

# Schichtenübergreifende Planung von Hochgeschwindigkeits- Telekommunikationsnetzen

Dissertation zur Erlangung des akademischen Grades eines  
Doktors der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.)  
des Departments Wirtschaftswissenschaften  
der Universität Hamburg

Vorgelegt von  
Dipl.-Wirtsch.-Ing. Holger Höller

Hamburg  
Januar 2008

1. Gutachter: Prof. Dr. Stefan Voß
2. Gutachter: Prof. Dr. Bernd E. Wolfinger

Datum der Disputation 4. Juni 2008

Hiermit erkläre ich an Eides Statt, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht.

Holger Höller  
Hamburg, im Januar 2008

# Inhaltsverzeichnis

Abbildungsverzeichnis	iv
Tabellenverzeichnis	vi
Symbolverzeichnis	vii
Abkürzungsverzeichnis	ix
<b>1 Einleitung</b>	<b>1</b>
1.1 Wirtschaftlicher Kontext und Entwicklung der Dienste . . . . .	2
1.2 Technologischer Kontext . . . . .	4
1.2.1 Leitungsvermittelnde Netze . . . . .	5
1.2.2 Paketvermittelnde Netze . . . . .	9
1.2.3 Konvergenz . . . . .	10
1.2.4 Schutzverfahren . . . . .	12
1.2.5 Netzschichten . . . . .	20
<b>2 Problemstellung</b>	<b>25</b>
<b>3 Modellierung</b>	<b>29</b>
3.1 Netzstruktur . . . . .	29
3.2 Geräte . . . . .	34
3.3 Bedarfe . . . . .	38
<b>4 Lösungsverfahren</b>	<b>40</b>
4.1 Exakte Lösungsverfahren . . . . .	40
4.1.1 Das ganzzahlige Basismodell . . . . .	43
4.1.2 Funktionale Erweiterungen . . . . .	47
4.1.3 Diversifizierte Lösungen . . . . .	58
4.1.4 Relaxationen . . . . .	62
4.1.5 Schnitte . . . . .	64
4.1.6 Praktischer Einsatz der ganzzahligen Modelle . . . . .	65

4.2	Heuristiken . . . . .	67
4.2.1	Greedy Randomized Adaptive Search Procedure . . . . .	68
4.2.2	Pilotmethode . . . . .	74
4.3	Zusammenspiel zwischen ganzzahligem Modell und Heuristiken	77
4.4	Ergänzende Literaturhinweise . . . . .	79
<b>5</b>	<b>Rechenergebnisse und Fallstudien</b>	<b>81</b>
5.1	Möglichkeiten und Grenzen der MIP-Modelle . . . . .	82
5.2	Leistungsfähigkeit der Heuristiken . . . . .	94
5.3	Fallstudien . . . . .	104
5.3.1	Fallstudie zum Einsatz von Carrier Ethernet . . . . .	105
5.3.2	Fallstudie zum Vergleich verschiedener Schutzverfahren	111
<b>6</b>	<b>Fazit</b>	<b>120</b>
<b>A</b>	<b>Programmdokumentation</b>	<b>122</b>
A.1	Programmstart und Systemanforderungen . . . . .	122
A.2	Die Bedienelemente . . . . .	123
A.3	Eingabedaten . . . . .	126
A.4	Ausgabedaten . . . . .	131
<b>B</b>	<b>Beispiel für ein AMPL-Modell</b>	<b>133</b>
	<b>Literaturverzeichnis</b>	<b>138</b>

# Abbildungsverzeichnis

1.1	Klassifizierung von Schutzmechanismen für Maschennetze. . . . .	14
1.2	1+1 Protection. . . . .	17
1.3	1:n Protection. . . . .	17
1.4	Das ISO/OSI-Referenzmodell. . . . .	21
1.5	Drei Beispiele für typische Schichtungen. . . . .	22
1.6	Beispiel für eine hierarchische Netzstruktur. . . . .	23
2.1	Planungsablauf. . . . .	26
2.2	Gestaltungsspielraum der Netzplanung. . . . .	27
3.1	Physikalische und logische Schichten eines Netzes. . . . .	30
3.2	Kantenabhängigkeit durch geographische Gegebenheiten. . . . .	31
3.3	Graphentransformation mittels eines Hilfsknotens. . . . .	31
3.4	Ein single point of failure. . . . .	33
3.5	Auflösen eines single point of failure. . . . .	33
3.6	Hilfsknoten bei Schutz über mehrere Teilnetze hinweg. . . . .	34
3.7	Grundlegende Komponenten des Gerätemodells. . . . .	35
3.8	Stufige Kostenfunktion und Glättung. . . . .	37
4.1	Grundsätzlicher Ablauf bei der Erstellung von Modellen für exakte Lösungsverfahren. . . . .	43
4.2	Beispiel für das Kreisproblem. . . . .	62
4.3	Grundsätzlicher Ablauf der Planung beim Einsatz von AMPL und CPLEX. . . . .	66
4.4	GRASP-Heuristik . . . . .	70
4.5	Glasfasergraph als Ergebnis einer Netzplanung. . . . .	74
4.6	Pilotsuche auf Basis des GRASP-Algorithmus. . . . .	76
4.7	Zusammenspiel der Komponenten. . . . .	78
5.1	Beispiel für die Auswirkung des Hoplimits auf die Laufzeit. . . . .	89
5.2	Einfluss der Gierigkeit der Lösungsstrategie. . . . .	102

5.3	Vorteil von GRASP gegenüber einem Verfahren ohne gesteu- erte lokale Suche. . . . .	103
5.4	Netzstruktur bei Szenario (I). . . . .	107
5.5	Netzstruktur bei den Szenarien (II) und (III). . . . .	107
5.6	Der Aufbau eines Netzknotens bei den Szenarien (II) und (III) im Detail. . . . .	108
5.7	Normierte szenariovariable Kosten. . . . .	111
5.8	Deutschlandweiter Backbone. . . . .	113
5.9	Regionalnetz. . . . .	114
5.10	Einfluss der Bedarfsgröße im Backbone. . . . .	117
5.11	Einfluss der Bedarfsspreizung im Backbone. . . . .	117
5.12	Einfluss der Bedarfsgröße in der Region. . . . .	118
5.13	Einfluss der logischen Vermaschung in der Region. . . . .	118
A.1	Bedienoberfläche. . . . .	123
A.2	Knotendaten mit Betriebsstellenschlüssel. . . . .	129
A.3	Glasfasertopologie. . . . .	130
A.4	Bedarfsmatrix. . . . .	131
A.5	Graphische Ausgabe. . . . .	132

# Tabellenverzeichnis

1.1	SDH: Bitraten der Virtual Containers. . . . .	6
1.2	SDH: Bitraten der Transportmodule. . . . .	7
5.1	Vergleich verschiedener Probleminstanzen mit 11 Knoten, 16 Kanten und 19 Bedarfen. . . . .	83
5.2	Vergleich der Laufzeiten für verschiedene Netzgrößen. . . . .	84
5.3	Vergleich der Laufzeiten zwischen 1+1 geschützten und un- geschützten Bedarfen. . . . .	85
5.4	Vergleich der Laufzeiten bei der Einführung verschiedener XC- Typen. . . . .	86
5.5	Vergleich verschiedener Probleminstanzen mit VC4-Grooming. . . . .	86
5.6	Laufzeiten für zwei SDH-Schichten und zwei XC-Typen. . . . .	87
5.7	Laufzeiten für Netze mit Restoration als Schutz. . . . .	87
5.8	Einfluss der Variablendeklaration von $f$ und $v$ . . . . .	90
5.9	Diversifizierung der Lösungen für eine kleine Probleminstanz. . . . .	90
5.10	Anzahl der besuchten Branch&Bound-Knoten. . . . .	92
5.11	Vergleich der Laufzeiten zwischen CPLEX 6 und CPLEX 8. . . . .	93
5.12	Vergleich der Lösungsgüte zwischen CPLEX 6 und CPLEX 8. . . . .	93
5.13	Beispiele für den Fortschritt durch das Pilotverfahren (GRASP). . . . .	97
5.14	Beispiele für den Fortschritt durch das Pilotverfahren (VNS). . . . .	98
5.15	Vergleich für kleine, künstlich generierte Netze. . . . .	98
5.16	Vergleich zwischen CPLEX und der Heuristik für mittelgroße Probleminstanzen. . . . .	99
5.17	Vergleich für große, reale Probleminstanzen. . . . .	100
5.18	Vergleich zwischen CPLEX und der Heuristik für 1+1 geschützte Probleminstanzen. . . . .	101

# Symbolverzeichnis

$b_{ij}$	Hilfsvariable, die den gesamten Restoration-Fluss über die Kante $(i, j)$ widerspiegelt
$C^C$	Kosten für einen WDM-Kanal (Transponderpaar)
$C_e^F$	Kosten einer Basisverbindung auf Kante $e$
$C_e^{F'}$	Kosten einer Verbindung erhöhter Geschwindigkeit auf Kante $e$
$C^O$	Kosten eines Cross-Connect Basisgerätes
$C^P$	Kosten für eine Portkarte (kurzer Reichweite) am Cross-Connect
$C^{P'}$	Kosten für eine Portkarte (erhöhte Geschwindigkeit, kurze Reichweite) am Cross-Connect
$C_e^W$	Kosten einer WDM-Strecke auf Kante $e$
$d_{st}$	Bedarf zwischen den Knoten $s$ und $t$
$d_{st}^l$	Bedarf niedriger Priorität zwischen den Knoten $s$ und $t$
$D$	Menge aller Bedarfe
$D^l$	Menge aller Bedarfe niedriger Priorität
$dis_{ij}^{st}$	Differenz der Routingvariablen von Bedarf $(s, t)$ auf Kante $(i, j)$ zwischen zwei Lösungen
$e$	Kante
$E$	Menge aller Kanten
$E'$	Menge aller Pfeile
$f_e$	Anzahl der Basisverbindungen (oder Gesamtkapazität) auf Kante $e$
$f'_e$	Anzahl der Basisverbindungen erhöhter Geschwindigkeit auf Kante $e$
$g_e$	Anzahl freier WDM-Kanäle auf Kante $e$
$h_n$	Anzahl freier Cross-Connect-Ports in Knoten $n$
$H$	Maximale Anzahl zulässiger Hops
$i, j, k, l$	laufende Variablen
$l_e$	Hilfsvariable für die Kapazität für Bedarfe niedriger Priorität und für Ersatzwege auf Kante $e$ bei 1:1 Protection
$M^O$	Kapazität eines Cross-Connects (Anzahl der Ports)
$M^W$	Kapazität eines WDM-Systems (Anzahl der Wellenlängen)
$n$	Knoten



$N$	Menge aller Knoten
$o_{ij}$	Hilfsvariable, die den gesamten primären Fluss über die Kante $(i, j)$ widerspiegelt (bei Restoration)
$r_{ij}^{kl}$	Restoration-Fluss über die Kanten $(k, l)$ , wenn die Kante $(i, j)$ ausfällt
$s$	Startknoten (source)
$t$	Zielknoten (target)
$t_n$	Transitflag für Knoten $n$
$v_e$	Anzahl der belegten Kanäle in den WDM-Systemen auf Kante $e$
$w_e$	Anzahl der WDM-Systeme auf Kante $e$
$y_n$	Anzahl der Cross-Connects in Knoten $n$
$z_{ij}^{st}$	Routingvariable von Bedarf $(s, t)$ auf Kante $(i, j)$
$za_{ij}^{st}$	Routingvariable für den Erstweg von Bedarf $(s, t)$ auf Kante $(i, j)$ bei 1:1 Protection
$zb_{ij}^{st}$	Routingvariable für den Ersatzweg von Bedarf $(s, t)$ auf Kante $(i, j)$ bei 1:1 Protection
$zl_{ij}^{st}$	Routingvariable für Bedarfe $(s, t)$ niedrigerer Priorität auf Kante $(i, j)$ bei 1:1 Protection
$zp_{ij}^{st}$	Routingvariable von Bedarf $(s, t)$ auf Kante $(i, j)$ in der Ausgangslösung (bei diversifizierten Lösungen)
$\omega$	Gewichtungsfaktor

# Abkürzungsverzeichnis

AMPL	A Mathematical Programming Language
ASCII	American Standard Code for Information Interchange
CWDM	Coarse Wavelength Division Multiplex
DSL	Digital Subscriber Line
DWDM	Dense Wavelength Division Multiplex
DXC	Digitaler Cross-Connect
EITO	European Information Technology Observatory
EXC	Elektrischer Cross-Connect
FCC	Federal Communications Commission
GAMS	General Algebraic Modeling System
GMPLS	Generalized Multi Protocol Label Switching
GRASP	Greedy Randomized Adaptive Search Procedure
IP	Internet Protocol
ISO	International Organization for Standardization
ITU	International Telecommunication Union
LCAS	Link Capacity Adjustment Scheme
LER	Label Edge Router
LSP	Label Switched Path
LSR	Label Switching Router
MAN	Metropolitan Area Network
MIP	Mixed Integer Programming
MPLS	Multi Protocol Label Switching
MP $\lambda$ S	Multi Protocol Lambda Switching
NP	Nichtdeterministisch Polynomiell
OADM	Optischer Add-Drop-Multiplexer
ODBC	Open Database Connectivity
OSI	Open System Interconnection
OTN	Optical Transport Network
OXC	Optical Cross-Connect
PXC	Photonischer Cross-Connect
RFC	Request For Comments

SDH	Synchronous Digital Hierarchy
SONET	Synchronous Optical Network
SRG	Shared Risk Group
STM	Synchronous Transfer Mode
UMTS	Universal Mobile Telecommunications System
VC	Virtual Container
VCAT	Virtual Concatenation
VLAN	Virtual Local Area Network
VNS	Variable Neighborhood Search
WDM	Wavelength Division Multiplex
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WWW	World Wide Web
XC	Cross-Connect

# Kapitel 1

## Einleitung

Zuverlässige und leistungsfähige Telekommunikationsdienstleistungen sind heutzutage aus vielen Bereichen des Lebens nicht mehr wegzudenken. Angefangen vom Telefon bis hin zu E-Mail und World Wide Web (WWW) werden sie täglich von vielen Millionen Menschen weltweit genutzt. Abgesehen von den Annehmlichkeiten der privaten Kommunikation sind gerade die Wirtschaftsunternehmen vieler Branchen in hohem Maße von einem reibungslosen Funktionieren der Telekommunikationssysteme abhängig. Dies betrifft nicht nur Unternehmen wie Amazon oder Google, deren gesamtes Geschäftsmodell auf der allgemeinen Verfügbarkeit von Internetdiensten basiert, sondern gerade auch die „klassische Industrie“. Auch wenn dies für den Endkunden oft auf den ersten Blick nicht erkennbar ist, so sind die einzelnen Unternehmen entlang der Lieferkette oft weitgehend elektronisch vernetzt.

Die stark steigende Nachfrage nach solchen Diensten führt zusammen mit dem technischen Fortschritt der Übertragungssysteme in vergleichsweise kurzen Zyklen zur Entstehung immer größerer und komplexerer Kommunikationsnetze. Die Planung dieser schnell wachsenden Netze stellt eine fortlaufende Herausforderung dar, die ab einer bestimmten Netzgröße nur noch sinnvoll mit Hilfe computergestützter Planungssysteme durchgeführt werden kann. Die Weiterentwicklung der mathematischen Verfahren, die solchen Planungssystemen zugrunde liegen, ist das Kernthema dieser Arbeit.

In der Einleitung wird zunächst der wirtschaftliche und technologische Kontext dieser Arbeit beschrieben. In diesem Rahmen werden auch die im Titel der Arbeit vorkommenden Begriffe „schichtenübergreifend“ und „Hochgeschwindigkeit“ definiert, da sie entscheidenden Einfluss auf die genaue Ausprägung der behandelten Problemstellung haben. Der eigentliche Kern der Arbeit (Kapitel 2, 3 und 4) beschäftigt sich dann mit der konkreten Ausprägung der Problemstellung, ihrer Modellierung und ihrer Lösung mittels

gemischt ganzzahliger Programmierung (MIP, mixed integer programming) und heuristischer Verfahren. Anschließend werden in Kapitel 5 eine Reihe von Rechenergebnissen und Fallstudien diskutiert. Während die Rechenergebnisse aus Berechnungen mit realen oder generierten Probleminstanzen gewonnen werden, um die Anwendbarkeit und Lösungsgüte der Verfahren zu überprüfen, gehen die Fallstudien noch einen Schritt weiter. Hier wird mit Hilfe der entwickelten Verfahren eine möglichst gute Lösung oder ein allgemeiner Erkenntnisgewinn für eine praktische Problemstellung gesucht. Abschließend werden in Kapitel 6 die gewonnenen Erkenntnisse zusammengefasst und ein Ausblick auf mögliche zukünftige, darauf aufbauende Forschungsgebiete gegeben.

Im Allgemeinen werden auch bei Fachausdrücken die deutschen Formulierungen gewählt. Einige englischsprachige Begriffe und Abkürzungen haben sich jedoch in einem Maße in der deutschsprachigen Literatur etabliert, dass eine Übersetzung, so denn eine hinreichend exakte überhaupt existiert, eher verwirrend als hilfreich wäre. Diese werden im Folgenden als englische Begriffe benutzt. Der in Abschnitt 3.1 eingeführte Begriff „single point of failure“ ist ein gutes Beispiel dafür.

## 1.1 Wirtschaftlicher Kontext und Entwicklung der Dienste

Der Telekommunikationsmarkt in Europa war in den 90er Jahren durch die beginnende Privatisierung der staatlichen Monopolanbieter und die Liberalisierung der Märkte geprägt. Mittlerweile gibt es in fast allen Bereichen eine Vielzahl von Anbietern, sowohl mit als auch ohne eigene Netzinfrastruktur. Trotzdem wird in vielen Ländern ein großer Teil der Infrastruktur von wenigen Unternehmen dominiert, das deutsche Festnetz sogar von einem einzigen, der Deutschen Telekom AG. In einigen Bereichen erfolgt deshalb eine staatliche Regulierung durch die Bundesnetzagentur.

Parallel dazu gab es in den 90er Jahren ein immenses Wachstum des Telekommunikationssektors. Viele der ehemaligen Monopolisten und einige der großen privaten Anbieter (letztere hauptsächlich im Mobilfunkbereich) versuchten auch international an Bedeutung zu gewinnen. Dies führte zu einer großen Zahl an freundlichen, aber auch feindlichen Unternehmensübernahmen. Aufgrund der Boomphase und der stark gestiegenen Aktienkurse waren viele dieser Übernahmen sehr teuer. Der Zusammenbruch der ehemals explodierten Aktienkurse im Jahr 2000 führte dann für viele der hoch verschuldeten Unternehmen zu schweren Problemen oder so-

gar Insolvenzen. Teilweise wurde dies auch durch schrumpfende Einnahmen, wie zum Beispiel durch eine zunehmende Sättigung des Mobilfunksektors in Verbindung mit hohen Lizenzgebühren und ein starkes Überangebot an Übertragungskapazität in internationalen Glasfasernetzen, verstärkt.

Bekanntere Beispiele für solche Insolvenzen sind MCI/Worldcom, eines der größten Telekommunikationsunternehmen der USA, Global Crossing, ein großer Betreiber internationaler Glasfaserleitungen, und KPNQwest, ein großer europäischer Betreiber optischer Langstreckennetze. Zusammen mit den Netzbetreibern sind auch die entsprechenden Gerätehersteller in Schwierigkeiten gekommen (z.B. Lucent, Nortel, Marconi und Siemens) und haben teilweise in großem Umfang Personal entlassen müssen.

In der darauf folgenden Konsolidierungsphase hat sich die Zahl der großen Unternehmen durch Fusionen weiter verringert. Zudem sind viele Investoren anderer Branchen aus dem Telekommunikationsbereich wieder ausgestiegen. Ein gutes Beispiel dafür ist das Engagement der Energieversorger in Deutschland, die zwischenzeitlich sowohl im Mobilfunk als auch im regionalen Festnetz sehr aktiv waren. Heute befindet sich die Telekommunikationsbranche wieder auf einem Wachstumspfad. Sie ist in Bezug auf Arbeitsplätze und Umsatz eine der größten Branchen weltweit.

Früher bestand der weitaus größte Teil der transportierten Daten aus Telefongesprächen. Etwa seit dem Jahr 2000 hat jedoch der Datenverkehr von Computersystemen vom Volumen her den klassischen Telefonverkehr überholt und ist weiterhin extrem angestiegen. In der Anfangsphase des Internet und der Datenfernübertragung waren Datendienste für Großkunden der entscheidende Faktor. Im Zuge des Digital Subscriber Line (DSL) Booms als schnellem Zugang zum Internet tragen die Privatkunden heute einen wesentlichen, wenn nicht inzwischen sogar den weitaus überwiegenden, Teil zum Gesamtvolumen des Datenverkehrs bei. Das aktuelle Stichwort dazu ist „Triple-Play“, also die Verbindung aus Telefon, Internet und Fernsehen über einen einzigen Anschluss, in Deutschland meistens DSL. Gerade die Tauglichkeit für technisch qualitativ hochwertiges Fernsehen erzwingt eine sehr leistungsfähige Infrastruktur, also Netze großer Kapazität. In diesem Zusammenhang wird auch der Ausbau der Glasfasernetze hin zum Kunden weiter vorangetrieben und die klassische Kupferdoppelader auf wenige letzte Meter zurückgedrängt.

Mobiltechnologien spielen vom Volumen her nach wie vor nur eine untergeordnete Rolle, abgesehen von der Überbrückung der letzten Meter vom Festnetzanschluss zum Computer. Diese wird heutzutage oft durch ein Wireless Local Area Network (WLAN) realisiert, sei es an einem öffentlichen Hotspot oder einem Privatanschluss. Gerade für anspruchsvolle Dienstbündel wie Triple Play sind die heute kommerziell verfügbaren, echten Mobiltechno-

logien, wie z.B. UMTS, einfach nicht leistungsfähig genug.

Für die Zukunft wird allgemein ein weiteres Wachstum der Datenkommunikation erwartet. Detaillierte Zahlen zur Entwicklung des Sektors Informations- und Kommunikationstechnologie in Europa finden sich beispielsweise in den Reports des European Information Technology Observatory (EITO), vgl. EITO (2005).

## 1.2 Technologischer Kontext

Neben der Nachfrage nach den Diensten stellen die technologischen Rahmenbedingungen einen weiteren zentralen Aspekt dar, der bei fast jeder Art von Netzplanung zu berücksichtigen ist. Um eine realistische Planung durchführen zu können, müssen zuerst die durch die zur Verfügung stehende Technologie gegebenen Randbedingungen analysiert werden. Hierzu ist ein prinzipielles Verständnis einiger technologischer Grundlagen notwendig. Von besonderer Bedeutung hierbei ist, dass es sich bei den betrachteten Netzen um „Hochgeschwindigkeitsnetze“ mit einer großen geographischen Ausdehnung handelt. Darunter sollen Netze verstanden werden, deren Übertragungsgeschwindigkeit in der Größenordnung von einem bis zu mehreren hundert Gigabit liegt und bei denen die Standorte, die zu vernetzen sind, mindestens einige Kilometer auseinander liegen. Für solche Netze gelten einige besondere Bedingungen, die im Folgenden aus den dort zum Einsatz kommenden Technologien heraus erklärt werden. Der Begriff der Kapazität einer Leitung oder eines Netzes ist synonym zu seiner Übertragungsgeschwindigkeit zu verstehen. Unter den Bedarfen, die durch ein solches Netz befriedigt werden, sollen grundsätzlich Verkehrsbeziehungen mit einer festen Datenrate zwischen zwei stationären Punkten verstanden werden.

