegenüberstellung der resultierenden Netzkosten für 1+1-Protection und DSP.

Szenario	DSP Lichtwege zu routen	Knoten-kosten	Link-kosten	Güte (in %)	Reduktion (in %) der Gesamtanzahl Lichtwege	Gesamt-kosten
$\frac{2}{3}$-Schutz	794	8567	2400	9.3	4.9	4.8
voller Schutz	1120	11557	3195	8.5	6.0	9.6

Tabelle 4: Ergebnisse für das erweiterte Netzwerk bei Anwendung von DSP.

rechneten Netzwerke ist zudem die Lösungsgüte (als relative Differenz zwischen den Gesamtkosten des ermittelten Netzwerkes und der besten erreichten Abschätzung unvermeidbarer Kosten) angegeben. Wie erwartet ist das Einsparpotential zwar geringer als die entsprechende Reduktion der Lichtweg-Anzahl, allerdings erlaubt das Konzept DSP deutlich günstigere Netze als die Anwendung von 1+1-Protection. Um das Potential der optimalen Ausnutzung der Netzstruktur weiter zu beleuchten, wird in einer abschließenden Untersuchung der Netzzusammenhang durch Hinzunahme der zwei in Abbildung 23(b) dargestellten zusätzlichen Links erhöht. Wir beschränken uns auf die beiden Szenarien mit höchstem Schutzbedarf, die durch eine erhöhte Vermaschung am meisten beeinflusst werden. Wie aus den Ergebnissen in Tabelle 4 entnommen werden kann, lassen sich durch geeignete Netzwerkerweiterungen nochmals Kostensenkungen erzielen, wobei das Potential mit steigenden Schutzanforderungen sogar steigt. Diesen zusätzlichen Einsparungen stehen allerdings auch die Investitionen zur Einrichtung der neuen Links gegenüber. Eine solche kostenbasierte Analyse kann folglich dazu dienen, fundierte Entscheidungen über mögliche Erweiterungen des Netzwerkes zu treffen.

5.4 Optimierung der Routinggewichte in IP-Netzen

In Abschnitt 2.4 wurde die Planungsaufgabe des Traffic Engineering für IP-Netze beschrieben. Ziel dabei war es, für gegebene Verkehrsanforderungen die Verkehrsflüsse durch eine geeignete Wahl der Routinggewichte an das Übertragungsnetz anzupassen. Die Topologie und die Kapazitäten des Übertragungsnetzes sind bei dieser operativen Aufgabenstellung meist vorgegeben und unveränderlich. In der Praxis ist dieses Problem zu lösen, wenn sich die Verkehrsanforderungen im existierenden Netz wesentlich ändern.

Das meistgenutzte Protokoll für das interne Routing in IP-Netzen ist OSPF. Die einzige Möglichkeit, beim OSPF-Routing die Verkehrsflüsse an das Netz anzupassen, besteht darin, die Routinggewichte zu

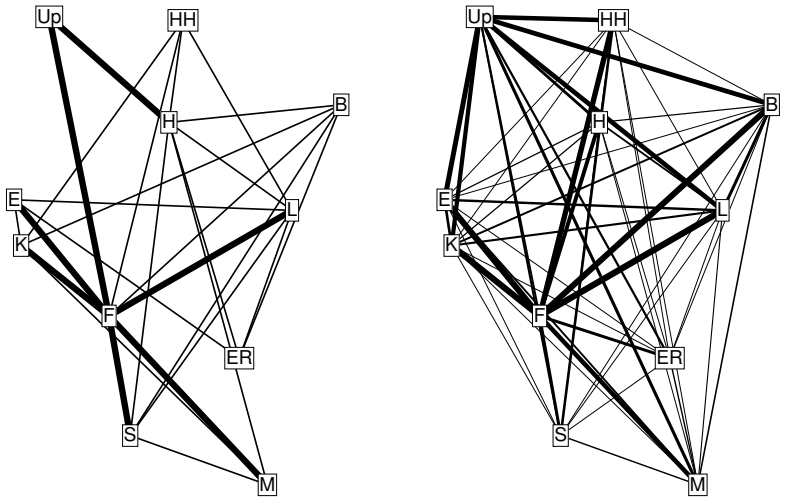


Abbildung 24: Topologie und Bedarfe für die 2. Ausbaustufe des Kernnetzes des G-WiN.