Inzwischen spielen mobile Technologien auch für die schnelle Datenübertragung eine immer wichtigere Rolle. Universal Mobile Telecommunications System (UMTS), Wireless Local Area Network (WLAN), Worldwide Interoperability for Microwave Access (WiMAX) sind prominente Beispiele dafür. Letztendlich sind diese Technologien zwar für immer mehr Nutzer der primär sichtbare Teil des Telekommunikationsnetzes, insgesamt gesehen füllen sie jedoch nur die letzten Meter oder Kilometer zwischen dem Festnetz und dem Nutzer. Der größte Teil einer Kommunikation zwischen zwei entfernten Mobilgeräten läuft in der Regel über eine fest installierte, kabelgebundene Infrastruktur. Trotz oder zum Teil auch gerade wegen des zunehmend mobilen Datenverkehrs bilden also die dadurch ebenso wachsenden festen Netze weiterhin den größten Teil der Infrastruktur. Somit ist ihre Planung auch in Zeiten der Mobilkommunikation eine aktuelle, interessante und

stetig wachsende Aufgabe.

Unabhängig von der genauen Technologie ist die Glasfaser das heutzutage bei weitem dominierende Übertragungsmedium. Kupferkabel, Satelliten oder Richtfunk spielen bei großräumigen Hochgeschwindigkeitsnetzen nur eine untergeordnete Rolle. Ein prinzipieller Nachteil der Glasfaser ist jedoch der hohe Aufwand für die Einrichtung neuer Trassen. Ein Netzbetreiber wird deshalb meistens bestrebt sein, keine neuen Glasfasern zu verlegen, sondern die bestehende Infrastruktur so gut wie möglich auszunutzen. In der Tat ist das Glasfasernetz heutzutage in den meisten Industrienationen so gut ausgebaut, dass die Verlegung längerer neuer Strecken nur sehr selten notwendig ist. Das notwendige Kapazitätswachstum kann durch eine bessere Ausnutzung bestehender Fasern mit Hilfe weiterentwickelter Übertragungsverfahren realisiert werden. Die verfügbare Glasfaserinfrastruktur kann also auch im Rahmen mittel- bis langfristiger Planung als nahezu konstant angenommen werden. Etwas anderes ist dies im Zugangsnetz, welches die letzten Kilometer bis in die Haushalte der Privatkunden umfasst. Hier werden zur Zeit in großem Umfang alte Kupferleitungen um Glasfasern ergänzt. Diese meistens stern- oder baumförmig angelegten Netze, oft als „letzte Meile“ bezeichnet, sind aber nicht Betrachtungsgegenstand dieser Arbeit.

### 1.2.1 Leitungsvermittelnde Netze

Das Prinzip der Leitungsvermittlung findet sich in vielen verschiedenen Arten von Kommunikationsnetzen. Es war für lange Zeit, und ist wahrscheinlich zumindest im Hochgeschwindigkeitsbereich noch immer, die vorherrschende Technologie. In leitungsvermittelnden Netzen wird eine physikalische Ende-zu-Ende-Verbindung hergestellt, bevor mit der eigentlichen Datenübertragung begonnen wird. Diese Verbindung bleibt dann bis zum Ende der Übertragung unverändert bestehen, sofern nicht Störungen ein Eingreifen erforderlich machen. Selbst wenn zeitweise keine Nutzdaten zu übertragen sind, bleibt die Verbindung aktiv und die entsprechende Kapazität reserviert. Ein solches Verfahren hat diverse Vorteile. Neben der bereits erwähnten zugesicherten Kapazität sind auch die Schalteinrichtungen zum Verbinden der Leitungen relativ einfach aufzubauen. Prinzipiell reicht eine einfache quadratische Schaltmatrix, in der jede Eingangsleitung mit jeder Ausgangsleitung verbunden werden kann, obwohl es natürlich effizientere Möglichkeiten der technischen Realisierung gibt. Ein gutes Beispiel für ein solches leitungsvermittelndes Netz ist das herkömmliche Telefonnetz. Mit dem Wählen der Nummer wird vor dem Gespräch eine feste Verbindung zwischen den Teilnehmern aufgebaut. Unabhängig davon, ob gerade gesprochen wird oder nicht,



belegt diese Verbindung eine konstante Kapazität im Netz, die in dieser Zeit von keinem anderen Gespräch genutzt werden kann.

Die meisten schnellen Langstreckennetze basieren auf einer solchen leitungsvermittelnden Technologie, der Synchronen Digitalen Hierarchie (SDH) bzw. dem amerikanischen Äquivalent Synchronous Optical Network (SONET), häufig in Kombination mit Dense Wavelength Division Multiplex (DWDM). SDH erlaubt eine (geschützte) Übertragung und Verschaltung von standardisierten Dateneinheiten im Größenbereich von 2 Mbit/s bis 40 Gbit/s. Die Übertragung der Signale findet optisch statt, zumindest für die heute gebräuchlichen Übertragungsraten von 622 Mbit/s an aufwärts, die Verarbeitung in den Netzknoten erfolgt jedoch durch elektrische Schaltkreise. Folglich muss an jedem Ein- und Ausgang eine elektrisch-optische Wandlung stattfinden. Ein solches Schaltsystem wird Cross-Connect genannt. SDH ist strikt bidirektional, d.h. für jeden Datenstrom von einem Punkt  $A$  zu einem anderen Punkt  $B$  existiert immer ein entgegengesetzter Datenstrom gleicher Größe von  $B$  nach  $A$ . Wie schon der Name deutlich macht, sind alle Datenflüsse in einem SDH-Netz synchron, d.h. es gibt einen zentralen Takt für die gesamte Kommunikation.

Das Basiselement jeder SDH-Datenübertragung ist der sogenannte Virtual Container (VC). Er besteht aus der eigentlichen Nutzlast und einem Protokollkopf. Tabelle 1.1 führt die gebräuchlichsten Containertypen und ihre jeweilige Nutzlast auf. Der VC 12 dient im Wesentlichen der Übertragung von Telefongesprächen und kleineren Datenanschlüssen. Der VC 4 ist die Basiseinheit für die meisten Übertragungen und kann auch für höhere Datenraten in fest gekoppelten Gruppen verwendet werden (Contiguous Concatenation). So ist z.B. ein VC 4-4c eine Verknüpfung von vier VC 4. Solche verknüpften Container können jedoch in den Cross-Connects nur als Ganzes verarbeitet werden; ein Zugriff auf einzelne VC 4 ist nicht möglich. Im Rahmen der hier betrachteten Hochgeschwindigkeitsnetze sind nur die Containergrößen von VC 4 an aufwärts von Interesse.

Typ	max. Nutzlast in Mbit/s
VC 12	2,304
VC 3	49,536
VC 4	150,336
VC 4-4c	601,344
VC 4-16c	2.405,376

Tabelle 1.1: SDH: Bitraten der Virtual Containers.

Für die eigentliche Übertragung über die Glasfaser werden diese virtuellen Container in sogenannte Synchronous Transport Modules (STM) eingebettet. Transportmodule gibt es in verschiedenen Bandbreiten, für die hier betrachteten Netze sind jedoch im wesentlichen nur die Größen 2,5 Gbit/s und 10 Gbit/s relevant. Auch wenn bereits ein Modul für noch höhere Datenraten spezifiziert ist (STM-256, 40 Gbit/s), finden sich solche Systeme bisher in der Praxis nur sehr selten. Tabelle 1.2 gibt einen Überblick über die standardisierten Modultypen und ihre Datenrate.

Typ	Bitrate in Mbit/s
STM-1	155,52
STM-4	622,08
STM-16	2.488,32
STM-64	9.953,28
STM-256	39.813,12

Tabelle 1.2: SDH: Bitraten der Transportmodule.

In der Literatur existieren verschiedene Abkürzungen für den Begriff Cross-Connect, die sehr uneinheitlich benutzt werden. Rein der Bedeutung der Worte nach könnten die beiden am häufigsten benutzten Begriffe „Elektrischer Cross-Connect (EXC)“ als ein System zum ausschließlich elektrischen Verarbeiten der Signale (Verschalten von Kupferleitungen mit Hilfe von elektrischen Schaltkreisen) und „Optischer Cross-Connect (OXC)“ als ein System zum ausschließlich optischen Verarbeiten der Signale (Verschalten von Wellenlängen in Glasfasern z.B. über bewegliche Spiegel) angesehen werden. Die Mehrzahl der aktuell im Betrieb befindlichen Cross-Connects arbeitet jedoch auf den Übertragungsleitungen optisch und für die Verarbeitung werden die Signale in elektrische umgewandelt. Diese Typen werden deshalb in der Literatur sowohl als EXC als auch als OXC bezeichnet. Für rein elektrische Systeme existiert daneben auch die Bezeichnung „Digitaler Cross-Connect (DXC)“ und für rein optische Systeme „Photonischer Cross-Connect (PXC)“. Um die verwirrende und uneinheitliche Nutzung dieser Begriffe zu vermeiden, wird in dieser Arbeit der Begriff Cross-Connect (XC) als Oberbegriff eingeführt und ausschließlich verwendet. Auf technologische Details wird nur eingegangen, wenn diese im Kontext eine spezielle Relevanz besitzen.

DWDM ist eine Technologie, die es ermöglicht, mittels Wellenlängenmultiplex viele verschiedene Quellsignale unabhängig voneinander auf einer Glasfaser zu transportieren. Hierzu wird jedes einzelne Quellsignal

in eine andere Wellenlänge (anschaulich gesprochen also in ein Licht unterschiedlicher Farbe) transformiert und dann werden alle Signale auf eine einzige Glasfaser gebündelt (gemultiplext). Dies ermöglicht eine um den Faktor  $n$  höhere Übertragungsgeschwindigkeit pro Glasfaser, wenn  $n$  die Anzahl an Signalen ist, die gemultiplext werden können. Gebräuchliche Systeme bündeln beispielsweise  $80 \times 2,5$  Gbit/s oder  $40 \times 10$  Gbit/s, allerdings sind auch Multiplexfaktoren bis zu 160 Kanälen oder sogar darüber hinaus möglich. Herkömmliche WDM-Systeme stellen jedoch immer nur eine Punkt-zu-Punkt-Verbindung als physikalischen Kanal zwischen Netzknoten anderer Technologien (z.B. SDH) her. Diese Verbindung ist darüber hinaus immer bidirektional und symmetrisch. Traditionell kommen WDM-Systeme aus der Langstreckenübertragung, beispielsweise basieren alle modernen Seekabel auf dieser Technologie. Im Zuge des wachsenden Bandbreitenbedarfes bei regionalen Netzen dringt WDM jedoch verstärkt auch in den Bereich der Metropolitan Area Networks (MAN) vor. In diesem Zusammenhang hat sich ein neuer Typ etabliert, das sogenannte Coarse Wavelength Division Multiplex (CWDM). Diese Systeme arbeiten mit deutlich weniger Wellenlängen und können dadurch kostengünstiger hergestellt werden. Die genaue Differenzierung zwischen WDM, DWDM und CWDM ist im Rahmen dieser Betrachtung jedoch nicht relevant, der Begriff WDM wird im Folgenden synonym für die gesamte Gruppe der Wellenlängenmultiplex-Technologien verwendet.

Aufbauend auf diesen etablierten Technologien kann das Optical Transport Network (OTN) als der nächste Schritt auf dem Weg von heutigen SDH/SONET-Netzen, die ihren Ursprung noch im elektrischen Verschalten der Signale haben, zum vollständig optischen Netz gesehen werden. OTN ist ein optisches Transportnetz, welches WDM als integralen Bestandteil hat und mit einer kleinsten Datenrate von 2,5 Gbit/s pro Kanal für den Kern zukünftiger Hochgeschwindigkeitsnetze ausgelegt ist.

Technische Details zu SDH werden z.B. in Lochmann (1995), Wilde (1999) oder Kiefer (1996) erklärt. Sexton und Reid (1997) und Wu (1992) sind zwei Standardwerke über SONET. Informationen über DWDM und OTN finden sich unter anderem in Krauss (2002) bzw. speziell zu WDM in Zang (2003). Die offiziellen ITU (International Telecommunication Union) Standards zu SDH und OTN, u.a. G.707, G.709 und G.872, sind auf der Homepage der ITU abrufbar (<http://www.itu.org>). Daneben decken natürlich diverse umfangreiche Lehrbücher diese Gebiete ab, so z.B. Lee und Kim (2002) und Ramaswami und Sivarajan (2002).

## 1.2.2 Paketvermittelnde Netze

Während die Langstreckennetze durch den ehemals dominierenden Telefonverkehr traditionell eher verbindungsorientiert arbeiten, haben im Bereich der Local Area Networks (LAN), vor allem innerhalb von Unternehmen und staatlichen Institutionen, und des Internet schon lange paketbasierte Netze zur Datenübertragung zwischen Computersystemen dominiert.

Als Paket wird eine kleine Einheit aus Nutzdaten und einem sogenannten Header bezeichnet. Der Header enthält Informationen wie Zieladresse, Zeitstempel usw. Pakete werden einzeln in das Netz eingespeist und von den Netzknoten auch individuell behandelt. Somit können unterschiedliche Pakete zwischen denselben Knoten ganz andere Wege durch das Netz nehmen. Damit ist eine paketbasierte Übertragung per Definition unidirektional und auf physikalischer Ebene auch verbindungslos. Natürlich kann eine bidirektionale Kommunikation mit Hilfe von zwei unidirektionalen Paketströmen aufgebaut werden, die Wege dieser Ströme sind jedoch im Prinzip voneinander unabhängig. Ebenso können auf logischer Ebene auch so genannte virtuelle Verbindungen aufgebaut werden. Klassisch arbeiten Paketnetze oft dezentral und selbstorganisierend.

Das mit Abstand dominierende Protokoll sowohl in lokalen Netzen als auch im Internet ist das Internet Protocol (IP). Die Geräte, die in den Netzknoten die besten Pfade für die Pakete ermitteln und sie dann entsprechend weiterleiten, werden IP-Router genannt. Sie sind in etwa das paketbasierte Gegenstück zu den SDH-Cross-Connects. Als eigentliche Transporttechnologie zwischen den Routern dient in lokalen Netzen meistens Ethernet.

Die Übertragungsgeschwindigkeiten bei Ethernet sind traditionell in der Regel um den Faktor Zehn gestiegen (10 Mbit/s, 100 Mbit/s, 1 Gbit/s, 10 Gbit/s). Bei höheren Geschwindigkeiten werden die Daten meistens optisch übertragen, bei niedrigeren über Kupferkabel. Damit entspricht die schnellste Ethernet-Übertragung mit 10 Gbit/s von den Rahmenbedingungen her gängigen SDH- bzw. OTN-Übertragungsgeschwindigkeiten (bei OTN bezogen auf eine Wellenlänge). Dies erleichtert prinzipiell das Zusammenspiel zwischen den Technologien. Allerdings treten hier in der Praxis dann doch wieder Probleme auf, weil die Datenraten von SDH und Ethernet bei der 10 Gbit/s Übertragung nicht ganz identisch sind, bzw. es zwei leicht unterschiedliche Ethernet Datenraten gibt, von denen eine so ausgelegt ist, dass sie genau über eine 10 Gbit/s SDH-Verbindung transportiert werden kann, die historisch ältere jedoch etwas zu groß dafür ist.

Für lange Strecken führt ohnehin kaum ein Weg an WDM vorbei. Auf kürzeren Strecken bzw. als Nutzsignal der WDM-Systeme tritt Ethernet jedoch auch bei großen Netzbetreibern zunehmend in Konkurrenz zur etablier-

ten SDH-Technologie. Geräte, welche die entsprechenden erweiterten Funktionen, vor allem aus den Bereichen Netzmanagement und Netzüberwachung, beherrschen, werden oft als Carrier Ethernet bezeichnet.

Gatrell *et al.* (2001) oder allgemeine Lehrbücher wie Lee und Kim (2002) bieten eine umfassende Einführung in das IP-Protokoll und damit zusammenhängende Bereiche. Der ursprüngliche Standard dazu ist im RFC 760 (Poestel (1994)) beschrieben, er wurde mittlerweile jedoch vielfach aktualisiert und ergänzt. Unter <http://www.ietf.org/> sind alle RFCs abrufbar.

### 1.2.3 Konvergenz

Ein wesentlicher Antrieb für die zunehmende Konvergenz ist die Tatsache, dass die traditionellen Übertragungssysteme für lange Strecken leitungsvermittelnd arbeiten, während der Datenverkehr, der diese Netze nutzt, inzwischen primär paketvermittelnd ist. Dazu kommt, dass die aus dem paketvermittelnden Bereich stammenden Technologien wie Ethernet auf Grund der enormen Stückzahlen inzwischen sehr kostengünstig erhältlich sind und mit neuen Funktionen auch in den Bereich der Netzbetreiber vordringen.

Trotz der oben beschriebenen fundamentalen Unterschiede zwischen paketvermittelnder und leitungsvermittelnder Übertragung gibt es folglich immer mehr Bestrebungen, beide Methoden zu kombinieren. Auf der einen Seite verfügen moderne SDH-Geräte über Funktionen, um mit paketbasiertem Verkehr umgehen zu können (Next Generation SDH). Auf der anderen Seite verfügen moderne paketbasierte Technologien über Mechanismen für virtuelle, geschaltete Verbindungen, z.B. Multiprotocol Label Switching (MPLS) und dessen Verallgemeinerung Generalized Multiprotocol Label Switching (GMPLS).

Um paketbasierte Nutzdaten besser handhaben zu können, sind eine Reihe von Erweiterungen zum ursprünglichen SDH-Standard entwickelt worden. Während SDH normalerweise einen kontinuierlichen Datenstrom erwartet, tritt paketbasierter Verkehr meistens in Bursts auf. Die Generic Framing Procedure (GFP) übernimmt die Aufgabe, verschiedenste Arten von Nutzdaten in SDH-Container einzubetten, bei Bursts werden die Lücken ggf. mit Füllbytes geschlossen. GFP arbeitet ähnlich wie der Digital Wrapper des OTN.

Weiterhin besteht das Problem, dass gerade den weit verbreiteten Ethernet-Datenraten von 100 Mbit/s und 1 Gbit/s keine entsprechend großen SDH-Container zur Verfügung stehen. Beispielsweise müsste für eine 1 Gbit/s Nutzlast ein kompletter VC4-16c vorgesehen werden (vgl. auch Tabelle 1.1). Damit würde aber ein Großteil der nominellen Kapazität von 2,5 Gbit/s

ungenutzt bleiben. Um dieses Problem zu umgehen, wurde über die sogenannte Virtual Concatenation (VCAT) die Möglichkeit geschaffen, die SDH-Basiscontainer (relevant ist hier primär der VC4) zu virtuellen Einheiten zusammenzufassen. Die Nutzlast wird also nicht in den nächstgrößeren passenden Containertyp eingebettet, sondern in mehrere Basiscontainer, die nur eine virtuelle Einheit bilden. Ein 1 Gbit/s Ethernet-Signal kann somit in sieben VC4 verpackt und als virtuelle Einheit VC4-7v transportiert werden (im Gegensatz zu den 16 untrennbar verknüpften VC4 bei Contiguous Concatenation, VC4-16c).

Virtual Concatenation erlaubt zusammen mit dem Link Capacity Adjustment Scheme (LCAS) zusätzlich eine dynamische Anpassung der Bandbreite. Während der Übertragung können nach Bedarf zusätzliche Container hinzugenommen oder entfernt werden. So kann beispielsweise aus einer VC4-7v-Verbindung bei nachlassender Bandbreiteneanforderung ohne Unterbrechung eine VC4-6v-Verbindung werden. Die frei gewordene Kapazität im Netz in Größe eines VC4 kann dann anderen Verbindungen zugeordnet werden.

Multiprotocol Label Switching wurde mit dem Ziel entwickelt, verschiedene Verbesserungen in paketbasierten Netzwerken einzuführen, primär eine Reduzierung der Routerlast und damit Erhöhung der Leistungsfähigkeit und erweitertes Traffic Engineering durch festgelegte Pfade, sogenannte Label Switched Paths (LSP). Im Prinzip kann MPLS als eine Art virtuelle Leitung innerhalb eines Paketnetzes angesehen werden.

Jedes Paket enthält ein Label. Anhand dieses Labels wird es dann in den Netzknoten auf einem festgelegten Pfad weitergeleitet, ohne dass der Router erst aufwändig den Layer 3 Header auswerten muss. LSPs können auch ineinander geschachtelt werden, somit kann ein LSP großer Bandbreite mehrere kleine enthalten. Router mit dieser Fähigkeit werden Label Switching Router (LSR) genannt. MPLS ermöglicht es also, festgelegte Pfade (Tunnel) durch das Netzwerk zu definieren. Dies kann beispielsweise genutzt werden, um virtuelle Netze aufzubauen, Quality of Service (QoS)-Garantien einzuhalten oder Ersatzwege bei Störungen vorzudefinieren. Somit kann mittels MPLS ein Traffic Engineering durchgeführt werden, welches weit über die Möglichkeiten der traditionellen Routingprotokolle und ihrer auf Link-Metriken basierenden Verfahren hinaus geht.

Aus der Idee, optische Transportnetze in MPLS mit einzubeziehen, ist MPAS entstanden, welches nach einiger Entwicklungszeit inzwischen allgemein als GMPLS bezeichnet wird. Vereinfacht ausgedrückt ermöglicht GMPLS das Management von Pfaden in Gesamtsystemen, die sowohl aus paket- als auch aus leitungsvermittelnden Subsystemen bestehen können, beispielsweise eine Kombination aus einem IP-Router-Netz und einem SDH/OTN/WDM-Netz. Es ermöglicht also eine technologieübergreifende

Netzsteuerung.

Der MPLS-Standard ist im RFC 3031 dargelegt (Rosen *et al.* (2001)), GMPLS wird in RFC 3471 (Berger (2003)) beschrieben. Zhu *et al.* (2003) und Cisco Systems (2003) geben einen Überblick über die wesentlichen Neuerungen bei Next Generation SDH.

Diese zunehmende Konvergenz der Technologien stößt mitunter auch auf praktische Hürden in der Organisationsstruktur der Netzbetreiber. Traditionell waren die Bereiche „Transportnetz“ mit den Wurzeln im klassischen Telefonnetz und „IP-Netze“ häufig getrennt. Jeder Bereich hat mehr oder weniger autark sein „eigenes“ Netz geplant. Bei einer zunehmenden Verknüpfung der tatsächlichen Netzstrukturen führt diese getrennte Planung möglicherweise zu einem suboptimalen Gesamtnetz. Für eine integrierte Planung muss in beiden Bereichen zumindest ein entsprechender grundlegender Sachverstand für die jeweils andere Technologie aufgebaut werden. Abgesehen davon kann aber auch der Fall auftreten, dass mit der teilweisen Substitution einer Technologie durch eine andere sich auch Kompetenzen und Personalbedarfe zwischen den Teilbereichen verschieben. Neben dem akademischen Problem einer optimalen, übergreifenden Planung gibt es also in der Praxis eine Reihe von möglichen Problemen, die ebenfalls beachtet werden müssen, die aber letztendlich mit entscheidend sind, ob der theoretische Erkenntnisgewinn auch praktischen Nutzen abwerfen kann.