verändern. Prinzipiell kann der Administrator eines IP-Netzes diese Gewichte für jeden Link beliebig zwischen 0 und 65535 wählen. Praktisch werden jedoch meist recht einfache Strategien zur Wahl der Routinggewichte verwendet. Alle großen Hersteller von IP-Routertechnik haben in ihren Geräten dafür Standardeinstellungen. Cisco-Router benutzen beispielsweise als Standardeinstellung Routinggewichte, die umgekehrt proportional zur Linkkapazität sind. Andere einfache Strategien, die häufig in der Praxis verwendet werden, sind die Wahl von Einheitsgewichten für alle Links, die Wahl der geografischen Länge oder eines gemessenen Delays des Links sowie Kombinationen dieser Strategien. Alle solchen Strategien wählen die Gewichte aber unabhängig von den tatsächlichen Verkehrsanforderungen und -flüssen. Das führt häufig zu ungleichmäßigen Auslastungen der Links im Netz, bis hin zur völligen Überlastung einzelner Links.

Eine funktionierende Möglichkeit, dieses Problem zu beheben, liegt in der dynamischen Anpassung der Routinggewichte an die aktuellen Verkehrsflüsse im Netz. Dabei werden in kurzen Abständen die Auslastungen der Links gemessen und die Routinggewichte auf stark frequentierten Links erhöht und auf weniger belasteten Links reduziert. Verschiedene Untersuchungen haben aber gezeigt, dass solche Gewichtsänderungen selbst bei stabilen Verkehrsanforderungen zu unerwünschten Oszillationen im Netz und bei zeitlich stark schwankenden Verkehren sogar zu instabilen Zuständen führen können, siehe z. B. [8]. Es ist für die Zuverlässigkeit der Netze fast immer besser, die Routinggewichte in regelmäßigen Abständen geeignet neu zu bestimmen und für einen gewissen Zeitraum unverändert zu lassen.

Die in Kapitel 4 vorgestellten Methoden der mathematischen Optimierung lassen sich auch zur Bestimmung optimaler Routinggewichte einsetzen. Unter Berücksichtigung der vorhandenen Linkkapazitäten und der aktuellen oder prognostizierten Verkehrsanforderungen können damit für die Links die tatsächlich bestmöglichen Routinggewichte bestimmt werden, also solche Gewichte, für die die größte daraus resultierende Linkauslastung im Netz minimal wird.

In Abbildung 24 sind die Topologie und die Bedarfe für ein Planungsszenario aus der zweiten Ausbaustufe des Level-1-Kernnetzes des G-WiN dargestellt. Bereits an diesem relativ kleinen Netz lässt sich erkennen, dass sowohl die Standardeinstellungen der Routerhersteller als auch die einfachen Strategien,

wie Einheitsgewichte, längenabhängige Gewichte oder Kombinationen davon teilweise sehr schlechte Routings liefern. Die Linkauslastungen, die sich für das Szenario mit diesen Verfahren ergeben, und die Linkauslastungen bei optimal gewählten Routinggewichten sind in Abbildung 25 gegenübergestellt. Für die optimierten Gewichte sind die Auslastungen deutlich ausgeglichener, und die maximale Auslastung ist nicht einmal halb so groß wie die maximale Auslastung, die bei Verwendung der einfachen Strategien auftritt.

Durch die zusätzliche Berücksichtigung möglicher Ausfallszenarien kann die Robustheit des Netzes noch weiter verbessert werden. So kann sichergestellt werden, dass durch den Ausfall eines einzelnen Links oder Routers nicht das gesamte Netz in Mitleidenschaft gezogen wird, weil etwa durch die Wahl der Routinggewichte viel Verkehr über eine hochkapazitative Leitung fließt, beim Ausfall dieser aber auf den Ersatzwegen nicht genügend Kapazität zur Verfügung steht. Werden die möglichen Ausfallszenarien bereits bei der Bestimmung der Routinggewichte einbezogen, so lassen sich die in Fehlersituationen auftretenden Linkauslastungen häufig deutlich reduzieren. In Abbildung 26 sind die maximalen Auslastungen für alle oben beschriebenen Strategien sowie für die Routinggewichte aus der Optimierung mit Berücksichtigung der Fehlerszenarien gegenübergestellt. Es ist offensichtlich, dass nur durch eine Optimierung der Routinggewichte auch bei kurzfristigen Routerausfällen eine effiziente Nutzung der verbliebenen Kapazitäten gewährleistet wird.