#### 1.2.4 Schutzverfahren

Die Gründe, warum eine Datenübertragung ausfällt, können vielfältig sein, z.B. zerstörte Glasfasern durch Bauarbeiten, der Ausfall eines Lasers oder der Absturz einer Software in einem Router. Bei vielen Bedarfen ist es für den Kunden jedoch sehr wichtig, dass diese gegen solche Ausfälle geschützt werden. Auch abseits der offensichtlich vollständig vom Internet abhängigen Firmen, wie z.B. Amazon oder E-Bay, ist eine zuverlässige Datenübertragung für sehr viele Unternehmen mittlerweile essentiell. Aber auch unabhängig von Kunden, die explizite Zusicherungen über eine hohe Verfügbarkeit verlangen und ggf. auch bereit sind, dafür einen entsprechenden Aufpreis zu bezahlen, gibt es starke Anreize für einen Netzbetreiber, eine möglichst hohe Zuverlässigkeit zu erreichen. Die Auswirkungen von Ausfällen, die eine große Zahl an Kunden gleichzeitig betreffen, auf das Image des Netzbetreibers können verheerend sein, selbst wenn er trotzdem die vertraglich zugesicherte Verfügbarkeit erfüllt. Und mit der enormen Datenmenge, die eine einzelne Glasfaser mittlerweile führen kann (400 Gbit/s pro Faser sind bereits weit verbreitet), steigen entsprechend auch die Auswirkungen eines Ausfalls

dieser Faser. Damit werden zuverlässige Schutzverfahren für die Datenverbindungen unumgänglich. In diesem Abschnitt werden einige der gängigsten Verfahren vorgestellt. Dabei geht es nicht um den Schutz spezieller Hardwarekomponenten gegen eventuelle Ausfälle sondern um das Aufrechterhalten der Datenverbindung als Ganzes.

Es gibt verschiedene Möglichkeiten, die Widerstandsfähigkeit des Netzes gegen Fehlerquellen zu stärken. Technische Maßnahmen wie gedoppelte Stromversorgungen oder redundante Schaltmatrizen sind natürlich ein wichtiger Aspekt. Aber alleine mit solchen Maßnahmen kann ein Netz nicht gesichert werden. Vor allem zerstörten Glasfasern oder größeren Ausfällen von Geräten durch Feuer o.ä. kann damit nicht sinnvoll entgegengewirkt werden. Aber auch in solchen Fällen sollte der Betrieb des Netzes aufrecht erhalten werden können, bis das Servicepersonal dazu in der Lage ist, den Fehler zu beheben. Im Allgemeinen werden kommerzielle Netze nur gegen Einzelfehler geschützt. Das Risiko des Auftretens eines weiteren Fehlers, bevor der erste behoben werden konnte, wird normalerweise als gering eingeschätzt. Lediglich bei langen Übertragungstrecken in Gegenden mit wenig vorhandener Infrastruktur, z.B. in einigen ländlichen Gegenden Chinas, wird teilweise ein Schutz gegen zwei simultane Fehler vorgesehen, weil die durchschnittlichen Reparaturzeiten sehr lang sein können. Daneben gibt es natürlich im militärischen Bereich Anforderungen für einen Schutz gegen eine größere Zahl gleichzeitig auftretender Fehler. Diese Sonderfälle sollen hier jedoch keine Rolle spielen.

Grundvoraussetzung für den Schutz einer Übertragungstrecke zwischen zwei Punkten ist, dass überhaupt ein zweiter Übertragungsweg existiert, der vollständig vom ersten unabhängig ist. Zwei Pfade werden kantendisjunkt genannt, wenn sie keine gemeinsame Kante nutzen, und knotendisjunkt, wenn sie keinen gemeinsamen Knoten nutzen. Dabei impliziert knotendisjunkt immer auch kantendisjunkt, aber nicht umgekehrt. Diese theoretische Bedingung ist in der Praxis leider oft nicht mehr so eindeutig wiederzufinden. Selbst logisch unabhängige Glasfasern können physikalisch voneinander abhängig werden, wenn sie zufällig im gleichen Kabelkanal verlaufen, z.B. über eine Brücke. Hinzu kommt teilweise das Problem, dass bei großen, historisch gewachsenen Netzen möglicherweise gar nicht alle Daten über solche Abhängigkeiten verfügbar sind. Auf das Problem solcher voneinander abhängiger Übertragungstrecken und die Auswirkungen auf die Netzplanung wird in Abschnitt 3 näher eingegangen.

Schutzverfahren, besonders in leitungsvermittelnden Netzen, können in zwei große Gruppen unterteilt werden: Protection und Restoration. Abbildung 1.1 zeigt eine Klassifikation der gängigsten Protection- und Restorationverfahren für Maschennetze. In den folgenden Abschnitten werden diese



im Detail erläutert. Gerade im Grenzbereich zwischen Protection und Restoration ist die Nomenklatur in der Literatur sehr unterschiedlich. In dieser Arbeit wird der Begriff Restoration benutzt, wenn die Ressourcen des Ersatzweges erst im Fehlerfall reserviert werden oder a priori keinem Erstweg exklusiv zur Verfügung stehen. Ob ein Ersatzweg schon vorab oder erst im Fehlerfall für einen Erstweg berechnet wird, ist dabei unerheblich.

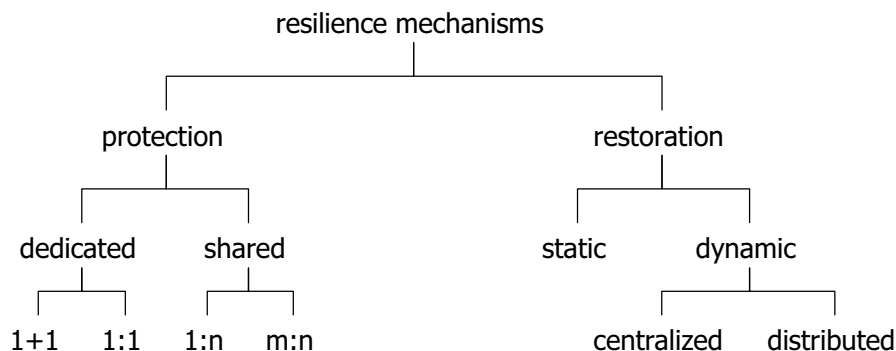


Abbildung 1.1: Klassifizierung von Schutzmechanismen für Maschennetze.

Die Schutzverfahren in paketvermittelnden Netzen lassen sich prinzipiell in die gleiche Klassifizierung einordnen. Allerdings ist die gängige Nomenklatur hier eine andere, sie wird im IP-Kontext vom Begriff „Routing“ dominiert. Das klassische IP-Routing entspricht im Fehlerfall in etwa einem dynamischen Restorationverfahren, während es in MPLS-Netzen zusätzlich auch Protection ähnliche Mechanismen gibt. In der Regel wird im Folgenden die SDH-Nomenklatur gebraucht und dann jeweils kurz auf die entsprechenden IP/MPLS-Äquivalente hingewiesen.

Ein wichtiger Aspekt beim Schutz mehrschichtiger Netze ist die Koordination der einzelnen Schichten. Einerseits muss gewährleistet sein, dass Ausfälle überbrückt werden, unabhängig davon auf welcher Schicht ein Gerät ausgefallen ist. Andererseits dürfen die Schutzmechanismen für die Schichten aber nicht miteinander konkurrieren. Neben einer Verschwendung von Kapazität könnte es auch zu einem instabilen Netz führen, wenn auf unterschiedlichen Ebenen gleichzeitig versucht wird, denselben Fehler zu beheben.

Angenommen bei einer IP-über-SDH-Übertragung würde die optische Leitung ausfallen. Die SDH-Schicht bemerkt diesen Ausfall und setzt ihr Schutzverfahren in Gang. Andererseits bemerkt die IP-Schicht ebenfalls, dass die Verbindung ausgefallen ist, und startet ihr eigenes Schutzverfahren. Während jetzt die normalerweise langsamer reagierende IP-Schicht noch mitten im Prozess der Umwegführung ist, hat die SDH Schicht die Verbindung bereits repariert. Die IP-Schicht bemerkt diese Reparatur und fängt nun an,

die eigenen Wege wieder entsprechend zurück zu konfigurieren. Ein solcher doppelter Schutz ist somit eher schädlich als sinnvoll. Eine einfache Lösung dafür wäre, dass höher gelegene Schichten (in diesem Beispiel IP) erst mit einer gewissen voreingestellten Verzögerung auf Ausfälle reagieren. Wenn diese Verzögerung so eingestellt wird, dass sie für die untere Schicht ausreicht, um den Fehler zu beheben, müsste die höhere Schicht gar nicht mehr aktiv werden. Wenn die untere Schicht aus irgendwelchen Gründen nicht in der Lage ist, den Fehler zu beheben, würde die höhere Schicht nach kurzer Verzögerung jedoch selber aktiv werden. Technisch aufwändigere, aber damit meistens auch effizientere Verfahren basieren auf einer schichtenübergreifenden, gezielten Steuerung der Reaktionen im Fehlerfall. Dies setzt jedoch eine zentralisierte Kenntnis über den Zustand aller dafür relevanten Schichten im Netz voraus.

### Protection

Grundsätzlich zeichnen sich Protection-Verfahren dadurch aus, dass die Ersatzkapazität im Voraus reserviert und explizit zu schützender Arbeitskapazität zugeordnet wird. Damit ist im Fehlerfall eine sehr schnelle Reaktion möglich; 50ms Umschaltzeit auf den Ersatzweg hat sich zu einer Art Industriestandard entwickelt. Normalerweise erfolgt die Reaktion auf Fehler direkt durch die Netzknoten, ohne dass ein zentrales Netzmanagement eingreifen müsste. Protection kann entweder Ende-zu-Ende für den ganzen Weg eines Bedarfes (Path Protection) oder in Bezug auf einzelne Kanten (Link Protection) implementiert werden.

Neben dieser Unterscheidung gibt es noch eine Reihe weiterer Kriterien, um Protection-Verfahren zu klassifizieren. Ein mögliches Kriterium ist die Unterscheidung zwischen exklusiv einem Bedarf zugeordneter (Dedicated Protection) und potentiell mehreren Bedarfen gemeinsam zur Verfügung stehender (Shared Protection) Reservekapazität. Die weit verbreitete 1+1 Path Protection ist ein Dedicated Protection-Verfahren. Hierbei wird zwischen Quelle und Senke ein zweiter, disjunkter Pfad geschaltet und für die Übertragung eines identischen Signals genutzt. An der Senke werden die beiden ankommenden Signale überwacht und das mit der besten Signalqualität ausgewählt. Sollte einer der beiden Wege ausfallen, so kann, fast ohne Verzögerung und von der Senke autonom, auf den anderen umgeschaltet werden, vgl. Abbildung 1.2. Somit ermöglicht 1+1 Path Protection eine hohe Verfügbarkeit, allerdings auf Kosten eines hohen Ressourcenverbrauchs (doppelte Übertragung). 1+1 Path Protection ist das traditionell vorherrschende Schutzverfahren in den meisten Hochgeschwindigkeitsnetzen mit Maschenstruktur, die auf SDH- und WDM-Technologie basieren. Damit ist eine Un-

terstützung von 1+1 Path Protection in allen Planungsverfahren dieser Arbeit selbstverständlich.

Eine etwas modifizierte Variante von 1+1 Path Protection stellt das sogenannte 1:1 Path Protection dar. Hier können die Ersatzwege von Verkehr niederer Priorität belegt werden, solange kein Fehler auftritt. Wenn der Bedarf hoher Priorität auf seinen Ersatzweg umgeschaltet werden muss, dann wird der dort fließende Verkehr niederer Priorität unterbrochen. Damit wird eine bessere Ausnutzung der Ressourcen auf Kosten einer leicht erhöhten Umschaltzeit im Fehlerfall ermöglicht. Natürlich macht dieses Verfahren nur Sinn, wenn es ausreichend Verkehr niederer Priorität gibt, um die brachliegenden Reservekapazitäten entsprechend nutzen zu können. Dieses Verfahren wird in Abschnitt 4.1.2 wieder aufgegriffen und als ganzzahliges Modell formuliert. Das Verfahren fällt in die Kategorie „Dedicated Protection“, weil jeder Erstweg einen für ihn vorbestimmten Ersatzweg besitzt. Die Kapazität der Ersatzwege kann, solange kein Fehler auftritt, dabei von beliebigen Bedarfen niederer Priorität, auch abschnittsweise, genutzt werden. Eine feste Zuordnung zwischen den Bedarfen hoher und niedriger Priorität bezogen auf ein Knotenpaar existiert nicht.

In der Praxis sind dafür verschiedene Anwendungsszenarien denkbar. Netzbetreiber könnten die normalerweise brachliegenden Reservekapazitäten an Kunden verkaufen, deren Datenverkehr keine hohe Priorität hat. Somit könnte zumindest ein gewisser zusätzlicher Umsatz mit den bestehenden Ressourcen generiert werden. Voraussetzung für 1:1 Path Protection ist natürlich, dass der Provider überhaupt Kunden findet, die eine so geringe Verfügbarkeit akzeptieren können und wollen. Schließlich ist diese noch geringer als bei einer einfachen, ungeschützten Leitung, da der Verkehr unterbrochen wird, wenn entweder die Reserveleitung selber oder die Leitung hoher Priorität gestört ist. Bei abschnittsweiser Nutzung der Reservekapazitäten reduziert sich die Verfügbarkeit für den Verkehr niederer Priorität sogar noch weiter, weil der Ausfall jedes einzelnen Bedarfes hoher Priorität, dessen Ersatzweg teilweise genutzt wird, zum Ausfall des Verkehrs niederer Priorität führt.

Auf ähnliche Weise wie ein Provider mit vielen Kunden könnte auch ein einzelner Kunde selber innerhalb eines angemieteten VPN verfahren. Der Verkehr zwischen seinen Standorten wird in zwei Prioritätsklassen sortiert und daraus ergeben sich dann die benötigten Bandbreiten auf den einzelnen Leitungen. Natürlich bleibt die Frage, ob es bei geschäftlicher Nutzung der Datennetze überhaupt Verkehr niederer Priorität gibt. Sogar E-Mail, das klassische Beispiel für zeitunkritische Daten, wird zunehmend für zeitkritische Kommunikation eingesetzt. Dazu trägt mit Sicherheit bei, dass die aktuelle Qualität der Netze in der Regel außerordentlich gut ist. Auch wenn

der Kunde laut Vertrag streng genommen nur für Best-effort-Übertragung bezahlt, treten nur selten ernsthafte Ausfälle oder Verzögerungen auf, so dass eine hohe Qualität allgemein als Standard angesehen wird. Trotzdem sind durchaus Szenarien denkbar, wo Verkehr niedriger Priorität vorhanden ist und im Sinne von 1:1 Path Protection zu einer wirtschaftlicheren Nutzung des Netzes insgesamt beitragen kann.

1: $n$  Path Protection (vgl. Abbildung 1.3) ist eine Variante, bei der mehrere Übertragungen zwischen den gleichen Punkten über einen einzigen Reservepfad gesichert werden. Im Unterschied dazu 1+1 Path Protection kann hier natürlich nicht simultan auf dem Reservepfad übertragen werden, da im Voraus nicht bekannt ist, welcher der Erstwege als nächster ausfallen wird. Trotzdem ist dieses Verfahren noch relativ schnell, da nur lokal an Quelle und Senke umgeschaltet werden muss. Ganz ähnlich funktioniert  $m:n$  Protection, bei dem  $m$  Ersatzwege  $n$  primären Übertragungswegen gemeinsam zur Verfügung stehen. Allerdings werden diese beiden Varianten im Vergleich zu 1+1 Path Protection in der Praxis eher selten eingesetzt. Ein mögliches Szenario für 1: $n$  Path Protection sind Unterseekabel. Hier können mehrere aktive Kabel durch ein einzelnes Reservekabel geschützt werden. Einerseits wäre das Errichten größerer Reservekapazitäten sehr teuer und andererseits ist der gleichzeitige Ausfall mehrerer Kabel unwahrscheinlich.



Abbildung 1.2: 1+1 Protection.

Abbildung 1.3: 1: $n$  Protection.

Allgemein basiert Path Protection auf der Annahme, dass es sehr unwahrscheinlich ist, dass der Erstweg und sein Ersatzweg gleichzeitig ausfallen. Dagegen bietet keine der angesprochenen Varianten einen Schutz. Von solchen Mehrfachfehlern abgesehen garantiert der Einsatz von Protection aber eine schnelle Reaktion auf Fehler und somit eine hohe Verfügbarkeit. Weiterhin bleibt das Netzmanagement auch bei sehr großen Netzen vergleichsweise

einfach. Der wesentliche Nachteil ist der verhältnismäßig hohe Ressourcenverbrauch. Besonders 1+1 Path Protection hat eine weite Verbreitung gefunden und ist deshalb integraler Bestandteil der in Abschnitt 4 vorgestellten Planungsverfahren. Ein weiterer Grund für die große Verbreitung ist, dass für Protection schon relativ lange herstellerübergreifende Standards existieren, während für die eigentlich effektiveren Restorationverfahren lange Zeit nur proprietäre Lösungen verfügbar waren. Vor allem für große Netzbetreiber, die Geräte von mehr als einem Hersteller einsetzen, ist dies ein entscheidender Aspekt.

In paketbasierten Netzen sind mittels MPLS-Schutzmechanismen realisierbar, die den hier beschriebenen Protection-Verfahren sehr ähnlich sind. Hier werden dann ebenfalls ganze Paketströme in kurzer Zeit auf vorbestimmte Ersatzwege umgeschaltet. Beispielsweise ermöglicht das sogenannte MPLS Fast Reroute eine schnelle Reaktion durch einen Link Protection-Ansatz. Aber auch die Realisierung eines Ende-zu-Ende-Schutzes mit einem zu 1+1 Protection ähnlichen Verfahren ist mit Hilfe von MPLS möglich.

## Restoration

Restoration ermöglicht normalerweise den Schutz von Verbindungen mit geringerem Ressourcenverbrauch als Protection und ist deshalb tendenziell kostengünstiger. Dies wird dadurch erreicht, dass mehrere verschiedene, voneinander unabhängige Bedarfe sich Reservekapazitäten teilen. Der Nachteil ist die gewöhnlich längere Reaktionszeit auf Fehler und die größere Komplexität des Netzwerk-Managements. Genau wie Protection kann Restoration entweder ganze Pfade oder nur einzelne Links schützen, letzteres kommt in der Praxis aber nur selten vor. Wenn ganze Pfade geschützt werden, dann können bei Ausfall eines einzelnen Links theoretisch die restlichen Teilstücke des Pfades entweder für das Backup oder sogar anderweitig freigegeben werden. Diese Funktion, „stub release“ genannt, kann dazu beitragen, die insgesamt im Netz benötigte Reservekapazität möglichst gering zu halten. Steht diese Funktion nicht zur Verfügung und wird der Pfad als Ganzes geschützt, dann bleiben diese Teilstücke des ursprünglichen Pfades im Fehlerfall ungenutzt. Restorationverfahren lassen sich in zwei wesentliche Untergruppen einteilen, statisch und dynamisch.

**Statisches Restoration** Bei diesem Restorationverfahren wird zu jedem primären Weg ein disjunkter Ersatzweg berechnet, aber noch nicht belegt. Erst im Fehlerfall des Erstweges wird auf den Ersatz umgeschaltet. Somit können sich Erstwege, die disjunkt sind, Reservekapazitäten teilen. Durch

die vorberechneten Ersatzwege folgt die Umschaltung im Fehlerfall gerade bei großen Netzen im Allgemeinen deutlich schneller als bei dynamischen Restorationverfahren, die erst im Fehlerfall den Ersatzweg berechnen. Die gemeinsame Nutzung von Reservekapazitäten bedingt allerdings, dass bei zwei unabhängigen Fehlern die benötigte Reservekapazität zum Zeitpunkt des zweiten Fehlers schon durch den ersten Ersatzweg belegt sein kann.

**Dynamisches Restoration** Bei dynamischen Restorationverfahren wird ein Ersatzweg erst dann berechnet, wenn tatsächlich ein Fehler auftritt. Diese Berechnung kann entweder durch ein zentrales Netzmanagementsystem oder, wie beim IP-Rerouting, verteilt durch die einzelnen Netzknoten auf Basis der aktuell freien Netzressourcen geschehen. Um überhaupt einen Ersatzweg nutzen zu können, muss natürlich ausreichend Reservekapazität im Netzwerk vorhanden sein. Diese ist aber nicht einzelnen Bedarfen oder Links zugeordnet, sondern sie steht global zur Verfügung. Dafür erlaubt das Fehlen einer vorab festgelegten Zuordnung eine sehr flexible Reaktion auf Fehler. Im Idealfall können so auch mehrere gleichzeitig auftretende Fehler oder der Ausfall eines Ersatzweges abgefangen werden. Dies ist bei den meisten anderen Verfahren prinzipiell nicht möglich, selbst wenn theoretisch noch Kapazität dafür vorhanden wäre. Allerdings benötigt diese flexible Berechnung des Ersatzweges deutlich mehr Zeit als das starre Umschalten bei Protection oder statischem Restoration. Tendenziell hängt diese Zeit auch von der Größe des Netzwerkes ab, so dass dynamisches Restoration selten in sehr großen Netzwerken angewendet wird.

### Literatur zu Schutzverfahren

Einen Literaturüberblick über Netzplanung unter Berücksichtigung von Schutzverfahren, jedoch mit Stand 1999, gewährt die Veröffentlichung von Soni *et al.* (1999). Einen neueren Literaturüberblick liefern Kerivin und Mahjoub (2005). Sexton und Reid (1997) geben einen grundlegenden Überblick über SDH-Schutzverfahren. Eine Einführung in Protection in WDM-Netzen und eine Klassifikation verschiedener Protection-Varianten findet sich in Ramamurthy und Mukherjee (1999a) und Maier *et al.* (2002). Ramamurthy und Mukherjee (1999b) enthält das gleiche für Restoration. Sowohl Protection als auch Restoration, speziell für Maschennetze, werden in Conte *et al.* (2003) behandelt. Der Standard RFC 3386 (Lai und McDysan (2002)) gibt einen Überblick über verschiedene Schutzverfahren für mehrschichtige Netze. Harrison *et al.* (2001) bietet eine Einführung in MPLS-Schutzmechanismen.

Das in der Praxis wichtige Konzept der „shared risk groups“, also der gegenseitigen Abhängigkeit bestimmter Netzelemente, insbesondere Glasfasern, in Bezug auf Ausfälle, wird detailliert in Ellinas *et al.* (2003) und Grover (2004) beschrieben. Studien zu den Ausfallgründen in realen Netzen gibt es relativ wenige, ein Beispiel dafür ist jedoch Kuhn (1997).

Ein Vorschlag für eine detaillierte und konsistente Klassifizierung der vielfältigen Schutzmechanismen in optischen Netzen findet sich in Barry *et al.* (2004). Diese Klassifizierung ist von der Idee her der in dieser Arbeit gewählten sehr ähnlich. Wenn die von Barry *et al.* (2004) veröffentlichte Klassifizierung oder eine andere ähnlich konsistente zu einem früheren Zeitpunkt der Arbeit verfügbar gewesen wäre, hätte es wahrscheinlich Sinn gemacht, diese zu übernehmen. Eine spätere Veränderung der einmal in den eigenen Veröffentlichungen eingeführten eigenen Klassifizierung erscheint jedoch wenig sinnvoll, da dies eine Inkonsistenz der Begrifflichkeiten zwischen den früheren eigenen Veröffentlichungen und dieser Arbeit zur Folge gehabt hätte.

### 1.2.5 Netzschichten

Der Begriff „Schicht“ kann im Bereich der Telekommunikationsnetze mit unterschiedlichen Bedeutungen belegt werden. Zunächst einmal gibt es das bekannte ISO/OSI-Referenzmodell, welches den Kommunikationsprozess in sieben Schichten einteilt, vgl. Abbildung 1.4. Da die deutschen Äquivalente für die Schichtnamen nicht so gebräuchlich sind und für einige Schichten auch mehrere, leicht voneinander abweichende Bezeichnungen existieren, werden in der Abbildung und im Text die englischen Originalbegriffe verwendet.

Die unterste Schicht, der sogenannte Physical Layer, ist für die Spezifikation der elektrischen bzw. optischen Übertragung zuständig. Die darüber liegenden Schichten organisieren den Medienzugriff, sind verantwortlich für die sichere Übertragung (Fehlerkorrektur), das Routing usw. Die anwendungsnahen Schichten 5 bis 7 sind für die Aufgabenstellung nicht mehr relevant. Ein wesentliches Ziel des ISO/OSI-Modells ist eine klare funktionale Trennung der einzelnen Elemente. Dies vereinfacht die Standardisierung und ermöglicht es, einzelne Komponenten separat zu modifizieren bzw. durch die klare Definition der Schnittstellen zwischen den Schichten auch aus unterschiedlichen Quellen zu beziehen (Kompatibilität).

Das ISO/OSI-Modell ist allerdings eher eine theoretische Struktur. Die in der Praxis weit verbreitete TCP/IP-Protokollfamilie, auf der auch das Internet maßgeblich basiert, passt nicht streng in das fest gefügte ISO/OSI-Schema. Vielmehr liegen diese Protokolle z.T. zwischen den Schichten oder

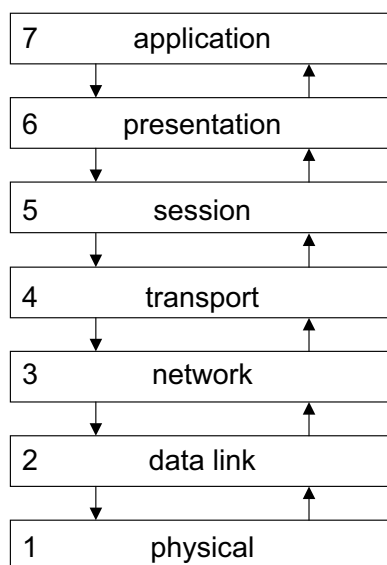


Abbildung 1.4: Das ISO/OSI-Referenzmodell.

umfassen mehr als eine Schicht.

Die für diese Arbeit genutzte Definition der Schichten basiert auf einer ähnlichen Idee, verfolgt aber eine etwas andere Zielsetzung. Die unterste Schicht ist immer die Glasfaser. Darüber wird eine Schicht jeweils durch eine Technologie und die darauf basierenden Protokolle gebildet. An welcher Stelle der Hierarchie sich eine Technologie befindet, wird dadurch definiert, ob sie eine andere Technologie (eine tiefer liegende Schicht) als Basis nutzt oder (eine höher liegende Schicht) transportiert. WDM als grundlegende Bitübertragungstechnologie ist dabei meistens direkt über der Glasfaser anzutreffen. Die Technologien, die diese WDM-Systeme nutzen, beispielsweise SDH oder Ethernet, bilden dann die nächsthöhere Schicht.

SDH und Ethernet sind ein gutes Beispiel dafür, dass die Einordnung in die Schichten im hier genutzten Modell nicht fixiert ist, sondern vom Kontext abhängt. In einem reinen Ethernet-Netz ist Ethernet gleich die erste Schicht über der Glasfaser. Genau so gut kann Ethernet aber auch über SDH-Systeme transportiert werden, in diesem Fall wäre Ethernet eine Schicht höher als SDH. Theoretisch ist hier sogar mit SDH über Ethernet eine Umkehrung möglich. Entscheidend ist immer, welche Technologie oder welches Protokoll welches einkapselt. Zusätzlich kann sich eine Technologie in mehrere Unterschichten aufteilen, beispielsweise wenn die Wahl zwischen 1 Gbit/s und 10 Gbit/s Ethernet besteht. Abbildung 1.5 zeigt drei typische Schichtungen, wie sie bei hier betrachteten Problemstellungen auftreten.



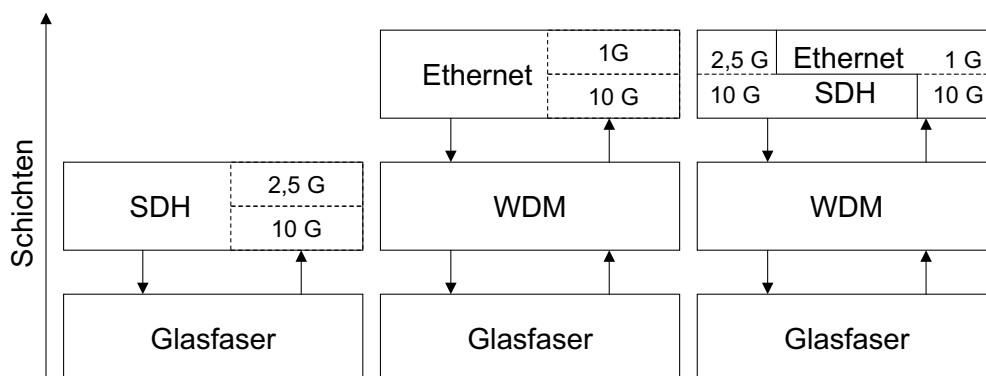


Abbildung 1.5: Drei Beispiele für typische Schichtungen.

Der wesentliche Unterschied zum ISO/OSI-Modell soll an einem Beispiel deutlich gemacht werden. Eine SDH-über-WDM-Übertragung über eine Glasfaserstrecke umfasst mit der hier gewählten Schichtendefinition drei Schichten. Von unten nach oben sind das die Glasfaser, das WDM-System, welches diese nutzt, und das SDH-System, welches wiederum WDM nutzt. Gemäß dem ISO/OSI-Modell gehören Teile der Portkarten des SDH-Systems, insbesondere die Laser und die optischen Steckverbindungen, bereits zum Physical Layer und damit zur untersten Schicht. Das WDM-System ist als reines Punkt-zu-Punkt-Bitübertragungssystem ebenfalls der Schicht 1 zuzuordnen, während die Glasfaser als eigentliches Übertragungsmedium sogar außerhalb des Modells liegt.

Der so definierte Begriff „Schicht“ ist folglich auch von einer hierarchischen Struktur des Netzes abzugrenzen, welche prinzipiell auch als eine Art Schichtung angesehen werden kann (vgl. Abbildung 1.6). Unter einem hierarchischen Netz ist eine Struktur zu verstehen, in der regionale Teilnetze über ein Kernnetz (im Folgenden Backbone genannt) verbunden sind. Eine Verbindung zwischen Regionen, also zwischen verschiedenen Teilnetzen der unteren Hierarchieebene, kann nur über die obere Ebene erfolgen; direkte Querverbindungen sind ausgeschlossen. Die umfassende Planung solch hierarchischer Netze ist nicht der Fokus dieser Arbeit. Jedoch sind bestimmte Aspekte dieser hierarchischen Struktur, wie sie z.B. klassisch im Telefonnetz vorkommt, auch für die hier betrachtete Planung relevant. Spätestens bei der Berücksichtigung von disjunkten Mehrwegführungen zur Erhöhung der Ausfallsicherheit resultieren aus der hierarchischen Struktur zusätzliche Anforderungen an die Netzplanung. Neben der Auswahl geeigneter Übergabepunkte zwischen den Netzteilen (Hub-Knoten), was im Rahmen dieser Arbeit nur randlich betrachtet wird, müssen die Routingverfahren natürlich in der Lage

sein, eine disjunkte Wegeführung Ende-zu-Ende, also über den Planungsbereich des jeweiligen Teilnetzes hinaus, sicherzustellen (vgl. Abschnitt 3.1). Das heuristische Planungstool (vgl. Abschnitt 4.2) ist ursprünglich explizit für die Planung zweistufig hierarchischer Netze entworfen worden.

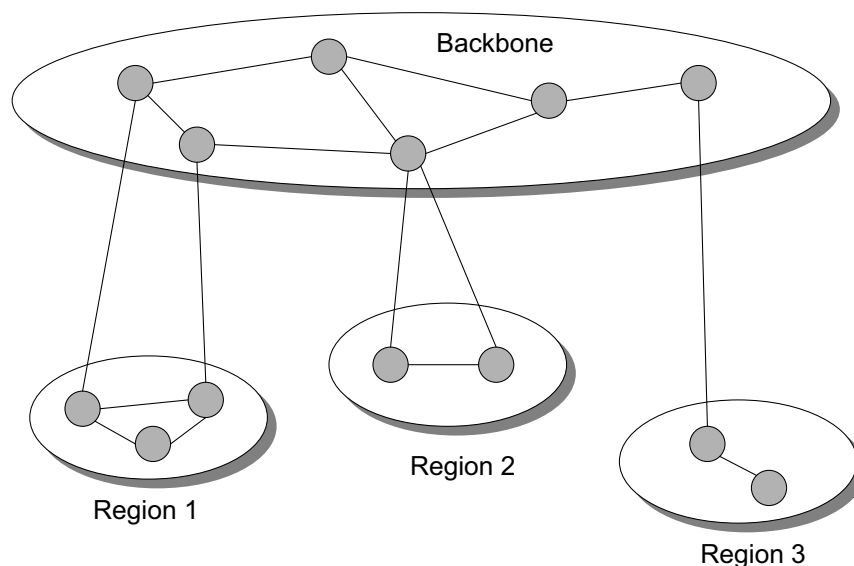


Abbildung 1.6: Beispiel für eine hierarchische Netzstruktur.

**Literatur zu schichtenübergreifender Planung** Die beiden Artikel von Demeester *et al.* (1999) und Doverspike *et al.* (2000) können als eine Art Ausgangspunkt im Bereich „schichtenübergreifende (Englisch: multi-layer) Planung“ gesehen werden. Sie legen viele konzeptionelle Grundlagen und werden dementsprechend von aktuellen Artikeln häufig zitiert. Demeester *et al.* (1999) erläutern am Beispiel eines ATM/SDH/WDM-Netzes grundlegende Ansätze zur schichtenübergreifenden Planung, die auch für ein IP/SDH/WDM größtenteils ihre Gültigkeit behalten. Das von ihnen aufgestellte Blockbild „Framework for multi-layer survivability“ wird in diversen nachfolgenden Publikationen und Vorträgen zitiert und als Ausgangspunkt verwendet. Doverspike *et al.* (2000) formulieren aus dem Blickwinkel des Netzbetreibers AT&T die Anforderungen an aus damaliger Sicht zukünftige mehrschichtige Netzwerke und bieten ein einfaches darauf ausgerichtetes Planungsmodell. Die grundlegenden Ideen hinter einer schichtenübergreifenden Planung für SDH- und WDM-Netze werden auch in Fricke und Höller (2003) und Höller *et al.* (2003) thematisiert.

Grundsätzliche Aspekte zum technischen Zusammenspiel der verschiedenen Protokolle finden sich in Lee und Kim (2002). Das Lehrbuch von Stern und Bala (1999) zeigt schon im Titel die konzeptionelle Ausrichtung auf geschichtete Netze, ist aber ebenfalls eher technisch orientiert. Andere aktuelle Lehrbücher, wie z.B. Grover (2004), streifen das Thema schichtenübergreifende Planung nur randlich.

Der IETF Request for Comments von Lai und McDysan (2002) befasst sich insbesondere mit Schutzverfahren in mehrschichtigen Netzen. Chiu und Strand (2001) beschreiben einen kombinierten IP/WDM-Restoration-Mechanismus, allerdings in sehr knapper Form, Phillips *et al.* (2002) greifen das Thema ebenfalls auf. Diese Publikationen sind beide von AT&T. Es existiert allgemein nur wenig Literatur zur Verfügbarkeit unter Berücksichtigung mehrerer Schichten, Jakab *et al.* (2000) und Jereb *et al.* (2002) bilden hier eine Ausnahme und behandeln die Verfügbarkeit bzw. Zuverlässigkeit von mehrschichtigen optischen Netzen.

Xu *et al.* (2000) bieten einen Überblick über verschiedene Möglichkeiten für IP-über-SONET-Architekturen (sowie „Optical Packet Switching“ und „Optical Burst Switching“). Leider beinhaltet die Arbeit keinen entsprechend umfassenden Literaturüberblick. Das Thema IP über SONET wird im Wesentlichen unter dem Aspekt der dabei genutzten Protokolle und Kapselmechanismen behandelt. Auch Theelen *et al.* (2002) beschränken sich auf konzeptionelle Überlegungen zu den Protokollschichten. Ghani (2000) beschreibt die Evolution von vielschichtigen IP over Optical-Netzen hin zu einem von ihm „lamda-labeling“ genannten Verfahren, welches IP/MPLS direkt auf der optischen Schicht aufsetzt.

Die meisten aktuellen in diesem Kontext relevanten Publikationen beschäftigen sich schwerpunktmäßig mit GMPLS. Zhang und Duresi (2002) behandeln sehr detailliert und anschaulich das Thema Schutzverfahren in GMPLS-Netzwerken. Sie beschränken sich dabei aber auf qualitative Aussagen. Die Dissertation von Pongpaibool (2004) enthält eine Fülle nützlicher Informationen zum Thema GMPLS bzw. IP over Optical Networks, z.B. Algorithmen zur dynamischen LSP-Suche oder Ansätze zur Verbesserung verschiedener im Rahmen von GMPLS genutzter Protokolle sowie ein umfangreiches Literaturverzeichnis.

Insgesamt ist die sehr umfassende, kommentierte Bibliographie für den Themenkomplex Netzplanung von Yuan (2001) erwähnenswert. Leider ist eine gezielte Suche nach schichtenübergreifender Planung auf Grund der gewählten Sortierung nicht möglich. Höller und Voß (2005) enthalten einige der zentralen Gedanken zu schichtenübergreifender Planung, wie sie in dieser Arbeit verstanden wird, in Kurzform.

# Kapitel 2

## Problemstellung

Aus dem starken Wachstum der Anforderungen, vor allem an die Übertragungsgeschwindigkeit der Telekommunikationsnetze, ergibt sich für die Netzbetreiber eine fortwährende Notwendigkeit, neue Netze aufzubauen und vorhandene zu erweitern. Das grundlegende Vorgehen dazu ist in Abbildung 2.1 dargestellt. Ausgehend von einer, meistens weitgehend vorhandenen, Glasfaserstruktur, den prognostizierten Bedarfen der Kunden und der zur Verfügung stehenden Technologie wird die eigentliche Netzplanung durchgeführt. Ergebnis der Planung sind dann eine oder mehrere mögliche Gerätekonfigurationen, also welche Geräte wo und in welcher Stückzahl im Netz aufgebaut werden müssen, und die geplante Führung der Bedarfe durch das Netz (Routing). Aus dieser allgemeinen Problemstellung ergeben sich für jeden konkreten Einzelfall vor allem durch die Fülle unterschiedlicher technischer Realisierungsmöglichkeiten und diverse Randbedingungen aus dem realen Kontext der Netzbetreiber sehr vielfältige, aber vor allem auch sehr komplexe Planungsprobleme.

Ziel der Planung ist es, alle Bedarfe möglichst kostengünstig zu führen, also die Kosten der benötigten Netzinfrastruktur zu minimieren. Die Gesamtkosten entsprechen hier der Summe aller Preise für die eingeplanten Geräte. Diese Definition wird auch in der Literatur zur Netzplanung durchgängig so benutzt, in englischsprachigen Publikationen wird der äquivalente Term „cost“ gebraucht. Nicht direkt in diesem Planungsprozess berücksichtigt werden eventuelle Betriebskosten für Personal, Wartung usw. Diese sind für Netzbetreiber selbst bei bestehenden Netzen nur schwer realistisch quantitativ einer bestimmten Netzkonfiguration zuzuordnen. Indirekt kann dies natürlich durch die Auswahl des bei der Planung berücksichtigten Geräteportfolios geschehen. So können sehr wartungsaufwändige oder stromhungrige Geräte mit höheren Kosten belegt oder gar ausgeschlossen werden.

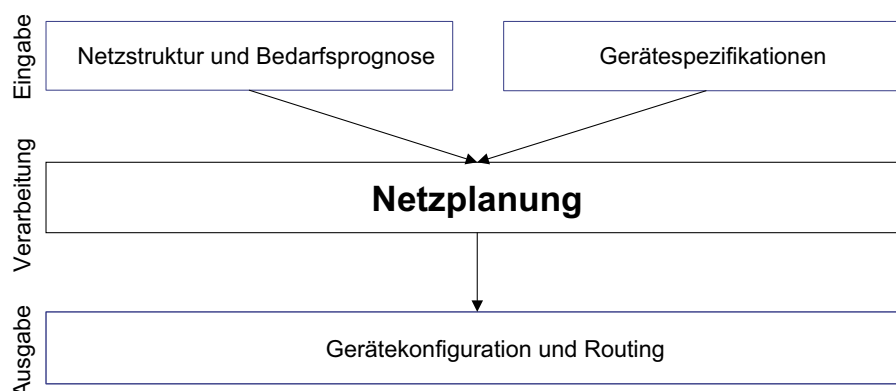


Abbildung 2.1: Planungsablauf.

Die hier bearbeiteten Problemstellungen stammen primär aus dem Bereich der strategischen Planung, also langfristiger Planung mehrere Jahre im Voraus. Nur vereinzelt werden auch taktische Planungsprobleme angesprochen. Für operative Planung eignen sich die eingesetzten Modelle durchweg nicht. Hier wäre ein noch höherer Detaillierungsgrad notwendig.

Aufgrund der enormen Entwicklungsgeschwindigkeit bei den Übertragungstechnologien ist auch eine sogenannte „Grüne-Wiese-Planung“, also eine Netzplanung ohne Berücksichtigung vorhandener Systeme, kein rein akademischer Fall. Einerseits kann es sein, dass bestehende Infrastruktur ohnehin bereits technologisch überholt und abgeschrieben ist. Andererseits können neue Dienste, wie z.B. Fernsehen über DSL, Anforderungen an die Übertragungsgeschwindigkeit stellen, die so deutlich über dem bisher Vorhandenen liegen, dass die bestehende Hardware vernachlässigt werden kann. Beides gilt jedoch auch im Bereich langfristiger Planung nur sehr bedingt für die Glasfaserinfrastruktur. Deshalb baut die Planung hier weitestgehend auf einer bestehenden Struktur auf. Abbildung 2.2 zeigt den primären Gestaltungsspielraum der Netzplanung. Die gewünschten Dienste beeinflussen natürlich maßgeblich die Anforderungen an das Netz. Sie sind ebenfalls eine Randbedingung, die für die eigentliche Infrastrukturplanung in der Regel fixiert ist.

Ein zentraler Aspekt bei der Planung mehrschichtiger Netze ist, dies soweit wie möglich simultan für die unterschiedlichen Schichten durchzuführen und nicht sequentiell. Durch die gegenseitigen Abhängigkeiten zwischen den Schichten führt eine Zerlegung des Problems und eine getrennte Planung ggf. zu suboptimalen Ergebnissen.

Die weiter unten liegenden Schichten stellen Kapazitäten für die höheren Schichten bereit. Bei einer sequentiellen Planung von oben nach unten können

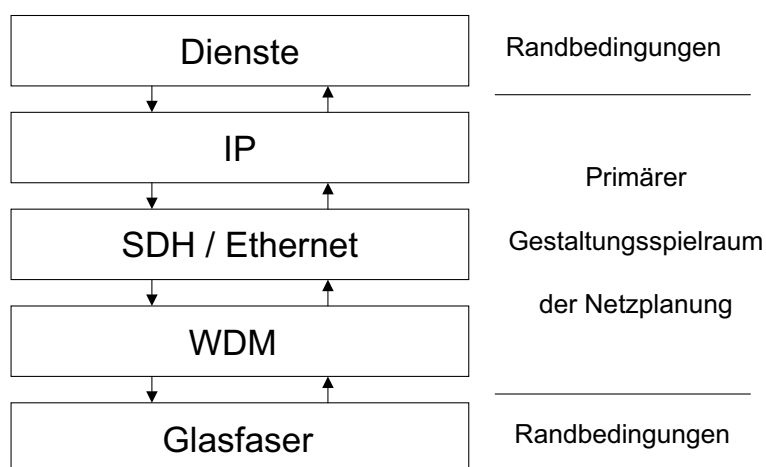


Abbildung 2.2: Gestaltungsspielraum der Netzplanung.

diese Kapazitäten jedoch nicht unbedingt optimal platziert und dimensioniert werden. Als Beispiel dafür soll die Kombination aus SDH und WDM dienen. Theoretisch ist es möglich, erst rein die SDH-Schicht zu planen und dann in einem zweiten Planungsschritt auf den besonders verkehrsreichen Kanten WDM vorzusehen. Dies berücksichtigt jedoch nicht vollständig die möglichen Einsparungen, welche durch die Skaleneffekte erzielt werden können. Möglicherweise ist es besser, auf der SDH-Ebene einige Umwege in Kauf zu nehmen, um so den Langstreckenverkehr gezielt zu bündeln und WDM deutlich effizienter einsetzen zu können. Rein aus SDH-Sicht wäre eine solche Lösung schlechter, in der Gesamtsicht über beide Schichten hinweg kann sie jedoch deutlich günstiger sein. Natürlich kann dies durch ein iteratives Vorgehen auch bei einer sequentiellen Planung berücksichtigt werden, ein von vornherein integrierter Ansatz umgeht diese Probleme jedoch vollständig und ist deshalb vorzuziehen.

Andere Gründe für einen übergreifenden Planungsprozess sind Funktionalitäten, die auf mehreren Schichten vorhanden sind. Beispielsweise erscheint eine Koordination der Schutzmechanismen von MPLS- und Transportschicht sinnvoll, einerseits um im Fehlerfall eine definierte Reaktion des Netzes zu garantieren und andererseits um ein doppeltes Vorhalten von Ersatzkapazitäten zu vermeiden. Darüber hinaus kann es technologische Gründe dafür geben, dass auf bestimmten Schichten Restriktionen bestehen, die auch auf andere Schichten einen Einfluss haben, z.B. eine Begrenzung der Anzahl möglicher Nachbarn eines Knotens oder Reichweitebeschränkungen.

Im Sinne einer langfristigen Planung ist es natürlich nicht wünschenswert eine Lösung zu erzielen, die zwar für die gegebenen Prognosen optimal ist,

aber bei Abweichungen davon wahrscheinlich nachteilig. Ein Beispiel dafür ist das Führen von Bedarfen auf Umwegen, um die Netzkapazität restlos auszunutzen. Für den statischen Fall kann damit kurzfristig die Investition in neue Übertragungstechnik auf einem direkteren Weg gespart werden. Langfristig ist dies bei steigenden Bedarfen jedoch sogar teurer, da oft ohnehin früher oder später eine Kapazitätserweiterung benötigt wird. Steht diese dann zur Verfügung, ist ein aufwändiges Rerouting der auf Umwegen geführten Bedarfe notwendig. Teilweise können diese gar nicht mehr umgelegt werden, weil den Kunden keine Umkonfiguration ihrer Geräte oder schlimmstenfalls sogar eine Unterbrechung der Übertragung zugemutet werden soll bzw. darf. Damit wird dauerhaft unnötig Kapazität belegt und der ursprüngliche kurzfristige Vorteil wird langfristig zu einem Nachteil.