6 Zusammenfassung

In diesem Artikel stellen wir die Anwendung und Methodik der mathematischen Optimierung vor. Anhand von vier ausgewählten konkreten Aufgabenstellungen aus der Festnetzplanung, die wir in Kooperation mit Industriepartnern in den letzten Jahren bearbeitet haben, werden neue Möglichkeiten zur Kosten- und Qualitätsoptimierung in Telekommunikationsnetzen aufgezeigt.

Zur Beschreibung des „mathematischen Ansatzes“ erläutern wir zunächst die grundlegenden Konzepte der mathematischen Modellbildung. Mit Hilfe der eingesetzten Abstraktionen können wir die planerisch sehr unterschiedlichen Aufgabenstellungen in einem einheitlichen Rahmen abbilden und gemeinsame Strukturen identifizieren. Auf dieser für die Mathematik charakteristischen Vorgehensweise basieren flexible und leistungsfähige Verfahren, die alle Entscheidungen umfassen und sämtliche Rahmenbedingungen berücksichtigen. Auf der Grundlage einfacher Enumerationstechniken werden zunehmend komplexere Bausteine moderner mathematischer Lösungsverfahren vorgestellt, bis hin zum hochentwickelten Branch&Cut&Price-Verfahren. Über die Berechnung guter Lösungen hinaus bieten diese Ansätze den Vorteil, zusätzliche Informationen bereitzustellen, mit denen eine fundierte Bewertung der Güte von Lösungsvorschlägen vorgenommen werden kann.

Als Beleg der Eignung und Güte der mathematischen Optimierung für die Praxis dienen vor allem die erzielten Ergebnisse für Planungsaufgaben, die das gesamte Spektrum von langfristigen strategischen Entscheidungen über den mittelfristigen Netzausbau sowie die Auswahl geeigneter Konzepte bis hin zur kurzfristigen operativen Steuerung abdecken. So wird deutlich,

- dass selbst für Netze mit mehr als 750 Standorten eine optimierte Unterteilung in Kernnetz und Zugangsnetze sowie deren bestmögliche Ausgestaltung vorgenommen werden kann,
- wie sich eine detaillierte Konfigurations- und Ausbauplanung für einzelne Netzschichten kostengünstig erreichen lässt,