Unabhängig davon ist eine Planung mit einer solch hohen Detaillierung wie dem exakten Routing eines einzelnen Bedarfes zwar akademisch interessant, um die Güte der Methoden zu verifizieren. Für die Praxis hat eine Abweichung der Lösung um beispielsweise 1% bei strategischer Planung jedoch keinerlei Bedeutung, da die Unsicherheit der Eingangsdaten (Bedarfsprognose, Gerätepreise und -spezifikationen) bei der strategischen Planung um ein Vielfaches höher liegt.

# Kapitel 3

## Modellierung

Um Planungsprobleme, nicht nur aus dem Bereich Telekommunikation, mit mathematischen Methoden behandeln zu können, muss eine geeignete mathematische Repräsentation der Problemstellung gefunden werden. Die allgemeine Modellierung der gesamten Problemstellung umfasst drei wesentliche Teile. Zuerst muss die Struktur des Netzes beschrieben werden. Weiterhin muss die zur Verfügung stehende Hardware modelliert werden. Und nicht zuletzt müssen natürlich die zu transportierenden Bedarfe der Kunden entsprechend abgeschätzt und abgebildet werden. Die spezielle, dem jeweiligen Lösungsverfahren spezifisch angepasste Modellierung wird nachfolgend in Abschnitt 4 bei den einzelnen Verfahren beschrieben.

### 3.1 Netzstruktur

Ein oft genutzter Weg für die Beschreibung von Netzwerkstrukturen, insbesondere im Bereich Telekommunikation, ist mit Hilfe von Graphen. Sehr umfassend befassen sich Ahuja *et al.* (1993) mit der Modellierung und Lösung von Flussproblemen verschiedenster Art auf der Basis von Graphen. In der Tat ergibt sich die Modellierung der Infrastruktur eines kabelgebundenen Kommunikationsnetzes, also verschiedene durch Glasfasern verbundene Standorte, als Graph, also verschiedene, über Kanten verbundene Knoten, fast intuitiv. Die Kantengewichtung des Graphen kann dabei zum Beispiel die Länge der Glasfasern oder die Kapazität der Übertragungsstrecke widerspiegeln.

Eine ähnliche Repräsentation kann man auch für die logische Netzebene, also die einzelnen Datenströme, nutzen. Hier können Kommunikationspartner durch Knoten dargestellt werden und eine Kante zeigt jeweils eine Ver-



bindung zwischen diesen Partnern an. Die Gewichtung kann hier beispielsweise das zu übertragende Datenvolumen angeben. Während die physikalische Netzebene in der Praxis häufig weitmaschig ist, ist die logische Netzebene tendenziell eher hoch vermascht oder sogar vollständig vermascht, vgl. Abbildung 3.1.

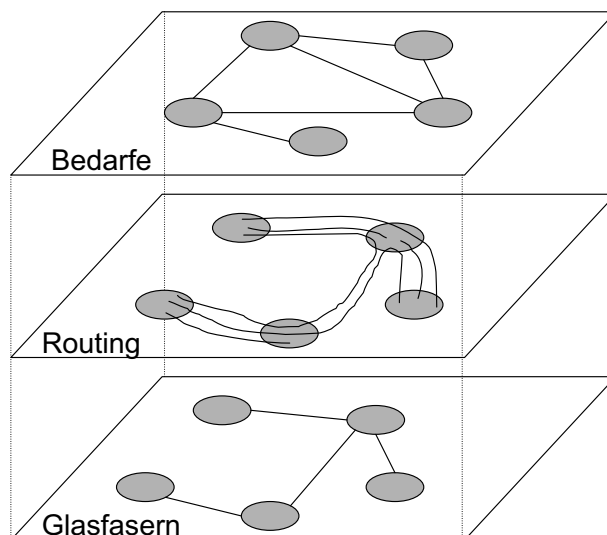


Abbildung 3.1: Physikalische und logische Schichten eines Netzes.

In der Praxis können einige Fälle auftreten, in denen eine einfache Repräsentation „Knoten = Betriebsstelle“, „Kante = Glasfaser“ nicht ausreichend ist, insbesondere bei disjunkter Wegführung. Bei einem schlichten Graphen sind per Definition erst einmal alle Kanten unabhängig voneinander. Dies gilt jedoch für reale Glasfaserstrecken nicht unter allen Umständen. Obwohl sie logisch voneinander unabhängig sind, können zwei Glasfasern (ggf. auch punktuell oder abschnittsweise) die gleiche Trassenführung besitzen, z.B. bei der Überbrückung eines Flusses (vgl. Abbildung 3.2). Bei einem Schaden an dieser gemeinsamen Trasse fallen beide Faserstrecken gleichzeitig aus; sie sind voneinander abhängig. Eine solche Abhängigkeit muss beim ausfallsicheren Routing disjunkter Wege berücksichtigt werden, lässt sich aber nicht ohne weiteres als Graph modellieren. Deswegen sind die üblichen Verfahren für disjunkte Wegesuche entweder gar nicht oder zumindest nur bedingt für Netzstrukturen mit solchen Kantenabhängigkeiten geeignet.

Das Einführen von Hilfsknoten löst das Problem der Abhängigkeiten auf den ersten Blick, hat aber unerwünschte Nebeneffekte. Durch die neuen Knoten ergeben sich zusätzliche Wege, die explizit für das Routing verboten werden müssen. So entsteht beispielsweise in Abbildung 3.3 ein neuer Weg von

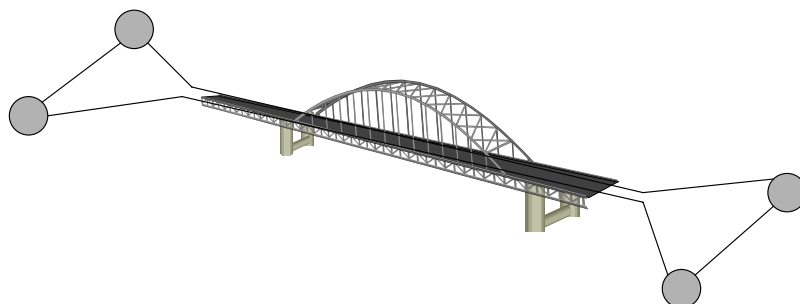


Abbildung 3.2: Kantenabhängigkeit durch geographische Gegebenheiten.

$A$  nach  $B$  über den Hilfsknoten  $E$ , zu dem keine Glasfaser im eigentlichen Netz existiert. Das Verbot einzelner Routingwege ist aber wiederum nicht oder nur sehr umständlich in die normalen Routingalgorithmen auf der Basis von Graphen einzubinden.

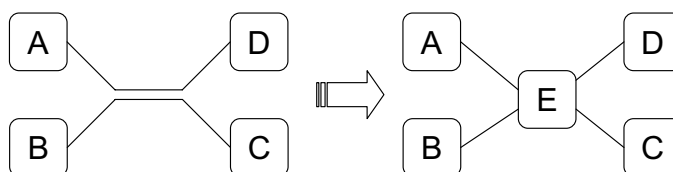


Abbildung 3.3: Graphentransformation mittels eines Hilfsknotens.

Ein gängiges Verfahren mit diesen Problemen umzugehen ist, solche unzulässigen Wegführungen, sei es über Hilfsknoten oder abhängige Fasern, in einer ersten Phase zuzulassen, aber mit hohen Strafkosten zu belegen. Nach Abschluss der Planung können dann die verbleibenden unzulässigen Wegführungen entweder manuell oder mit Hilfe einer spezialisierten Nachverarbeitung beseitigt werden. Allerdings führt dieser Ansatz in der Regel zu einer suboptimalen Lösung des Ausgangsproblems.

Es gibt eine Reihe von Veröffentlichungen, die sich detailliert mit dem Problem solcher Kantenabhängigkeiten („shared risk groups“, SRG) auseinandersetzen, beispielsweise Ellinas *et al.* (2003). Im Rahmen dieser Arbeit wird diese Problematik nicht näher betrachtet sondern vorausgesetzt, dass das Routingverfahren frei von eventuell vorhandenen Abhängigkeiten agieren kann. Diese werden entweder in einem Vorverarbeitungsschritt bereinigt oder in einem (ggf. iterativen) Nachverarbeitungsprozess individuell behandelt. Eine gesicherte Aussage, ob voneinander abhängige Wegführungen in der Praxis eher die Ausnahme sind und deshalb richtigerweise vernachlässigt werden können, oder ob sie ein sehr häufig auftretendes Problem sind, lässt

sich nicht treffen. Solche Daten, so sie denn überhaupt konsolidiert vorhanden sind, werden von Netzbetreibern naturgemäß nicht öffentlich kommuniziert.

Schutzverfahren, wie in Kapitel 1.2.4 vorgestellt, können im Allgemeinen nur vollständig funktionieren, wenn die zugrunde liegende physikalische Topologie des Netzwerkes mindestens zwei disjunkte Wege für jede Datenverbindung bereitstellen kann. Diese Eigenschaft eines Netzwerkes bzw. allgemeiner eines Graphen wird zweifacher Zusammenhang genannt. Dazu reicht es nicht aus, dass jeder Knoten mindestens zwei Nachbarn hat, wie in Abbildung 3.4 zu sehen ist. Der Knoten in der Mitte hat vier Nachbarn, er stellt jedoch trotzdem einen sogenannten „single point of failure“ (oder bezogen auf den Graphen einen Artikulationspunkt) dar. Ein solcher Punkt ist dadurch gekennzeichnet, dass das Netzwerk in zwei getrennte Teile zerfällt, wenn er entfernt wird. Mitunter wird in der Literatur zwischen „zweifachem Zusammenhang“ und „2-fachem Zusammenhang“ unterschieden, so z.B. bei Grover (2004) auf Seite 180, wo „biconnected“ in Bezug auf Knoten und „two-connected“ in Bezug auf Kanten verwendet wird. Allgemein üblich ist diese Nomenklatur jedoch nicht.

Im Fall einer bestehenden Infrastruktur muss diese vor Beginn der Planung auf zweifachen Zusammenhang getestet werden. Dies kann mit Hilfe eines Algorithmus, wie er z.B. in Reiß und Viebeck (2000) vorgestellt wird, in polynomialer Zeit durchgeführt werden. Sollten dabei Artikulationspunkte gefunden werden, so gibt es verschiedene Möglichkeiten, diese unkompliziert zu beseitigen, vgl. Abbildung 3.5. Entweder kann eine neue direkte Verbindung (also Glasfaser) zwischen den beiden Netzteilen geschaffen werden (Fall *a*) oder der Artikulationspunkt wird intern in zwei unabhängige Knoten aufgeteilt (Fall *b*). Dies ist in der Praxis aufwändiger, als es den Anschein hat, da die beiden Teile in allen wichtigen Belangen wie z.B. Stromversorgung und Brandschutz getrennt sein müssen. In Einzelfällen sind diese beiden einfachen Lösungen praktikabel. Sollte jedoch eine große Zahl an Artikulationspunkten im Netz vorhanden sein, kann es sinnvoll sein, eine vollständige Überarbeitung der gesamten Glasfaserstruktur in Erwägung zu ziehen.

Analog dazu muss ein Netzwerk  $n$ -fach zusammenhängend sein, um Schutz gegen  $n - 1$  gleichzeitige Fehler bieten zu können. Da der Fall  $n > 2$  in der Praxis eher selten ist und ein entsprechend hoher Zusammenhang des Netzes oft hohe Kosten verursacht, werden die meisten kommerziellen Netze nur für den Schutz gegen einzelne Fehler dimensioniert. Lediglich im militärischen Bereich oder bei Netzen mit sehr langen Reparaturzeiten wird teilweise ein Schutz zumindest gegen Doppelfehler gefordert.

Ein weiterer zu berücksichtigender Punkt bei geschützten Netzen ergibt sich, wenn das Netz aus mehreren Teilen besteht, die nicht als Ganzes geplant werden sollen oder können. Ein Beispiel dafür sind verschiedene Regionalnet-

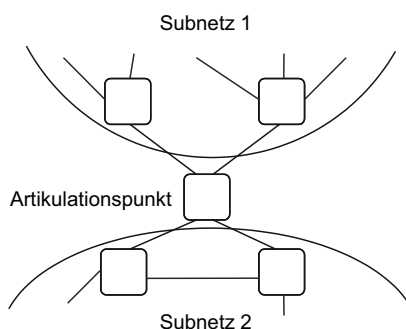


Abbildung 3.4: Ein single point of failure.

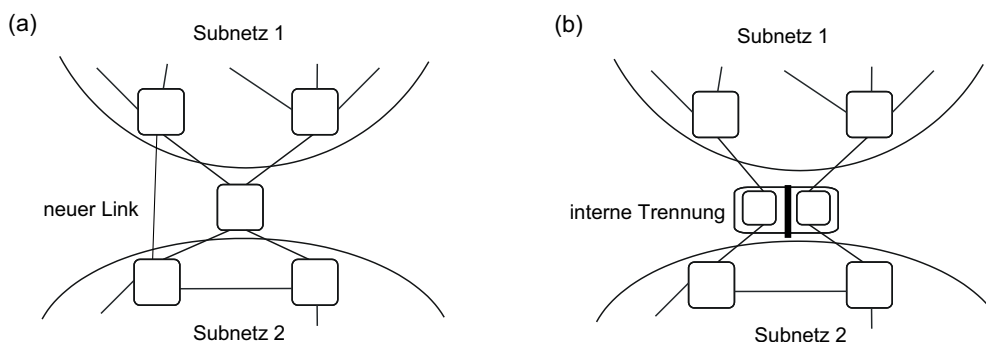


Abbildung 3.5: Auflösen eines single point of failure.

ze, die über einen landesweiten Backbone verbunden werden. Ein geschützter Bedarf muss folglich an zwei verschiedenen Stellen an den anderen Netzteil übergeben werden („dual homing“). Wenn zwischendurch ein eigenständiger Netzbereich nur im Transit durchlaufen wird, dann bestehen dort nur zwei einzelne Bedarfe, jeweils zwischen den beiden Übergabepunktpaaren am Ein- und Ausstieg. Diese beiden Bedarfe müssen jedoch disjunkt geführt werden, um den Schutz von Ende zu Ende sicherzustellen, vgl. Abbildung 3.6. Eine Möglichkeit, dies zu modellieren, ist das Einfügen von Hilfsknoten in die Netztopologie, durch die die eigentlich unabhängigen Übergabepunkte zwischen den Netzen logisch gekoppelt werden. Die Kanten zu den Hilfsknoten und die Knoten selber können von den Bedarfen kostenlos genutzt werden. Für den Planungsalgorithmus sieht ein solcher Bedarf dann wie ein herkömmlicher geschützter Bedarf aus und kann auch entsprechend behandelt werden. Am Ende der Planung werden die Hilfsknoten und ihre angrenzenden Kanten einfach gelöscht.

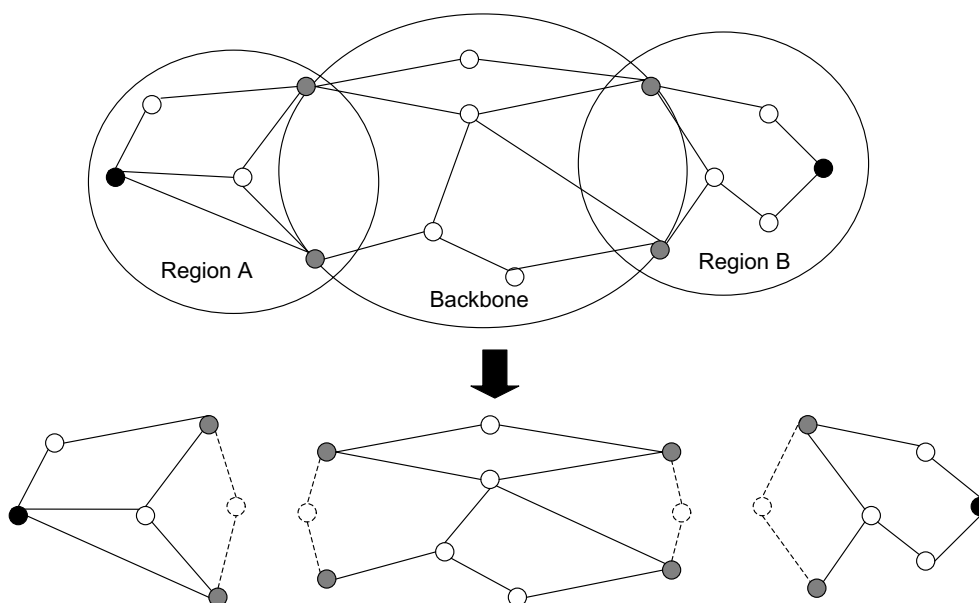


Abbildung 3.6: Hilfsknoten bei Schutz über mehrere Teilnetze hinweg.

## 3.2 Geräte

Für die meisten aktuellen Technologien im Bereich optischer Datenübertragung gibt es eine fast unüberschaubare Zahl verschiedener Geräte von einer Vielzahl von Herstellern. Ziel dieser Planung ist es, daraus ein übersichtliches Portfolio generischer Geräte abzuleiten, welches alle relevanten Funktionen abbilden kann. Das Augenmerk liegt also nicht auf (ggf. sogar herstellerspezifischen) Details, sondern auf den Grundfunktionalitäten, die im Wesentlichen sogar technologieunabhängig sind. Möglicherweise notwendige problemspezifische Anpassungen werden dann erst bei Bedarf je nach Problemstellung vorgenommen und hier nicht allgemein diskutiert.

Eine grundlegende Unterscheidung, die in Einklang mit der Modellierung der Netze als Graphen steht, ist die Unterteilung in Kanten und Knoten. Es gibt also bestimmte Geräte, die den Netzknoten zugeordnet werden und andere, die den Übertragungsstrecken zwischen diesen Knoten zugeordnet werden. Abbildung 3.7 zeigt beispielhaft alle primären Komponenten.

Die Anschlüsse der Kunden sowie die Übertragungsleitungen zu den anderen Netzknoten laufen in eine universelle Schaltmatrix, hier gemäß der SDH-Terminologie Cross-Connect genannt. Bei Ethernet wäre dies dementsprechend ein MPLS-Switch. Von diesem Grundgerät kann es verschiedene Varianten geben, die sich in ihrer Kapazität (Größe der Schaltmatrix bzw.

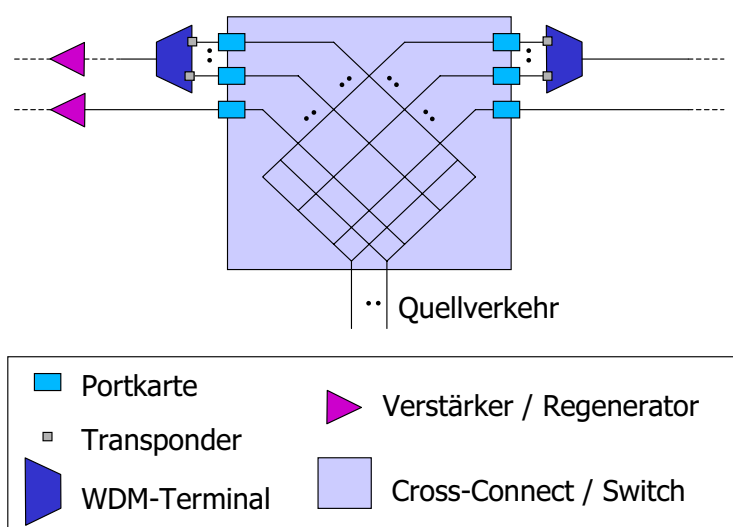


Abbildung 3.7: Grundlegende Komponenten des Gerätemodells.

daraus abgeleitet Anzahl und Kapazität der Ports) und in den Gerätekosten unterscheiden. Diese Geräte erlauben grundsätzlich eine beliebige Verschaltung aller Ports. Varianten mit Beschränkung der Freiheitsgrade bei der Verschaltung, wie z.B. optische Add-Drop-Multiplexer (OADM), werden nicht berücksichtigt. Diese Basisgeräte entsprechen den Knoten des Graphen, wobei jeder Knoten prinzipiell eine beliebige Anzahl Basisgeräte unterschiedlicher Kapazität aufnehmen kann.

Die Modellierung der Kanten ist deutlich aufwendiger und auch stärker technologieabhängig. Die eigentlichen Glasfasern werden im Normalfall als vorhanden vorausgesetzt. Bei der Beschaltung der Glasfasern gibt es verschiedene Möglichkeiten. Bei kurzen Übertragungsstrecken können direkt Portkarten angeschlossen werden, welche die Verbindung zu den Cross-Connects herstellen. Abhängig von der Übertragungsgeschwindigkeit und der notwendigen Reichweite kommen hier unterschiedliche Typen zum Einsatz. Wenn sehr große Reichweiten benötigt werden, müssen auf der Übertragungsstrecke selber in bestimmten Abständen (z.B. alle 80 km) zusätzlich Verstärker und/oder Regeneratoren eingesetzt werden. Gängige Übertragungsgeschwindigkeiten sind beispielsweise 1 Gbit/s, 2,5 Gbit/s, 10 Gbit/s und 40 Gbit/s.

Für noch höhere Übertragungsgeschwindigkeiten auf einer einzelnen Glasfaser können mehrere optische Signale mit Hilfe der Wellenlängenmultiplex-Technologie gebündelt werden. In diesem Fall werden nur Portkarten sehr kurzer Reichweite („intra-office“) benötigt, die das Signal aus dem Cross-

Connect zum WDM-Multiplexer am gleichen Standort transportieren. Dort wird in einem Transponder die Wellenlänge umgewandelt, und dann werden in der eigentlichen Multiplexereinheit die unterschiedlichen Signale zusammengefasst. Eine solche Multiplexereinheit kann eine feste Zahl von Kanälen verarbeiten, beispielsweise 16 oder 40. Unabhängig von der tatsächlich genutzten Kanalzahl wird immer ein vollständiger Multiplexer benötigt, lediglich die Transponder müssen nur für die genutzten Kanäle bestückt werden. Neben der hohen Übertragungskapazität zeichnen sich WDM-Systeme gewöhnlich auch durch ein sehr hohe Übertragungsreichweite aus.

Natürlich kann ein solches Modell über den hier gewählten Detaillierungsgrad hinaus nahezu beliebig verfeinert werden. Konkret denkbar wäre zum Beispiel die Einbeziehung von Modulbaugruppen für die Schnittstellen. Sowohl die Schnittstellen der Cross-Connects als auch die Transponder der WDM-Systeme sind oft nicht als einzelne Karten ausgeführt, sondern in Modulen zusammengefasst. Beispielsweise könnte der Einbau von 2,5 Gbit/s Karten in einen Cross-Connect über ein 4x2,5 Gbit/s zu 1x10 Gbit/s Konvertermodul erfolgen. Eine ähnliche Zwischenebene ist auch bei einigen WDM-Systemen vorhanden. Beispielsweise wird ein maximal 160 Kanäle fassendes Basissystem ggf. erst bei Bedarf mit Modulgruppen (z.B. für je 40 Kanäle) bestückt, in die dann wiederum bei Bedarf einzelne Transponder eingesteckt werden können. Somit ergibt sich eine mehrfach abgestufte Kapazität. Auch die einzelnen Transponder können selber wiederum Multiplexfunktionen übernehmen, beispielsweise 4x2,5 Gbit/s auf 1x10 Gbit/s. Diese Details können jedoch sowohl im Laufe der Entwicklung der Geräte als auch von Hersteller zu Hersteller stark variieren. Da DXCs und OXCs sich in bezug auf den modularen Aufbau mit Basisgerät und Portkarten nicht prinzipiell unterscheiden, eignet sich diese Modellierung sowohl für elektrische als auch für gemischte und rein optische Systeme. OTN als Nachfolgetechnologie zu SDH kann somit ebenfalls abgebildet werden.

Neben den technischen Daten des jeweiligen Gerätes ist ein weiterer elementarer Bestandteil der Eingabedaten natürlich sein Preis. Schließlich sollen am Ende die Gesamtkosten der Netzinfrastruktur minimiert werden und diese bestimmen sich aus der Summe der Preise aller benötigten Geräte. Natürlich stehen hier auf Grund des Zeithorizontes nur geplante Listenpreise aus der Roadmap der Hersteller zur Verfügung und noch keine etablierten Marktpreise. Dies ist jedoch weitgehend unproblematisch, solange das Preisverhältnis zwischen den einzelnen Geräten davon unbeeinflusst bleibt.

Einige spezifische Eigenschaften der Gerätekosten haben einen besonderen Einfluss auf den Planungsprozess. Einerseits unterliegen die Gerätepreise oft einer starken Economies of Scale in Abhängigkeit von der Übertragungsgeschwindigkeit. Als Faustformel gilt die Aussage, dass eine

vierfache Geschwindigkeit lediglich die doppelten Kosten verursacht. Dies führt dazu, dass es sich lohnen kann, gewisse Umwege in der Wegeführung in Kauf zu nehmen, um auf langen Strecken große Systeme einsetzen zu können. Weiterhin ist dadurch ein großes, aber nicht sehr gut ausgelastetes Gerät oft wirtschaftlicher als mehrere kleine Geräte mit einer höheren Auslastung. Der Auslastungsgrad eines Netzes ist also nicht unbedingt ein brauchbares Kriterium für dessen Effizienz. Ausgenommen davon ist häufig die jeweils aktuellste Technologiegeneration, die sogar im Gegenteil überproportional teuer sein kann. Aktuell wäre dies beispielsweise 40 Gbit/s Übertragungstechnik.

Aus der festen Modulgröße der Cross-Connects und Multiplexer folgt eine stufige Kostenfunktion in Abhängigkeit von der Kapazität. Um einen einzigen Bedarf führen zu können, wird bereits eine relativ hohe Grundinvestition für die Basisgeräte notwendig. Weitere Bedarfe können dann allerdings sehr günstig durch einfaches Hinzufügen von Portkarten bzw. Transpondern mitgeführt werden, bis die Kapazität des Basisgerätes erschöpft ist. Dann wird wiederum eine erneute, hohe Grundinvestition notwendig. Diese Stufigkeit der Gerätekosten kann bei sequentiellen Routingverfahren problematisch sein, vgl. Abschnitt 4.2.1. Deshalb werden neben der Stufenfunktion für die Planung zusätzlich zwei daraus abgeleitete, geglättete Funktionen der Gerätekosten in Abhängigkeit von der benötigten Bandbreite erstellt, eine für die Knoten und eine für die Kanten (vgl. Abbildung 3.8).

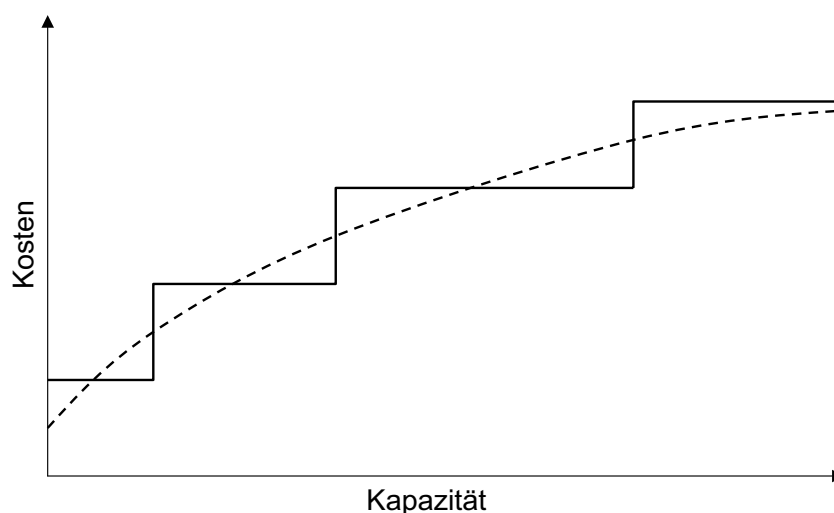


Abbildung 3.8: Stufige Kostenfunktion und Glättung.

Das in der Praxis ebenfalls relevante Kapazitätserweiterungsproblem kann nur sehr bedingt abgebildet werden. Restkapazitäten bereits vorhandener Infrastruktur sind verhältnismäßig einfach im Rahmen der hier genutzten



Ansätze modellierbar (vgl. z.B. Abschnitt 4.1.1). Nicht so einfach lässt sich das Verschieben von Geräten abbilden. Beispiel: Ein Knoten hat bisher zwei kleine Cross-Connects und nach neuer Planung einen großen. Die beiden kleinen werden in diesem Knoten abgebaut und dafür in einem anderen Knoten wieder aufgebaut.

Letztendlich geht es darum, ein Modell zu erstellen, welches genau genug ist, um gute Ergebnisse erzielen zu können, aber sich nicht in Details verliert, welche für die Gesamtplanung entweder unerheblich sind, oder die Lösung des Problems bzw. die Beschaffung sinnvoller Eingangsdaten unnötig erschweren oder gar unmöglich machen. Deshalb werden einige der genannten Punkte, wie z.B. die Multiplextransponder, in dieser Arbeit bewusst aus der Modellierung herausgelassen. Der gewählte Ansatz ist insgesamt jedoch flexibel genug, um bei Bedarf um solche Aspekte ergänzt zu werden, wenn sie sich für eine konkrete Aufgabenstellung als notwendig erweisen sollten.

### 3.3 Bedarfe

Während bei der Netztopologie noch relativ sichere Aussagen getroffen werden können und auch die Geräteentwicklung sich zumindest gut abschätzen lässt, ist das Erstellen einer Bedarfsprognose für einige Jahre im Voraus mit einer großen Unsicherheit behaftet. Große Änderungen der von den Nutzern nachgefragten Datenvolumen können sich beispielsweise durch neue Dienste (z.B. Filesharing), Wettbewerbseinflüsse (z.B. Preisverfall) oder Technologiesprünge (z.B. DSL) ergeben und sind oft schlecht über lange Zeiträume zu prognostizieren. In der Vergangenheit hat sich das nachgefragte Datenvolumen in der Regel deutlich erhöht und nicht, bzw. wenn dann nur kurzfristig, verringert. In der Praxis wird diese Prognoseunsicherheit oft abgebildet, indem mit verschiedenen Szenarien gerechnet wird, beispielsweise dem eigentlich erwarteten Szenario, einem worst-case und einem best-case Szenario. Daraus ergibt sich dann ein Korridor aus Eingangsdaten und somit auch ein Korridor in dem sich die Ergebnisse für das Netzdesign bewegen.

Von dieser Prognoseunsicherheit abgesehen kann sich auch noch das Problem ergeben, dass das Datenaufkommen nicht konstant ist, sondern selbst im Laufe kurzer Zeiträume stark schwanken kann, beispielsweise im Laufe eines Tages oder sogar im Laufe einiger Minuten. Im Rahmen der operativen Planung von Netzen kann dieses Verhalten mit stochastischen Methoden beschrieben werden.

Für die strategische Planung, insbesondere in Backbonenetzen, kann auf eine so detaillierte Modellierung der Verkehrsströme jedoch in der Regel verzichtet werden. Großen Kunden wird in der Regel eine bestimmte konstante

Datenrate vertraglich zugesichert, diese muss also ohnehin jederzeit vorgehalten werden (sogenannte „leased lines“). Der Verkehr der vielen Klein- und Privatkunden durchläuft bereits eine oder mehrere Aggregationsebenen bevor er in die Backbonenetze gelangt. Dadurch ist er bereits stark verdichtet und im Verlauf deutlich konstanter als weiter außen im Netz. In der Praxis reicht es für die taktische und strategische Planung also meistens aus, mit festen Werten (in der Regel Spitzenwerten) zu rechnen.

Somit können die Bedarfe für jedes Knotenpaar des Netzes für einen Planungsdurchlauf durch einen einzigen Zahlenwert angegeben werden. Selbst wenn dahinter verschiedene Kunden stehen, wird aus betrieblichen Gründen davon ausgegangen, dass alle Daten zwischen zwei Endpunkten auf dem gleichen Weg geführt werden. Lediglich bei Technologien, die unidirektionale Übertragungen ermöglichen, wie z.B. MPLS, müssen die beiden Richtungen natürlich getrennt berücksichtigt werden. Hier wird dann ein Wert für die zu übertragende Datenmenge von  $A$  nach  $B$  angegeben und ein weiterer Wert für die umgekehrte Richtung von  $B$  nach  $A$ . Daraus ergibt sich dann bei einem Netz mit  $n$  Knoten letztendlich eine Bedarfsmatrix mit  $n \cdot (n - 1)$  (bzw.  $n \cdot (n - 1)/2$  im bidirektionalen Fall) Einträgen als Eingangsdaten.

# Kapitel 4

## Lösungsverfahren

Aufbauend auf der in Abschnitt 3 beschriebenen Modellierung werden in diesem Kapitel zwei verschiedene Lösungsstrategien für die Problemstellung vorgestellt: exakte Verfahren für kleine Probleminstanzen und heuristische Verfahren für große Probleminstanzen. Die hier genutzten Probleminstanzen reichen von kleinen, zufällig generierten Instanzen bis hin zu großen Praxisproblemen aus einer Zusammenarbeit mit Netzbetreibern und Geräteherstellern. Die umfangreiche Problemdatenbank von Wessály und Pióro (2006) wurde leider erst verfügbar, als es für eine Einbeziehung in diese Arbeit zu spät war.

### 4.1 Exakte Lösungsverfahren

Aus der in Abschnitt 3 vorgestellten, grundlegenden Modellierung lassen sich die relevanten Aspekte für die Netzplanung fast unmittelbar als eine Kombination mathematischer Formeln ableiten. Die Zielfunktion ist die Summe aller Kosten für die genutzten Geräte. Die einzelnen Nebenbedingungen bilden die Kapazitätsrestriktionen und das Routing ab. Hilfreich für die Lösbarkeit ist die Tatsache, dass sowohl in der Zielfunktion als auch in den Nebenbedingungen alle Entscheidungsvariablen nur durch lineare Operationen verknüpft sind. Durch die Ganzzahligkeitsbedingungen für die meisten Entscheidungsvariablen wird aus dem ursprünglich rein linearen Optimierungsproblem dann ein (gemischt) ganzzahliges Problem. Ein solches Modell lässt sich prinzipiell mit bekannten mathematischen Standardverfahren wie dem Simplex-Algorithmus, beim Auftreten ganzzahliger Entscheidungsvariablen unter Zuhilfenahme von Branch&Bound, lösen.

Speziell für Netzplanungsprobleme eignen sich zwei verschiedene Ansätze, um das zugrunde liegende mathematische Problem zu formulieren. Bei der

Pfadformulierung wird jedem möglichen Pfad jedes Bedarfes eine Entscheidungsvariable zugewiesen. Ein solches Modell enthält bedingt durch die große Zahl möglicher Pfade sehr viele Variablen, aber nur relativ wenige Nebenbedingungen. Bei der sogenannten Flussformulierung wird jeder Kante für jeden Bedarf eine Entscheidungsvariable zugeordnet. Ein solches Modell enthält dementsprechend deutlich weniger Variablen, aber eine große Anzahl an Nebenbedingungen (Flussbedingungen), um die Pfade durch das Netz zu beschreiben. Details zu beiden Modelltypen finden sich beispielsweise in Lee und Lewis (2006). Eine grundsätzliche Überlegenheit einer der beiden Modellvarianten kann nicht festgestellt werden, vielmehr hängt es stark vom Einzelfall ab. Für diese Arbeit wird das weit verbreitete Flusskonzept benutzt, im Wesentlichen weil es sich relativ einfach um bestimmte, hier geforderte Funktionalitäten, wie z.B. Schutzverfahren, erweitern lässt.

Die grundlegenden Komponenten eines Flussmodells sollen hier anhand eines einfachen Minimierungsproblems dargestellt werden. Gegeben ist ein Graph  $G = (N, A)$  bestehend aus einer Menge an Knoten  $N$  und einer Menge an Pfeilen  $A$ , welche die Knoten miteinander verbindet. Jeder Pfeil  $(i, j) \in A$  ist weiterhin durch ein nichtnegatives Gewicht  $c_{ij}$ , welches die Kosten für den Transport einer Einheit entlang dieses Pfeils beschreibt, und eine maximale Kapazität  $\hat{u}_{ij}$ , welche den kumulierten Transport entlang dieses Pfeils begrenzt, gekennzeichnet. Jeder Knoten  $i \in N$  hat einen bestimmten Bedarf  $d(i)$  (oder ein bestimmtes Angebot, je nach Vorzeichen von  $d(i)$ ). Die Variablen  $x_{ij}$  beschreiben den Fluss entlang der Pfeile  $(i, j)$ .

Ziel ist es, die Bedarfe aller Knoten möglichst kostengünstig zu befriedigen, ohne die Kapazitätsgrenzen zu verletzen. Eine dazu passende mathematische Formulierung kann wie folgt aussehen: Die Zielfunktion (4.1) dient der Minimierung aller anfallenden Kosten auf den Pfeilen. Die Nebenbedingungen (4.2) sind die sogenannten Flussbedingungen. Sie stellen sicher, dass der Bedarf eines jeden Knotens genau der Summe der ein- und ausgehenden Flüsse entspricht. Die Nebenbedingungen (4.3) dienen der Einhaltung der Kapazitätsgrenzen. Diese drei wesentlichen Komponenten, Minimierungsfunktion, Flussbedingungen und Kapazitätsbedingungen finden sich in abgewandelter Form in allen in dieser Arbeit vorgestellten Modellen.

Minimiere

$$\sum_{(i,j) \in E} c_{ij} \cdot x_{ij} \tag{4.1}$$

u.d.N.:

$$\sum_j x_{ij} - \sum_j x_{ji} = d(i) \quad \forall i \in N \quad (4.2)$$

$$0 \leq x_{ij} \leq \hat{u}_{ij} \quad \forall (i, j) \in A \quad (4.3)$$

Eine große Zahl unterschiedlicher ganzzahliger Modelle im Kontext Telekommunikation, insbesondere WDM, finden sich beispielsweise in Bhandari (1999) und Zang (2003). Ein umfassendes Standardwerk zum Thema Flussformulierungen ist Ahuja *et al.* (1993). Flussformulierungen werden aber nicht nur im Telekommunikationsbereich erfolgreich genutzt, sondern unter anderem in den Bereichen Logistik und Produktionsplanung. Beispielsweise zeigen Glover *et al.* (1992) neben einer Einführung in Netzwerke, Flüsse und lineare Modelle diverse praktische Anwendung auf. Ein sehr umfassendes Werk zur linearen Modellierung ist Dantzig und Thapa (1997).

Ein wie auch immer geartetes mathematisches Modell muss zur Lösung durch Computer mit Hilfe einer entsprechenden Modellierungssprache beschrieben werden. Hier gibt es eine Reihe frei verfügbarer und kommerzieller Produkte, zum Teil getrennt von der Software zur eigentlichen Problemlösung (Solver), zum Teil sehr eng damit verbunden. Die hier genutzte Sprache ist „A Mathematical Programming Language“ (AMPL), vgl. AMPL Optimization LLC (2006). AMPL wurde ursprünglich von AT&T entwickelt, wird inzwischen aber unabhängig davon weitergepflegt. Eine gute Einführung in AMPL bietet Fourer *et al.* (2003). Andere Beispiele für Modellierungssprachen sind LINGO (LINDO Systems Inc. (2006)) oder das „General Algebraic Modeling System“ (GAMS), vgl. GAMS Development Corporation (2006). Bei Modellierungssprachen, die keinen eigenen Solver mitbringen, wie z.B. AMPL, besteht wiederum eine große Auswahl sowohl an kommerziellen als auch an frei verfügbaren Solvern. Die hier genutzte Kombination aus AMPL und CPLEX ist im wissenschaftlichen Bereich weit verbreitet und kann von der Leistungsfähigkeit her durchaus als Referenz angesehen werden.

Das grundsätzliche Vorgehen bei der Anwendung solcher Modellierungssprachen und Solver ist immer ähnlich. Aus der allgemeinen, realen Problemstellung werden die Zielfunktion, die Nebenbedingungen sowie die Variablen und Parameter abgeleitet. Mit anderen Worten: Das reale Problem wird als ein mathematisches Modell formuliert. Dieses Modell wird dann in die gewählte Modellierungssprache (in unserem Fall AMPL) umgesetzt, so dass es von einem Computer verarbeitet werden kann (vgl. auch Anhang B). Solch ein Modell kann dann mit Hilfe eines Solvers, z.B. CPLEX, für die jeweilige konkrete Ausprägung der Problemstellung (Probleminstanz) gelöst werden. Dieses Vorgehen ist in Abbildung 4.1 als Ablaufplan dargestellt.

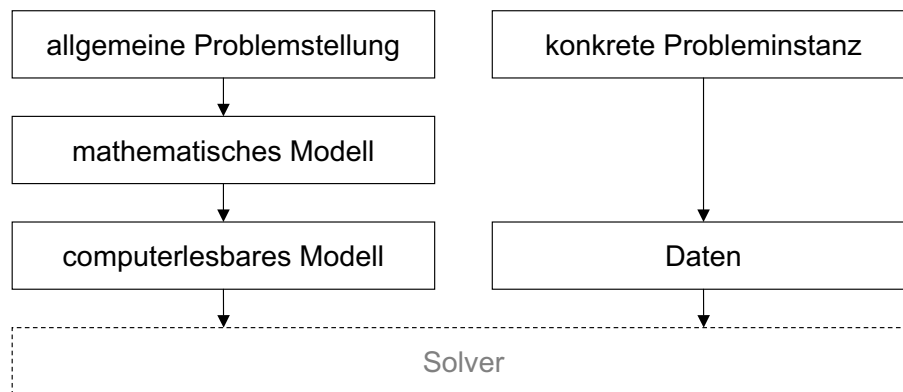


Abbildung 4.1: Grundsätzlicher Ablauf bei der Erstellung von Modellen für exakte Lösungsverfahren.

Problematisch an ganzzahligen Modellen für die hier betrachtete Art von Problemstellungen ist, dass bereits Teilprobleme NP-schwer sind, wie sich aus dem in Garey und Johnson (1979) aufgezeigten Nachweis der NP-Vollständigkeit der dazugehörigen Entscheidungsprobleme ableiten lässt. Somit stoßen auf ganzzahligen Modellen aufbauende, exakte Lösungsverfahren schnell an ihre Grenzen. Viele reale Probleminstanzen lassen sich selbst auf den schnellsten verfügbaren Computern nicht mehr in endlicher Zeit beweisbar optimal lösen. In Abschnitt 5 wird der z.T. extreme Anstieg der Rechenzeit und des Speicherplatzbedarfs bei nur geringer Vergrößerung der Probleminstanz deutlich.

#### 4.1.1 Das ganzzahlige Basismodell

Im Folgenden wird nun aus Gründen der Übersichtlichkeit zunächst ein Basismodell vorgestellt, welches das grundlegende Planungsproblem abbildet. Die grundsätzliche Idee, die Problemstellung auf diese Art abzubilden, wurde durch das Modell von Melian *et al.* (2004) inspiriert, welches allerdings eine Beschränkung für die Anzahl der möglichen Pfade enthält. Diese wurde dann später in Melian *et al.* (2005a) allerdings ebenfalls aufgegeben. Die Benennung der einzelnen Variablen und Parameter ist größtenteils deckungsgleich mit der in den beiden genannten Literaturstellen. Dies bietet sich alleine schon aus Gründen der Konsistenz mit gemeinsamen Veröffentlichungen, z.B. Melian *et al.* (2006), an. In der Tat haben sich aber viele Bezeichner, wie z.B.  $N$  für die Menge der Knoten oder  $D$  für die Menge der Bedarfe (abgeleitet aus den jeweiligen englischen Begriffen „nodes“ und „demands“), allgemein in

der Literatur zu diesem Themenkomplex durchgesetzt. Grundsätzlich basiert das Modell auf weit verbreiteten Flussformulierungen, wie sie beispielsweise im Standardwerk Ahuja *et al.* (1993) allgemein eingeführt werden. Eine Übersicht über verschiedene Flussformulierungen im Telekommunikationskontext bietet Minoux (2006). Anschließend an das Basismodell werden verschiedene Erweiterungen behandelt. Für die jeweilige praktische Anwendung müssen dann ein oder mehrere dieser Erweiterungen in das Modell integriert werden.

Dieses Basismodell unterstützt einen Bedarfstyp (z.B. 2,5 Gbit/s) und die Grundgeschwindigkeit der Übertragungssysteme hat die gleiche Granularität, Grooming ist also nicht erlaubt. Der gesamte Bedarf zwischen zwei Knoten wird auf einem einzigen Weg geführt; ein Aufteilen ist nicht zulässig. Diese Restriktion für das Routing hat sich ursprünglich aus den Randbedingungen der betrachteten realen Problemstellung ergeben und wurde dann allgemein so beibehalten. Als zur Verfügung stehende Geräte werden ein Typ von Cross-Connect mit einer festen Anzahl an Ports, ein Typ von WDM-Multiplexer mit einer festen Zahl an Kanälen, sowie die entsprechenden Portkarten, Verstärker, Transponder und ggf. Regeneratoren berücksichtigt. Verstärker-, Regenerator- und ggf. Glasfaser- und Baukosten tauchen nicht einzeln im Modell auf, sondern sind individuell in den jeweiligen Kantenkosten beinhaltet. Für eine gegebene Distanz lassen sich diese Kosten bezogen auf die verschiedenen zur Wahl stehenden Übertragungsgeschwindigkeiten im Voraus und unabhängig vom eigentlichen Optimierungsmodell ermitteln. Sie hängen nur von der Länge der Übertragungsstrecke ab und stellen somit für das Modell Konstanten dar.

Das Modell unterstützt in einfacher Form auch die Kapazitätserweiterungsplanung, indem über  $g_e$  und  $h_n$  vorhandene Restkapazitäten für Knoten und Kanten berücksichtigt werden können. Da es sich bei  $g_e$  und  $h_n$  lediglich um Konstanten handelt, können diese ohne positive oder negative Auswirkungen auf die Rechenzeit im Modell belassen werden, auch wenn sie für die konkrete Anwendung gerade nicht benötigt werden. Eine für die Praxis nutzbare Kapazitätserweiterungsplanung lässt sich damit aber nur bedingt durchführen. Der häufig auftretende Fall, dass Geräte umgebaut werden, ist so nicht abzubilden. Beispielsweise kann es sinnvoll sein, einen Cross-Connect, der an einem Standort durch einen größeren ersetzt wird, stattdessen an einem anderen Standort wieder aufzubauen.