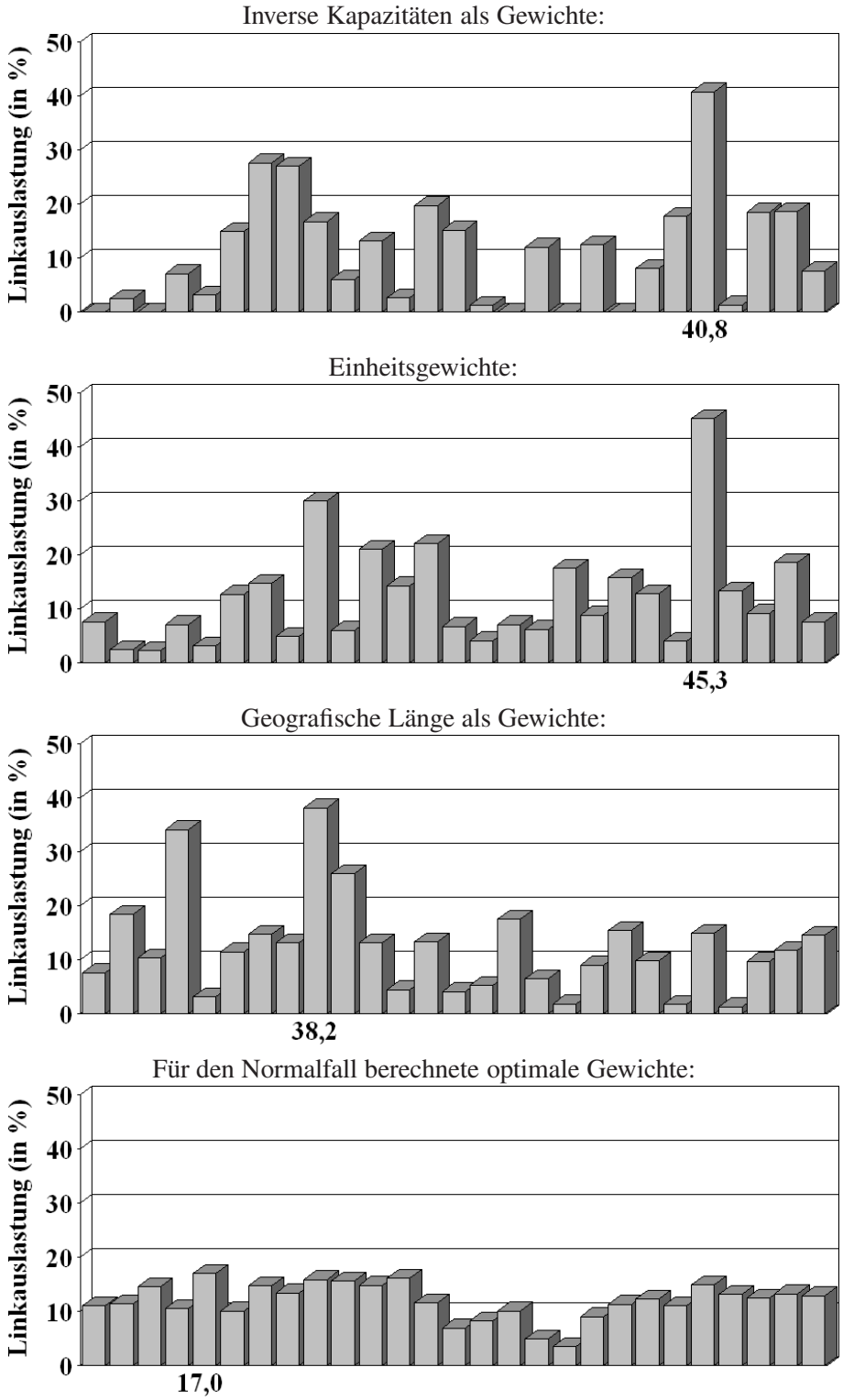


Abbildung 25: Linkauslastungen für verschiedene Routinggewichte.

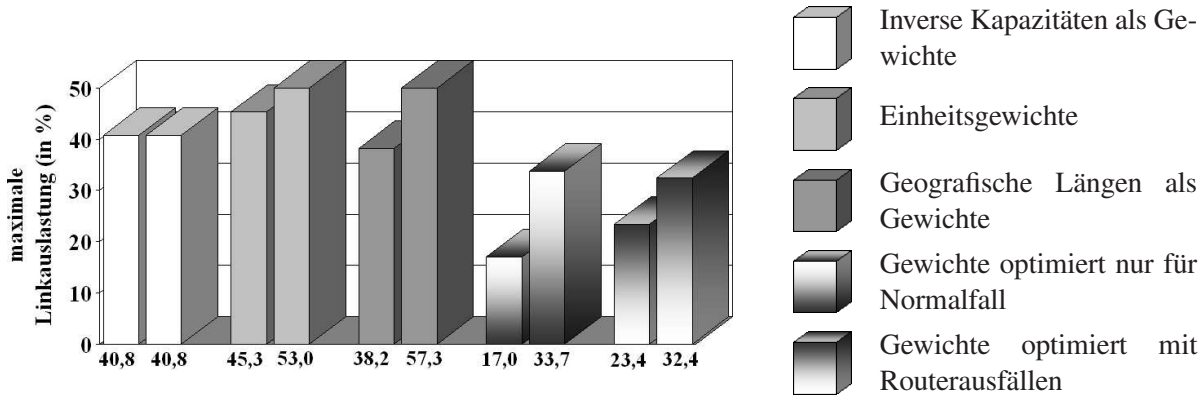


Abbildung 26: Vergleich maximaler Linkauslastungen jeweils ohne/mit Routerausfällen für verschiedene Routinggewichte.