Die gegebene Infrastruktur besteht aus einer Menge von Knoten  $N$  und einer Menge ungerichteter Kanten  $E$ , die diese Knoten verbinden. Für die Modellformulierung ist es zusätzlich hilfreich, aus der Menge der Kanten  $E$  eine gerichtete Menge  $E'$  abzuleiten, bei der jede Kante durch ein Paar entgegengesetzter Pfeile ersetzt wird. Die Bedarfe, die durch das Netz trans-

portiert werden müssen, werden als Menge  $D$  gegeben, wobei die einzelnen Elemente  $d_{st}$  dieser Menge die Größe des Bedarfes zwischen den Knoten  $s$  und  $t$  in Vielfachen der Basiseinheit angeben. Bezogen auf das obige Beispiel von 2,5 Gbit/s als Basiseinheit, würde  $d_{25} = 4$  bedeuten, dass ein Bedarf von insgesamt 10 Gbit/s zwischen den Knoten 2 und 5 geführt werden muss. In der Grundform des Modells sind alle Bedarfe bidirektional.

#### Eingabedaten - Gerätekosten:

$C_e^F$	Kosten einer Basisverbindung (z.B. 2,5 Gbit/s SDH) auf Kante $e$ (Aufpreis für Portkarten langer Reichweite, Verstärker, Regeneratoren)
$C_e^W$	Kosten einer WDM-Strecke auf Kante $e$ (Multiplexer, Verstärker)
$C^{OS}$	Kosten eines Cross-Connect-Basisgerätes
$C^C$	Kosten für einen WDM-Kanal (Transponderpaar)
$C^P$	Kosten für eine Portkarte (kurzer Reichweite) am Cross-Connect

#### Eingabedaten - Kapazitäten der Übertragungstechnik:

$M^W$	Kapazität eines WDM-Systems (Anzahl der Wellenlängen)
$M^{OS}$	Kapazität eines Cross-Connects (Anzahl der Ports)
$g_e$	Anzahl freier WDM-Kanäle auf Kante $e$ (Restkapazität des vorhandenen Netzes)
$h_n$	Anzahl freier Cross-Connect Ports in Knoten $n$ (Restkapazität des vorhandenen Netzes)

#### Entscheidungsvariablen:

$f_e$	Anzahl der Basisverbindungen (z.B. 2,5 Gbit/s SDH) auf Kante $e$
$w_e$	Anzahl der WDM-Systeme auf Kante $e$
$v_e$	Anzahl der belegten Kanäle in den WDM-Systemen auf Kante $e$
$y_n^S$	Anzahl der Cross-Connects in Knoten $n$
$z_{ij}^{st}$	1 wenn der Bedarf $(s, t)$ über die Kante $(i, j)$ geroutet wird; sonst 0

#### Zielfunktion:

$$\text{Minimiere } \sum_{e \in E} ((C_e^F + 2C^P)f_e + C_e^W w_e + (C^C + 2C^P)v_e) + \sum_{n \in N} C^{OS} y_n^S$$

u.d.N.:



$$\sum_{j \in N} z_{ji}^{st} - \sum_{j \in N} z_{ij}^{st} = \begin{cases} -1 & i = s \\ 0 & \forall i \neq s, t \\ +1 & i = t \end{cases} \quad \forall (s, t) \in D \quad (4.4)$$

$$\sum_{(s,t) \in D} d_{st}(z_{ij}^{st} + z_{ji}^{st}) \leq v_e + f_e \quad \forall e \in E \text{ mit } i \text{ und } j \text{ angrenzend an } e \quad (4.5)$$

$$v_e \leq M^W w_e + g_e \quad \forall e \in E \quad (4.6)$$

$$\sum_{e \text{ angrenzend an } n} (v_e + f_e) \leq M^{OS} y_n^S + h_e \quad \forall n \in N \quad (4.7)$$

$$y_n^S \geq 0 \text{ und ganzzahlig} \quad \forall n \in N \quad (4.8)$$

$$f_e, w_e, v_e \geq 0 \text{ und ganzzahlig} \quad \forall e \in E \quad (4.9)$$

$$z_{ij}^{st} \in \{0, 1\} \quad \forall (i, j) \in E', (s, t) \in D \quad (4.10)$$

Die Zielfunktion minimiert die Summe aller Gerätekosten. Dabei beschreibt die linke Summe die Kantenkosten und die rechte die Knotenkosten. Die Nebenbedingungen (4.4) sind die sogenannten Flusserhaltungsbedingungen. Sie garantieren, dass jeder Bedarf auf einem ununterbrochenen Pfad von der Quelle zur Senke geführt wird. (Bei den hier unterstellten bidirektionalen Bedarfen sind Quelle und Senke beliebig austauschbar.) Quelle und Senke müssen also jeweils eine abgehende bzw. ankommende Kante besitzen, die den Bedarf führt, was +1 bzw. -1 in den Flussbedingungen entspricht. Alle anderen Knoten werden von einem Bedarf entweder nicht berührt, dann ist sowohl die linke als auch die rechte Summe gleich Null, oder der Bedarf fließt auf genau einer Kante in den Knoten hinein und auf genau einer anderen Kante wieder heraus. In diesem Fall sind beide Summen gleich Eins, so dass die Differenz wiederum Null ergibt. Theoretisch lässt es diese Art der Flussformulierung zu, dass sich außerhalb des eigentlichen Pfades für den Bedarf in sich geschlossene Kreise bilden. Allerdings wird dies implizit durch die daraus resultierenden Kosten verhindert, welche sich in der zu minimierenden Zielfunktion niederschlagen. Details dazu werden in Abschnitt 4.1.3 diskutiert.

Die Nebenbedingungen (4.5) stellen sicher, dass die installierte Übertragungskapazität auf jeder Kante ausreicht, um die jeweiligen Bedarfe führen zu können. Bei der hier vorhandenen Identität von Basisübertragungskapazität und Basisbedarfsgröße könnte an Stelle des „ $\leq$ “

auch ein „=“ gesetzt werden. Für spätere Erweiterungen ist „ $\leq$ “ jedoch unabdingbar, deshalb wird es bereits hier eingeführt. Die Nebenbedingungen (4.6) stellen für die eingeplanten WDM-Kanäle die notwendige Anzahl an Multiplexern zur Verfügung. Die Nebenbedingungen (4.7) passen die Kapazität der Cross-Connect-Systeme an die Kapazitäten der angrenzenden Kanten an. Die Bedingungen (4.8) bis (4.10) stellen die Ganzzahligkeit der Entscheidungsvariablen sicher.

### 4.1.2 Funktionale Erweiterungen

Das Modell kann durch eine Vielzahl von Veränderungen und Erweiterungen an die jeweiligen Gegebenheiten realistischer Problemstellungen angepasst werden. Die wichtigsten davon werden in den folgenden Abschnitten vorgestellt.

#### Hop Limitierung

Die maximal zulässige Anzahl an Hops, also die maximale Zahl an Kanten, die ein Bedarf auf dem Weg von der Quelle zur Senke durchlaufen kann, kann über die zusätzlichen Nebenbedingungen (4.11) auf den Wert  $H$  begrenzt werden. Eine solche Begrenzung kann aus zwei unterschiedlichen Gründen Sinn machen. Einerseits kann es technologische Gründe geben,  $H$  zu begrenzen, beispielsweise weil jeder zusätzliche Hop zu einer Verzögerung des Signals führen kann. Andererseits führt eine Begrenzung der Anzahl von Hops auch zu einer Verkleinerung des Lösungsraumes und somit ggf. zu einer besseren Lösbarkeit der Problemstellung. Die genauen Auswirkungen der Hop-Beschränkung auf die Rechenzeit von CPLEX werden in Abschnitt 5.1 diskutiert.

Ebenfalls denkbar ist eine individuelle Hop-Beschränkung für jeden Bedarf, beispielsweise auf die Anzahl Hops des kürzesten Weges plus einer Konstante. Hierzu muss lediglich die einzelne Konstante  $H$  durch eine Matrix aus Konstanten  $H^{st}$  ersetzt werden. Die Anzahl der Hops des kürzesten Weges kann vorab einfach über den Dijkstra-Algorithmus berechnet werden. Die in Abschnitt 4.2.1 beschriebene Heuristik bietet diese Funktion ohnehin implizit mit an.

$$\sum_{(i,j) \in E} (z_{ij}^{st} + z_{ji}^{st}) \leq H \quad \forall (s, t) \in D \quad (4.11)$$

### 1+1 Path Protection geschützte Bedarfe

Das in Abschnitt 1.2.4 vorgestellte Schutzverfahren 1+1 Path Protection kann sehr einfach in das Modell aufgenommen werden. Hierzu müssen die Flussformulierungen (4.5) um einen zweiten, disjunkten Pfad ergänzt werden, indem „-1“ und „+1“ auf der rechten Seite der Gleichungen durch „-2“ und „+2“ ersetzt werden. Damit verlässt der Bedarf die Quelle auf zwei verschiedenen Kanten und erreicht die Senke ebenfalls auf zwei verschiedenen Kanten. Die Flussbedingungen für die anderen Knoten ändern sich für eine kantendisjunkte Wegführung nicht. Um jedoch knotendisjunkte Wege zu erhalten, müssen zusätzlich noch die Nebenbedingungen (4.12) hinzugefügt werden. Diese stellen sicher, dass ein Bedarf keinen Knoten außer der Senke auf mehr als einer ankommenden Kante erreicht.

$$\sum_{j \in N} z_{ji}^{st} \leq 1 \quad \forall i \neq s, t \quad \forall (s, t) \in D \quad (4.12)$$

Bei der Erweiterung für geschützte Bedarfe ist zu beachten, dass eine Kombination mit der Hop-Beschränkung die Anzahl der Kanten begrenzt, die beide Wege zusammen benutzen dürfen. Eine auf einzelne Pfade bezogene Beschränkung ist in dieser einfachen Form nicht möglich. Aus diesem Grund kann die planerische Funktion der Hop-Beschränkung nicht genutzt werden; lediglich der Laufzeitvorteil bei der Berechnung verbleibt.

### Cross-Connects unterschiedlicher Größen

Analog zur Definition des Cross-Connects im Basismodell können relativ einfach zusätzliche XC-Typen abweichender Kapazität definiert werden. In diesem Fall müssen für jeden XC-Typ eigene Konstanten  $C^{Ox}$ ,  $M^{Ox}$  und Variablen  $y_n^x$  definiert werden. Eine Erweiterung des Modells auf drei unterschiedliche Typen, einen kleinen, einen mittleren und einen großen Cross-Connect, würde wie folgt aussehen:

$C^{OS}$	Kosten eines kleinen Cross-Connect-Basisgerätes
$C^{OM}$	Kosten eines mittleren Cross-Connect-Basisgerätes
$C^{OL}$	Kosten eines großen Cross-Connect-Basisgerätes
$M^{OS}$	Kapazität eines kleinen Cross-Connects (Anzahl der Ports)
$M^{OM}$	Kapazität eines mittleren Cross-Connects (Anzahl der Ports)
$M^{OL}$	Kapazität eines großen Cross-Connects (Anzahl der Ports)
$y_n^S$	Anzahl der kleinen Cross-Connects in Knoten $n$
$y_n^M$	Anzahl der mittleren Cross-Connects in Knoten $n$
$y_n^L$	Anzahl der großen Cross-Connects in Knoten $n$

Diese neuen Variablen und Konstanten müssen an jeder Stelle des Basismodells eingefügt werden, an der die ursprünglichen  $C^{OS}$ ,  $M^{OS}$  und  $y^S$  auftauchen. Dadurch wird erreicht, dass die jeweils beste Kombination an XC-Typen für jeden Knoten eingeplant wird. Die ursprünglichen Nebenbedingungen (4.7) für die Kapazität der Cross-Connects beispielsweise würden dann zu (4.13) verändert werden.

$$\sum_{e \text{ angrenzend an } n} (v_e + f_e) \leq M^{OS} y_n^S + M^{OM} y_n^M + M^{OL} y_n^L + h_e \quad \forall n \in N \quad (4.13)$$

### Grooming

Als Grooming wird allgemein die Fähigkeit bezeichnet, mehrere kleine Bedarfe für eine Übertragungsstrecke zu einem großen Bedarf zusammenzufassen und am Ende wieder zu entbündeln. Somit können viele kleine Bedarfe effizient über wenige Leitungen hoher Kapazität transportiert werden. Das setzt voraus, dass die Cross-Connects am Anfang und am Ende der Übertragungsstrecke in einer feinen Granularität schalten und nicht nur eine ankommende Leitung als Ganzes verarbeiten können. Dies erhöht den technischen Aufwand in den Cross-Connects, ermöglicht dafür aber im Gegenzug eine bessere Auslastung der Übertragungsstrecken. Bei einer Technologie wie SDH erfolgt Grooming auf Containerebene, hier werden mehrere Container kleiner Kapazität zu einem größerer Kapazität zusammengefasst. Bei einer Technologie wie IP wird ohnehin immer auf der Paketebene als kleinster Einheit gearbeitet. Hier ist Grooming zwangsläufig integraler Bestandteil des

Systems.

Das bisher vorgestellte Basismodell unterstützt nur eine Bedarfsgröße. Die Erweiterung des Modells auf Grooming soll jetzt anhand eines Beispiels für ein SDH-Netz beschrieben werden. Bei unveränderter Bandbreite der Übertragungsleitungen von 2,5 Gbit/s wird die Granularität der Bedarfe von 2,5 Gbit/s (VC4-16c) auf 155 Mbit/s (VC4) reduziert. Ein Großteil der am Markt erhältlichen Cross-Connects ist ohnehin in der Lage, auf VC4-Ebene zu schalten und mehrere dieser kleinen Bedarfe zu größeren Datenströmen zusammenzufassen.

Die Bedarfe  $d_{st}$  werden folglich auf eine Granularität von 155 Mbit/s undefiniert, größere Einzelbedarfe werden durch eine entsprechende Zahl an VC4-Äquivalenten repräsentiert. Da sichergestellt ist, dass zwischen zwei Knoten nur genau ein Weg für alle Bedarfe gewählt wird, ist trotz dieser theoretischen Aufteilung großer Bedarfe gewährleistet, dass sie in der Lösung unverändert als Ganzes geführt werden. Weiterhin müssen die Nebenbedingungen (4.5) modifiziert werden. Der neu eingeführte Multiplikator spiegelt wider, dass ein 2,5 Gbit/s System bis zu 16 VC4-Äquivalente führen kann. Die Aufteilung dieser Gesamtkapazität auf verschiedene Bedarfe ist variabel, damit wird die Grooming-Funktionalität ermöglicht.

$$\sum_{(s,t) \in D} d_{st} z_e^{st} \leq 16 * (v_e + f_e) \quad \forall e \in E$$

### Zusätzliche Übertragungsgeschwindigkeiten

Das Basismodell unterstützt nur eine einzige Geschwindigkeit für die Übertragungsleitungen. Eine Erweiterung auf mehrere Geschwindigkeiten, z.B. 2,5 Gbit/s und 10 Gbit/s, ist jedoch einfach in das Modell zu integrieren. Die neue Schicht mit der höheren Übertragungsgeschwindigkeit wird analog zur vorhandenen implementiert. Dazu werden die entsprechenden Portkosten  $C^{P'}$ , Kosten für die Übertragungsstrecken  $C_e^{F'}$  und die Entscheidungsvariablen für die Anzahl der schnellen Systeme auf den einzelnen Kanten  $f'_e$  definiert. In diesem Zusammenhang bietet es sich an, die Geschwindigkeit der WDM-Kanäle an die höhere der beiden Übertragungsgeschwindigkeiten anzupassen. Die erhöhte Übertragungsgeschwindigkeit kann einfach über die Einführung des konstanten Multiplikators in den Nebenbedingungen (4.5) geschehen, die damit zu (4.15) werden. Damit wird dann implizit auch Grooming von der normalen auf die hohe Geschwindigkeit ermöglicht. Der neue XC-Typ wird in den Nebenbedingungen (4.17) eingeführt. Das um die neue Übertragungsschicht und einen zusätzlichen XC-Typ erweiterte Modell sieht dann wie folgt aus:

**Eingabedaten - Gerätekosten:**

$C_e^F$	Kosten einer Verbindung normaler Geschwindigkeit (z.B. 2,5 Gbit/s SDH) auf Kante $e$ (Aufpreis für Portkarten langer Reichweite, Verstärker, Regeneratoren)
$C_e^{F'}$	Kosten einer Verbindung erhöhter Geschwindigkeit (z.B. 10 Gbit/s SDH) auf Kante $e$ (Aufpreis für Portkarten langer Reichweite, Verstärker, Regeneratoren)
$C_e^W$	Kosten einer WDM-Strecke auf Kante $e$ (Multiplexer, Verstärker)
$C^{OS}$	Kosten eines kleinen Cross-Connect-Basisgerätes
$C^{OL}$	Kosten eines großen Cross-Connect-Basisgerätes
$C^C$	Kosten für einen WDM-Kanal (Transponderpaar)
$C^P$	Kosten für eine Portkarte (normale Geschwindigkeit, kurze Reichweite) am Cross-Connect
$C^{P'}$	Kosten für eine Portkarte (erhöhte Geschwindigkeit, kurze Reichweite) am Cross-Connect

**Eingabedaten - Kapazitäten:**

$M^W$	Kapazität eines WDM-Systems (Anzahl der Wellenlängen)
$M^{OS}$	Kapazität eines kleinen Cross-Connects (Anzahl der Ports)
$M^{OL}$	Kapazität eines großen Cross-Connects (Anzahl der Ports)
$g_e$	Anzahl freier WDM-Kanäle auf Kante $e$ (Restkapazität des vorhandenen Netzes)
$h_n$	Anzahl freier Cross-Connect Ports in Knoten $n$ (Restkapazität des vorhandenen Netzes)

**Entscheidungsvariablen:**

$f_e$	Anzahl der Basisverbindungen normaler Geschwindigkeit (z.B. 2,5 Gbit/s SDH) auf Kante $e$
$f'_e$	Anzahl der Basisverbindungen erhöhter Geschwindigkeit (z.B. 10 Gbit/s SDH) auf Kante $e$
$w_e$	Anzahl der WDM-Systeme auf Kante $e$
$v_e$	Anzahl der belegten Kanäle in den WDM-Systemen auf Kante $e$
$y_n^S$	Anzahl der kleinen Cross-Connects in Knoten $n$
$y_n^L$	Anzahl der großen Cross-Connects in Knoten $n$
$z_{ij}^{st}$	1 wenn der Bedarf $(s, t)$ über die Kante $(i, j)$ geroutet wird; sonst 0

**Zielfunktion:**

$$\begin{aligned} \text{Minimiere } & \sum_{e \in E} ((C_e^F + 2C^P)f_e + (C_e^{F'} + 2C^{P'})f'_e + C_e^W w_e + (C^C + 2C^P)v_e) \\ & + \sum_{n \in N} (C^{OS}y_n^S + C^{OL}y_n^L) \end{aligned}$$

u.d.N.:

$$\sum_{j \in N} z_{ji}^{st} - \sum_{j \in N} z_{ij}^{st} = \begin{cases} -1 & i = s \\ 0 & \forall i \neq s, t \\ +1 & i = t \end{cases} \quad \forall (s, t) \in D \quad (4.14)$$

$$\sum_{(s,t) \in D} d_{st}(z_{ij}^{st} + z_{ji}^{st}) \leq 4*(v_e + f'_e) + f_e \quad \forall e \in E \text{ mit } i \text{ und } j \text{ angrenzend an } e \quad (4.15)$$

$$v_e \leq M^W w_e + g_e \quad \forall e \in E \quad (4.16)$$

$$\sum_{e \text{ angrenzend an } n} (v_e + f_e + f'_e) \leq M^{OS}y_n^S + M^{OL}y_n^L + h_e \quad \forall n \in N \quad (4.17)$$

$$\sum_{e \in E} z_e^{st} \leq X \quad \forall (s, t) \in D \quad (4.18)$$

$$y_n^S, y_n^L \geq 0 \text{ und ganzzahlig} \quad \forall n \in N \quad (4.19)$$

$$f_e, f'_e, w_e, v_e \geq 0 \text{ und ganzzahlig} \quad \forall e \in E \quad (4.20)$$

$$z_{ij}^{st} \in \{0, 1\} \quad \forall (i, j) \in E', (s, t) \in D \quad (4.21)$$

**Stellplatzbegrenzungen**

Möglicher Weise kann eine aus Kostenaspekten optimale Netzkonfiguration nicht gewählt werden, weil andere, im Rahmen dieser Planung harte, Randbedingungen dem entgegenstehen. Die einzelnen Netzknoten verfügen mitunter über sehr unterschiedlich dimensionierte Stellflächen für die Cross-Connects. So kann der Fall eintreten, dass ein aus planerischer Sicht sehr günstiger Knoten nicht die gewünschte Menge an Transitverkehr führen kann, weil er in der Realität nur eine kleine Außenstelle mit wenig Stellfläche darstellt. Da

für die Standorte oft langfristige Immobilienverträge bestehen, ist es auch im Rahmen einer taktischen oder strategischen Planung durchaus gerechtfertigt, diese Einschränkungen in erster Instanz als harte Randbedingungen anzusehen.

Neben der Stellfläche stellt auch die von den Geräten produzierte Wärme eine zunehmende Einschränkung dar. Moderne Geräte sind mitunter sehr kompakt gebaut, erzeugen aber eine so starke Abwärme, dass sie nicht so dicht gestellt werden können wie Geräte älteren Typs. Deshalb benötigen sie teilweise effektiv deutlich mehr Platz, als es von der Bauform her den Anschein hat. Restriktionen dieser oder ähnlicher Art können über eine Begrenzung der zulässigen Werte für die einzelnen  $y_n$  Variablen in das Modell aufgenommen werden. Komplexere Zusammenhänge für mehrere Gerätetypen können auch als zusätzliche Nebenbedingungen in das Modell integriert werden, beispielsweise in der Form  $y_n^S + 3 \cdot y_n^L \leq 16$  für eine Platzbegrenzung von 16 Raumeinheiten und einen Platzbedarf von einer Raumeinheit für einen kleinen XC und drei Raumeinheiten für einen großen XC.

### 1:1 Path Protection

Die Basisidee bei 1:1 Protection ist, wie in Abschnitt 1.2.4 beschrieben, Ressourcen (und damit Geld) einzusparen, indem Reservekapazitäten im Normalbetrieb für Verkehr niedriger Priorität genutzt werden. Im Fehlerfall würde dieser Verkehr dann aus dem Netz verdrängt und die Reservekapazität genutzt, um den höher priorisierten Verkehr zu transportieren.

Der Verkehr hoher Priorität wird auf einem einzelnen Weg von der Quelle zur Senke geführt und besitzt einen kantendisjunkten Ersatzweg, der ihm explizit zugeordnet ist. Solange der Ersatzweg nicht benötigt wird, kann die dort vorhandene Kapazität, auch abschnittsweise, von beliebigen Bedarfen niedriger Priorität belegt werden.