- dass flexible Modelle und integrierte Lösungsansätze es ermöglichen, unterschiedliche Konzepte (z. B. für Ausfallsicherheit) quantitativ miteinander zu vergleichen, und
- wie durch optimiertes Traffic Engineering die maximale Linkauslastung in einem IP-Netz um mehr als die Hälfte reduziert werden kann.

Als Fazit kann festgehalten werden, dass die mathematische Optimierung in der Lage ist, die Netzplaner bei ihren vielfältigen strategischen und operativen Planungsaufgaben effektiv zu unterstützen.

Summary

The purpose of this article is to introduce the methodology and application of mathematical optimization in telecommunications and to highlight its striking potential for saving costs and improving qualities. The discussion is guided by four practical planning task cooperations with industrial partners.

The presentation of the ”mathematical approach” starts with a discussion of the basic concepts of mathematical modeling. The systematic use of abstraction assists greatly in identifying similar problem structures. Such structures are characteristic for mathematical models and enable us to formulate planning tasks of diverse nature within a common framework. This integrated model serves as basis for the development of flexible and efficient solution methods which comprise all decisions to be taken and to account for all relevant restrictions and relations. We incrementally introduce several components of advanced mathematical optimization solvers, from familiar enumeration techniques up to the elaborate branch&cut&price method. Besides good solutions, these sophisticated algorithms also provide information to assess the quality of any feasible solution.

For practice, the skill and quality of mathematical optimization is documented by the results achieved for planning tasks that cover the whole range from long-term strategic decisions, via mid-term network expansion and selection of concepts, to short-term operational control. The findings reveal

- that an optimized partition into and configuration of backbone and access networks can even be realized for large instances with more than 750 nodes,

- how a detailed and cost-efficient configuration and expansion of single network layers can be achieved,
- that flexible models and integrated approaches allow for the quantitative comparison of different concepts (e.g., for survivability), and
- how optimized traffic engineering can reduce the maximum load of links in IP-networks by more than 50 percent.

The bottom line is: mathematical optimization suits best to support network planners in effectively accomplishing their varying strategic and operational planning tasks.

Résumé

Dans cet article, la méthodologie et l'application d'optimisation mathématique ont été décrites. En présentant quatre problèmes réels de planification de réseaux de télécommunication sur lesquels on a travaillé pendant les dernières années en coopération avec des partenaires industriels, le potentiel de l'optimisation mathématique pour la réduction des coûts et l'amélioration de la qualité de solutions ont été démontrés.

Pour décrire l'abord mathématique, nous avons commencé avec une explication de techniques fondamentales de la modélisation mathématique. A l'aide des abstractions utilisées, on a identifié un cadre commun dans lequel les tâches de planification très différentes ont pu être représentées. Cet approche, qui est typique pour les mathématiques, sert de base pour des méthodes d'optimisatiyon très flexibles et puissantes qui prennent en considération toutes les décisions et contraintes.

Basé sur des méthodes d'énumération assez simples, des techniques d'optimisation de plus en plus complexes ont été introduits, qui sont utilisées dans les logiciels d'optimisation mathématique d'aujourd'hui. Au-delà de bonnes solutions, ces techniques révèlent des informations supplémentaires, ce qui nous permet d'évaluer la qualité de ces solutions.

L'utilité de l'optimisation mathématique pour des problèmes pratiques a été démontré par les solutions obtenues pour des tâches de planification de réseaux réels si différents que des décisions stratégiques à long terme, l'extension de réseaux à moyen terme et le choix de concepts, et le contrôle opérationel à court terme.

Les résultats montrent

- que même pour des réseaux avec 750 nœuds, une partition optimisée de ces nœuds en un réseaux à longue distance et des réseaux d'accès peut être déterminée en combinaison avec la configuration des nœuds,
- combien la configuration et l'extension détaillée d'une couche de réseau peut être effectué à un prix raisonnable,
- que des modèles flexibles et des abords intégrés permettent de comparer quantitativement de différents techniques (par example, pour protecter des réseaux contre des pannes techniques),

- comment par du traffic engineering optimisé, la congestion maximale des liens dans un réseau IP a pu être réduite par plus que la moitié.

Dressant le bilan, on note que l'optimisation mathématique peut effectivement soutenir des planificateurs de réseaux dans leurs multiples tâches stratégiques et opérationelles.

Literatur

[1] ALEVRAS, D., M. GRÖTSCHEL, P. JONAS, U. PAUL und R. WESSÄLY: *Survivable Mobile Phone Network Architectures: Models and Solution Methods*. IEEE Communications Magazine, 36(3):88–93, 1998.

[2] ALEVRAS, D., M. GRÖTSCHEL und R. WESSÄLY: *Cost–Efficient Network Synthesis from Leased Lines*. Annals of Operations Research, 76:1–20, 1998.

[3] BERTSIMAS, D. und R. S. FREUND: *Data Models and Decisions*. South-Western Educational Publishing, 2000. ISBN: 0538859067.

[4] BLEY, A., M. GRÖTSCHEL und R. WESSÄLY: *Design of Broadband Virtual Private Networks: Model and Heuristics for the B-WiN*. In: DEAN, N., D. F. HSU und R. RAVI (Hrsg.): *Robust Communication Networks: Interconnection and Survivability*, Bd. 53 d. Reihe *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, S. 1–16. AMS, 1998.

[5] BLEY, A. und T. KOCH: *Optimierung des G-WiN*. DFN-Mitteilungen, 54:13–15, 2000.

[6] BLEY, A. und T. KOCH: *Integer Programming Approaches to Access and Backbone IP-Network Planning*. Techn. Ber. ZR 02-41, Konrad-Zuse-Zentrum für Informationstechnik Berlin, 2002.

[7] CHVÁTAL, V.: *Linear Programming*. W. H. Freeman and Company, New York, 1983. ISBN: 0716715872.

[8] CROWCROFT, J. und Z. WANG: *Analysis of shortest-path routing algorithms in a dynamic network environment*. ACM SIGCOM Computer Communication Review, 22(2):63–71, 1992.

[9] DASH OPTIMIZATION: *Xpress-MP*, 2003. http://www.dashoptimization.com.

[10] DOMSCHKE, W. und A. DREXL: *Einführung in Operations Research*. Springer Verlag, 2002. ISBN: 3540429506.

[11] ELLINGER, T., G. BEUERMANN und R. LEISTEN: *Operations Research. Eine Einführung*. Springer Verlag, 2003. ISBN: 3540004777.

[12] *Gesellschaft für Operations Research Dissertationspreis 2002: A. Eisenblätter*. http://www.gor-ev.de.

[13] ILOG: *CPLEX*, 2003. http://www.ilog.com/products/cplex.

[14] *INFORMS Telecommunications Dissertation Award 2002: A. Eisenblätter*. http://www.informs.org/Prizes/tele.html#diss.

[15] *Jahresbericht 2002*. Konrad-Zuse-Zentrum für Informationstechnik Berlin, 2002. http://www.zib.de/bib/jb/.

[16] KERIVIN, H., B. LIAU und T. T. L. PHAM: *Survivable Capacitated Networks – Comparison of Shared Protection Mechanisms*. In: *Proceedings of Networks 2002*, S. 379–388, München, 2002.

[17] KORTE, B. und J. VYGEN: *Combinatorial Optimization. Theory and Algorithms*. Springer, Berlin, 2002. ISBN: 3540431543.

[18] KOSTER, A. M. C. A., A. ZYMOLKA, M. JÄGER, R. HÜLSERMANN und C. GERLACH: *Demand-wise Shared Protection for Meshed Optical Networks*. In: *Proceedings of the 4th Conference on the Design of Reliable Communication Networks (DRCN 2003)*, Banff, Kanada, 2003.

[19] KRÖLLER, A.: *Network Optimization: Integration of Hardware Configuration and Capacity Dimensioning*. Diplomarbeit, TU Berlin, 2003.

[20] MULTITERANET: *BMBF-Förderschwerpunkt*. http://www.pt-dlr.de/PT-DLR/kt/MultiTeraNet.htm.

[21] NEMHAUSER, G. L. und L. A. WOLSEY: *Integer and Combinatorial Optimization*. Wiley-Interscience, 1999. ISBN: 0471359432.

[22] OPNET: *WDM Guru*. http://www.opnet.com.

[23] ROCKAFELLAR, R. T.: *Lagrangian Multipliers and Optimality*. SIAM Review, 35(2):183–238, 1993.

[24] VANDERBEI, R. J.: *Linear Programming: Foundations and Extensions*. Kluwer Academic Publishers, Boston, 2001. ISBN: 0792373421.

[25] *Vodafone Innovationspreis 2001: R. Wessäly*. http://www.vodafone.de/unternehmen/presse/6518_8563.html.

[26] WESSÄLY, R.: *DISCNET – DImensioning Survivable Capacitated NETworks*. Doktorarbeit, Technische Universität Berlin, 2000.

[27] WOLSEY, L. A.: *Integer Programming*. Wiley-Interscience, 1998. ISBN: 0471283665.

[28] ZIB: *SOPLEX*, 2002. http://www.zib.de/Optimization/Software/Soplex/.

[29] ZYMOLKA, A., A. M. C. A. KOSTER und R. WESSÄLY: *Transparent optical network design with sparse wavelength conversion*. In: *Proceedings of ONDM 2003*, S. 61–80, Budapest, Hungary, 2003. The 7th IFIP Working Conference on Optical Network Design & Modelling.

A Verzeichnis der Abkürzungen und Symbole

Abkürzungen

ADM	Add-Drop-Multiplexer
ATM	Asynchronous Transfer Mode
B-WiN	Breitband-Wissenschaftsnetz
DFN	Deutsches Forschungsnetz
DXC	Digital Cross-Connect
DSP	Demand-wise Shared Protection
G-WiN	Gigabit-Wissenschaftsnetz
IP	Internet Protocol
MSC	Mobile Switching Center
MSP	Multiplex Section Protection
OSPF	Open Shortest Path First
OXC	Optical Cross-Connect
SDH	Synchronous Digital Hierarchy
SNCP	Sub-Network Connection Protection
STM	Synchronous Transport Module
VC	Virtual Container
VPN	Virtual Private Network
WDM	Wavelength Division Multiplexing
ZIB	Zuse-Institut Berlin

Symbole

g	Index für beliebigen Knoten oder Link
ℓ	Index für beliebigen Link
u	Index für beliebigen Knoten
k	Index für beliebige Komponente
r	Index für beliebige Ressource
q	Index für beliebige Bedarfsanforderung
K_g	Menge aller installierbaren Komponenten in Knoten oder Link g
$c_{k,g}$	Kosten für die Installation einer Komponente vom Typ k in Knoten oder Link g
$A(k,r)$	Angebot der Ressource r durch Komponente vom Typ k
$V(k,r)$	Verbrauch der Ressource r durch Komponente vom Typ k
s_q	Quelle des Bedarfes q
t_q	Senke des Bedarfes q
b_q	Basis-Routingeinheit des Bedarfes q
d_q	Bandbreitenanforderung des Bedarfes q in Basiseinheit b_q
$\mathcal{P}^q$	Menge aller für einen Bedarf q zulässigen Wege zwischen Quelle s_q und Senke t_q
$\mathcal{P}^q_\ell$	Teilmenge der Wege aus $\mathcal{P}^q$, die den Link ℓ enthalten
$\mathcal{P}^q_u$	Teilmenge der Wege aus $\mathcal{P}^q$, die den Knoten u enthalten
$x_{k,g}$	Ganzzahlige Variable, deren Wert die Anzahl der Komponenten vom Typ k angibt, die in Knoten oder Link g installiert werden
f^q_P	Ganzzahlige Variable, deren Wert die Anzahl der Basiseinheiten b_q des Bedarfes q angibt, die über den Weg P geroutet werden
$A_x(B,r)$	Gesamtangebot der Ressource r aller gemäß x installierten Komponenten auf einer Menge B von Knoten und Links
$V_x(B,r)$	Gesamtverbrauch der Ressource r aller gemäß x installierten Komponenten auf einer Menge B von Knoten und Links
C_ℓ	Verfügbare Routingkapazität auf Link ℓ, $C_\ell = A_x(\ell, \texttt{cap})$
A_ℓ	Kontinuierliche Variable, deren Wert die Auslastung von Link ℓ angibt
A	Kontinuierliche Variable, deren Wert die maximale Auslastung über alle Links angibt