Für eine solche Problemstellung soll jetzt eine vereinfachte Version des Basismodells, erweitert um Bedarfe niedriger Priorität, vorgestellt werden. Genau wie beim Basismodell bildet  $N$  die Menge aller Knoten,  $E$  die Menge aller Kanten und  $D$  die Menge aller Bedarfe. Zusätzlich wird eine Menge  $D^l$  eingeführt, die die Bedarfe niedriger Priorität  $d_{st}^l$  enthält. Diese werden ggf. unterbrochen, wenn ein Bedarf hoher Priorität im Fehlerfall die entsprechenden Ressourcen für seinen Ersatzpfad benötigt.



**Entscheidungsvariablen:**

$f_e$	Gesamtkapazität auf Kante $e$
$za_{ij}^{st}$	1 wenn der Erstweg von Bedarf $(s, t)$ über die Kante $(i, j)$ führt; sonst 0
$zb_{ij}^{st}$	1 wenn der Ersatzweg von Bedarf $(s, t)$ über die Kante $(i, j)$ führt; sonst 0
$zl_{ij}^{st}$	1 wenn der Bedarf niederer Priorität $(s, t)$ über die Kante $(i, j)$ geführt wird; sonst 0
$l_e$	Hilfsvariable, die das Maximum aus der für Bedarfe niederer Priorität benötigten Kapazität und der für Ersatzwege benötigten Kapazität auf Kante $e$ bildet

**Zielfunktion:**

$$\text{Minimiere } \sum_{e \in E} f_e$$

u.d.N.:

$$\sum_{(s,t) \in D} d_{st}(za_{ij}^{st} + za_{ji}^{st}) + l_e \leq f_e \quad \forall e \in E \text{ mit } i \text{ und } j \text{ angrenzend an } e \quad (4.22)$$

$$\sum_{j \in N} za_{ji}^{st} - \sum_{j \in N} za_{ij}^{st} = \begin{cases} -1 & i = s \\ 0 & \forall i \neq s, t \\ +1 & i = t \end{cases} \quad \forall (s, t) \in D \quad (4.23)$$

$$\sum_{j \in N} zb_{ji}^{st} - \sum_{j \in N} zb_{ij}^{st} = \begin{cases} -1 & i = s \\ 0 & \forall i \neq s, t \\ +1 & i = t \end{cases} \quad \forall (s, t) \in D \quad (4.24)$$

$$\sum_{j \in N} zl_{ji}^{st} - \sum_{j \in N} zl_{ij}^{st} = \begin{cases} -1 & i = s \\ 0 & \forall i \neq s, t \\ +1 & i = t \end{cases} \quad \forall (s, t) \in D^l \quad (4.25)$$

$$\sum_{j \in N} (za_{ji}^{st} + zb_{ji}^{st}) \leq 1 \quad \forall i \neq s, t \quad \forall (s, t) \in D \quad (4.26)$$

$$l_e \geq \sum_{(s,t) \in D} d_{st}(zb_{ij}^{st} + zb_{ji}^{st}) \quad \forall e \in E \quad (4.27)$$

$$l_e \geq \sum_{(s,t) \in D} d_{st}^l(zl_{ij}^{st} + zl_{ji}^{st}) \quad \forall e \in E \quad (4.28)$$

$$f_e, l_e \geq 0 \text{ und ganzzahlig} \quad \forall e \in E \quad (4.29)$$

$$za_{ij}^{st}, zb_{ij}^{st} \in \{0, 1\} \quad \forall (i, j) \in E', (s, t) \in D \quad (4.30)$$

$$zl_{ij}^{st} \in \{0, 1\} \quad \forall (i, j) \in E', (s, t) \in D^l \quad (4.31)$$

Die Zielfunktion dieses vereinfachten Modells ist die Minimierung der Gesamtkapazität. Die Kapazitätsbedingungen (4.22) summieren die für die Erstwege benötigten Kapazitäten zusammen mit der Hilfsvariable  $l_e$  auf, um die Gesamtkapazität jeder Kante zu beschreiben. Abgesehen von der Aufteilung in die verschiedenen Bedarfe unterscheiden sich die Flussformulierungen (4.23) bis (4.25) nicht von denen des Basismodells. Die Nebenbedingungen (4.23) und (4.24) stellen die Flusserhaltung für die Erstwege und die Ersatzwege der Bedarfe hoher Priorität sicher. Die Aufteilung in zwei getrennte Nebenbedingungssysteme im Vergleich zum Basismodell wird notwendig, weil nur die Ersatzwegkapazität von den Bedarfen niedriger Priorität genutzt werden darf. Die Nebenbedingungen (4.25) sind die Flusserhaltungsbedingungen für die Bedarfe niedriger Priorität.

Die Nebenbedingungen (4.26) stellen, wie bereits beim Modell für 1+1 Path Protection, die knotendisjunkte Wegführung sicher. Die Nebenbedingungen (4.27) und (4.28) dienen zur Bestimmung der Hilfsvariablen  $l_e$ . Sie berechnen für jede Kante jeweils das Maximum aus der für Bedarfe niedriger Priorität benötigten Kapazität und der für Ersatzwege benötigten Kapazität. Die Nebenbedingungen (4.29) bis (4.31) legen die Ganzzahligkeit der Entscheidungsvariablen fest.

## Restoration

Unter dem Begriff Restoration werden viele verwandte, aber doch deutlich unterschiedliche Schutzmechanismen subsummiert, vgl. Abschnitt 1.2.4. In diesem Abschnitt wird ein ganzzahliges Modell für dynamisches Restoration vorgestellt. Es werden also keine festen Ersatzwege berechnet, sondern es wird lediglich sichergestellt, dass bei jedem beliebigen einzelnen Kantenausfall genug Kapazität im Netz vorhanden ist, um die ausgefallenen Bedarfe anderweitig führen zu können. Diese Reservekapazität kann sich dabei auf beliebig viele Wege verteilen. Bei einem Doppelfehler kann es möglich sein, dass genug Kapazitäten zum Beheben beider Fehler vorhanden sind. Dies ist jedoch nicht in der Zielfunktion verankert und damit in keiner Weise garantiert.

Die Eingabedaten, die Entscheidungsvariablen, die Zielfunktion und die Nebenbedingungen sind zum großen Teil identisch mit denen des Basis-

modells. Hinzu kommen einige Variablen und Nebenbedingungen, um den Restoration-Fluss abzubilden.

**Zusätzliche Variablen:**

$r_{ij}^{kl}$	Restoration-Fluss über die Kanten $(k, l)$ , wenn die Kante $(i, j)$ ausfällt
$o_{ij}$	Hilfsvariable, die den gesamten primären Fluss über die Kante $(i, j)$ widerspiegelt
$b_{ij}$	Hilfsvariable, die den gesamten Restoration-Fluss über die Kante $(i, j)$ widerspiegelt

**Zielfunktion:**

$$\text{Minimiere } \sum_{e \in E} ((C_e^F + 2C^P)f_e + C_e^W w_e + (C^C + 2C^P)v_e) + \sum_{n \in N} C^{OS} y_n^S$$

u.d.N.:

$$\sum_{j \in N} z_{ji}^{st} - \sum_{j \in N} z_{ij}^{st} = \begin{cases} -1 & i = s \\ 0 & \forall i \neq s, t \\ +1 & i = t \end{cases} \quad \forall (s, t) \in D \quad (4.32)$$

$$o_{ij} + b_{ij} \leq v_e + f_e \quad \forall e \in E \text{ mit } i \text{ und } j \text{ angrenzend an } e \quad (4.33)$$

$$o_{ij} = \sum_{(s,t) \in D} d_{st}(z_{ij}^{st} + z_{ji}^{st}) \quad \forall (i, j) \in E \quad (4.34)$$

$$b_{kl} \geq r_{ij}^{kl} + r_{ij}^{lk} \quad \forall (i, j) \in E, \quad \forall (k, l) \in E \quad (4.35)$$

$$\sum_{(i,j) \in E, i \neq k, j \neq k} r_{ij}^{ik} = o_{ij} \quad \forall (i, j) \in E' \quad (4.36)$$

$$\sum_{(i,j) \in E, i \neq k, j \neq k} r_{ij}^{ki} = 0 \quad \forall (i, j) \in E' \quad (4.37)$$

$$\sum_{(i,j) \in E, i \neq k, j \neq k} r_{ij}^{kj} = o_{ij} \quad \forall (i, j) \in E' \quad (4.38)$$

$$\sum_{(i,j) \in E, i \neq k, j \neq k} r_{ij}^{jk} = 0 \quad \forall (i, j) \in E' \quad (4.39)$$

$$\sum_{(n,k) \in E} r_{ij}^{lk} - \sum_{(k,n) \in E} r_{ij}^{kl} \quad \forall (i, j) \in E, \forall n \in N, n \neq i \text{ und } n \neq j \quad (4.40)$$

$$v_e \leq M^W w_e + g_e \quad \forall e \in E \quad (4.41)$$

$$\sum_{e \text{ angrenzend an } n} (v_e + f_e) \leq M^{OS} y_n^S + h_e \quad \forall n \in N \quad (4.42)$$

$$y_n^S \geq 0 \text{ und ganzzahlig} \quad \forall n \in N \quad (4.43)$$

$$f_e, w_e, v_e, o_e, b_e \geq 0 \text{ und ganzzahlig} \quad \forall e \in E \quad (4.44)$$

$$z_{ij}^{st} \in \{0, 1\} \quad \forall (i, j) \in E', (s, t) \in D \quad (4.45)$$

$$r_{ij}^{kl} \geq 0 \quad \forall (i, j) \in E', (k, l) \in E \quad (4.46)$$

Die Zielfunktion entspricht genau der des Basismodells, das Gleiche gilt für die Flusserhaltungsbedingungen (4.32). Für die Kapazitätsbedingungen (4.33) muss nun neben der Kapazität für den Erstweg der Bedarfe  $o_{ij}$  auch noch die Reservekapazität für den Fehlerfall  $b_{ij}$  berücksichtigt werden. Die Nebenbedingungen (4.34) und die Variablen  $o_{ij}$ , welche die Erstwegkapazität definieren, dienen ausschließlich der Übersichtlichkeit. Der entsprechende Term auf der rechten Seite könnte auch direkt in die anderen Nebenbedingungen eingefügt werden. Die in den Nebenbedingungen (4.35) definierte Backup-Kapazität muss für jede Kante  $(i, j)$  gleich dem größten Fluss sein, der im Fehlerfall auf diese Kante umgeleitet würde. Die Menge der möglichen Fehlerfälle besteht in diesem Modell aus der Menge der Kantenausfälle.

Die Nebenbedingungen (4.36) stellen sicher, dass für jede Kante  $(i, j)$  bei deren Ausfall auf den anderen an  $i$  angrenzenden Kanten genug Kapazität vorgesehen wird, um den Ausfall kompensieren zu können. Mit anderen Worten muss genug Kapazität vorhanden sein, um den ursprünglich von  $i$  nach  $j$  fließenden Verkehr auf anderen Kanten vom Knoten  $i$  wegzubefördern. Gleiches stellen die Nebenbedingungen (4.38) analog für den ankommenden Verkehr am anderen Ende der Kante  $(i, j)$  sicher. Die Flüsse in Gegenrichtung müssen jeweils auf Null fixiert werden, um eine doppelte Zählung zu vermeiden (vgl. Nebenbedingungen (4.37) und (4.39)). Weiterhin muss für den verbleibenden Teil des Netzes, also alle Knoten, die nicht direkt an die ausfallende Kante  $(i, j)$  angrenzen, dafür gesorgt werden, dass der umgeleitete Fluss nicht im Netz verloren geht. Dies wird durch die Nebenbedingungen (4.40) erreicht.

Die Nebenbedingungen für WDM-Systeme und Cross-Connects, (4.41) und (4.42), entsprechen von der Idee her wieder dem Basismodell, die Ganzzahligkeits- und Nichtnegativitätsbedingungen (4.43) bis (4.46) unter Berücksichtigung der neuen Variablen ebenso.

### 4.1.3 Diversifizierte Lösungen

Alle bisher vorgestellten Modelle führen bei ihrer Lösung zu einem einzigen Netzdesign für eine gegebene Probleminstanz. Jedoch gibt es in der Praxis, wie in Greistorfer *et al.* (2005) allgemein beschrieben, eine Reihe guter Gründe, mehrere verschiedene Lösungen erhalten zu wollen. So haben bei der Planung von Telekommunikationsnetzen die Betreiber des Netzes oft bestimmte Vorstellungen, wie ein „gutes“ Netz aussehen sollte. Solange der mathematisch berechnete monetäre Nachteil nur gering ist, wird eine übersichtliche, dem planenden und betreibenden Personal sinnvoller erscheinende Netzstruktur wahrscheinlich einer nicht intuitiv begreifbaren, aber dem Modell nach etwas kostengünstigeren, Netzstruktur vorgezogen werden. Unabhängig davon bildet jedes Modell natürlich nur einen Teil der Realität ab, und viele Planer haben gerne eine Reihe verschiedener guter Lösungen, aus der sie dann manuell nach mathematisch nur schwer beschreibbaren Kriterien eine zu realisierende Lösung auswählen. Auch aus akademischer Sicht kann es interessant sein zu analysieren, wie sich verschiedene gute Lösungen unterscheiden und was eine gute Lösung letztendlich charakterisiert. Zwei mögliche Ansätze, diversifizierte Lösungen zu erzeugen, ein sequentieller und ein paralleler, werden im Folgenden mit dem oben vorgestellten ganzzahligen Modell beispielhaft umgesetzt.

Eine sehr wichtige Frage ist in diesem Zusammenhang natürlich, wie sich der Unterschied (oder Abstand) zwischen zwei Lösungen sinnvoll messen lässt. Das oben vorgestellte Basisproblem hat eine Reihe verschiedener Entscheidungsvariablen, die dafür in Frage kommen. Deshalb ist die Definition des Abstandes nicht so offensichtlich wie z.B. beim klassischen Knapsack-Problem. Sowohl die Variablen, die das Routing beschreiben, als auch die Variablen für die Geräteplanung der Kanten sind prinzipiell von Interesse. Allerdings sind diese sehr stark voneinander abhängig, so dass es aus Anwendungssicht unproblematisch ist, sich lediglich auf eine der beiden Möglichkeiten zu beschränken. In diesem Fall scheinen die Routingvariablen aus Modellierungssicht geeigneter. Durch den binären Charakter ist der Unterschied zwischen zwei Lösungen deutlich einfacher zu bestimmen, als dies mit Hilfe der ganzzahligen Gerätevariablen möglich wäre, die bei Einsatz unterschiedlicher Technologien und Bandbreiten auch noch entsprechend ge-

geneinander gewichtet werden müssten. Die Entscheidungsvariablen für die Anzahl der Cross-Connects sind in diesem Zusammenhang nur von untergeordneter Bedeutung. Jeder Knoten, der einen Quellbedarf hat, benötigt sowieso mindestens einen Cross-Connect. Die Unterschiede in der Cross-Connect-Installation zwischen verschiedenen Lösungen sind im Allgemeinen eher gering.

Die meisten verwendeten Bezeichner sind identisch mit denen im Basismodell. Lediglich der Gewichtungsfaktor  $\omega$  kommt neu hinzu. Er beeinflusst das Verhältnis zwischen Kosten und Distanz in der Zielfunktion. Außerdem wird ein neuer Satz von binären Entscheidungsvariablen  $dis_{ij}^{st}$  eingeführt, der angibt, ob die Belegung der Kante  $ij$  durch Bedarf  $st$  in den beiden Lösungen identisch ist oder nicht.

Beim sequentiellen Ansatz wird zunächst mit Hilfe des bekannten Basismodells eine Referenzlösung erzeugt. Das Routing dieser Lösung dient dann als Eingangsdaten für das sequentielle Modell, mit dem dann eine zweite, in Abhängigkeit von  $\omega$  mehr oder weniger stark abweichende Lösung errechnet wird. Die Zielfunktion (4.1.3) minimiert die Kosten der neuen Lösung abzüglich der Distanz zur alten Lösung. Das entsprechende sequentielle Modell dazu sieht wie folgt aus:

**Kosten und Kapazitäten:**

$C_e^F$	Kosten einer Basisverbindung (z.B. 2,5 Gbit/s SDH) auf Kante $e$ (Aufpreis für Portkarten langer Reichweite, Verstärker, Regeneratoren)
$C_e^W$	Kosten einer WDM-Strecke auf Kante $e$ (Multiplexer, Verstärker)
$C^C$	Kosten für einen WDM-Kanal (Transponderpaar)
$M^W$	Kapazität eines WDM-Systems (Anzahl der Wellenlängen)
$C^{OS}$	Kosten eines Cross-Connect-Basisgerätes
$M^{OS}$	Kapazität eines Cross-Connects (Anzahl der Ports)

**Eingangsdaten aus der Basislösung:**

$zp_{ij}^{st}$	Routingvariablen; 1 wenn Bedarf $(s, t)$ die Kante $(i, j)$ nutzt; sonst 0
----------------	--

**Entscheidungsvariablen:**

$f_e$	Anzahl der Basisverbindungen (z.B. 2,5 Gbit/s SDH) auf Kante $e$
$w_e$	Anzahl der WDM-Systeme auf Kante $e$
$v_e$	Anzahl der belegten Kanäle in den WDM-Systemen auf Kante $e$
$y_k^S$	Anzahl der Cross-Connects in Knoten $n$
$z_{ij}^{st}$	1 wenn Bedarf $(s, t)$ die Kante $(i, j)$ nutzt; sonst 0
$dis_{ij}^{st}$	1 wenn sich der Wert von $z_{ij}^{st}$ vom Wert von $zp_{ij}^{st}$ unterscheidet; sonst 0

**Zielfunktion:**

$$\text{Minimiere } \sum_{e \in E} (C_e^F f_e + C_e^W w_e + C^C v_e) + \sum_{k \in N} C^{OS} y_k^S - \omega \sum_{\substack{(i,j) \in E' \\ (s,t) \in D}} dis_{ij}^{st}$$

u.d.N.:

$$\sum_{j \in N} z_{ji}^{st} - \sum_{j \in N} z_{ij}^{st} = \begin{cases} -1 & i = s \\ 0 & \forall i \neq s, t \\ +1 & i = t \end{cases} \quad \forall (s, t) \in D \quad (4.47)$$

$$\sum_{(s,t) \in D} d_{st} (z_{ij}^{st} + z_{ji}^{st}) \leq v_e + f_e \quad \forall e \in E \text{ mit } i \text{ und } j \text{ angrenzend an } e \quad (4.48)$$

$$v_e \leq M^W w_e \quad \forall e \in E \quad (4.49)$$

$$\sum_{e \text{ angrenzend an } k} (v_e + f_e) \leq M^{OS} y_k^S \quad \forall k \in N \quad (4.50)$$

$$dis_{ij}^{st} \geq zp_{ij}^{st} - z_{ij}^{st} \quad \forall (s, t) \in D, (i, j) \in E' \quad (4.51)$$

$$dis_{ij}^{st} \geq z_{ij}^{st} - zp_{ij}^{st} \quad \forall (s, t) \in D, (i, j) \in E' \quad (4.52)$$

$$dis_{ij}^{st} \leq zp_{ij}^{st} + z_{ij}^{st} \quad \forall (s, t) \in D, (i, j) \in E' \quad (4.53)$$

$$dis_{ij}^{st} \leq 2 - zp_{ij}^{st} - z_{ij}^{st} \quad \forall (s, t) \in D, (i, j) \in E' \quad (4.54)$$

$$y_k^S \geq 0 \text{ und ganzzahlig} \quad \forall k \in N \quad (4.55)$$

$$f_e, w_e, v_e \geq 0 \text{ und ganzzahlig } \forall e \in E \quad (4.56)$$

$$z_{ij}^{st} \in \{0, 1\} \quad \forall (i, j) \in E', (s, t) \in D \quad (4.57)$$

In der Zielfunktion werden, wie im Basismodell, die Gesamtkosten minimiert. Zusätzlich wird hier die gewichtete Distanz abgezogen; je größer die Distanz, desto besser ist die Lösung. Die Nebenbedingungen (4.47) bis (4.50) sind die bekannten Flussformulierungen und Kapazitätsbedingungen aus dem Basismodell. Neu hinzugekommen sind die Nebenbedingungen (4.51) bis (4.54), welche zur Berechnung des Abstandes dienen und der Idee aus Greistorfer *et al.* (2005) folgend linearisiert sind.

Beim parallelen Ansatz werden beide Lösungen simultan aus einem umfassenden Modell heraus berechnet. Dieses Modell ist dem Basismodell sehr ähnlich, alle Entscheidungsvariablen und Nebenbedingungen (4.47) bis (4.50) existieren jetzt doppelt, für jede Lösung ein Mal. Die zusätzlichen Nebenbedingungen (4.51) bis (4.54) für die Abstandsberechnung sind identisch zum sequentiellen Modell. Allerdings werden die hier enthaltenen Routingvariablen diesmal im Modell selber generiert und nicht wie beim sequentiellen Ansatz als im Voraus berechnete Daten von außen eingelesen. Die Zielfunktion (4.58) minimiert die Summe aus den beiden Kostenwerten der Lösungen und subtrahiert davon deren Distanz. Während beim sequentiellen Ansatz also eine gute (oder ggf. sogar optimale) Lösung und eine nachrangige, etwas schlechtere Lösung gesucht wird, sind hier beide Lösungen gleichrangig.

$$\text{Minimiere } \text{Gesamtkosten } a + \text{Gesamtkosten } b - \omega \sum_{\substack{(i,j) \in E' \\ (s,t) \in D}} dis_{ij}^{st} \quad (4.58)$$

Die Auswirkungen auf Laufzeit und Lösungsgüte der beiden Ansätze werden in Abschnitt 5.1 im Rahmen der Ergebnisse des Basismodells diskutiert. Ein grundlegendes Problem dieser Modellierung bleibt jedoch unabhängig von der steigenden Rechenzeit bestehen: Die diversifizierten Lösungen sind oft für die Praxis nicht brauchbar. Der Grund dafür liegt darin, dass der Unterschied in den Lösungen oft aus mehreren kurzen Kreisen besteht, die unabhängig vom eigentlichen Pfad des Bedarfes entstehen (vgl. Abbildung 4.2).

Während solche Kreise im Basismodell noch implizit über die daraus resultierenden Kosten vermieden werden, kann dies bei beiden Modellen für diversifizierte Lösungen durch die daraus resultierende Zunahme der Distanz kompensiert werden. Streng der hier gewählten Definition der Distanz als Unterschied der Routingvariablen nach ist dies sogar richtig. In der Praxis haben solche Lösungen aber natürlich keinen Wert.



Über einfache Ergänzungen der Flussformulierungen lässt sich dieses Problem nur abmildern, aber nicht vollständig beseitigen. Dafür wäre ein eigener Satz an speziellen Nebenbedingungen zur Vermeidung von Kreisen notwendig. Dies ist im Prinzip nichts Ungewöhnliches, Nebenbedingungen zum Verhindern von Kreisen oder sogenannten Kurzzyklen sind auch bei anderen Problemen durchaus gängig, so z.B. beim Travelling Salesman-Problem (vgl. hierzu z.B. die grundlegende Arbeit von Dantzig *et al.* (1954)). Jedoch hat dies im hier betrachteten Kontext mit Sicherheit weiteren negativen Einfluss auf die Laufzeit und damit die Lösbarkeit. Zumindest für den sequentiellen Ansatz erscheint es aber lohnend, dieses in Zukunft genauer zu untersuchen.

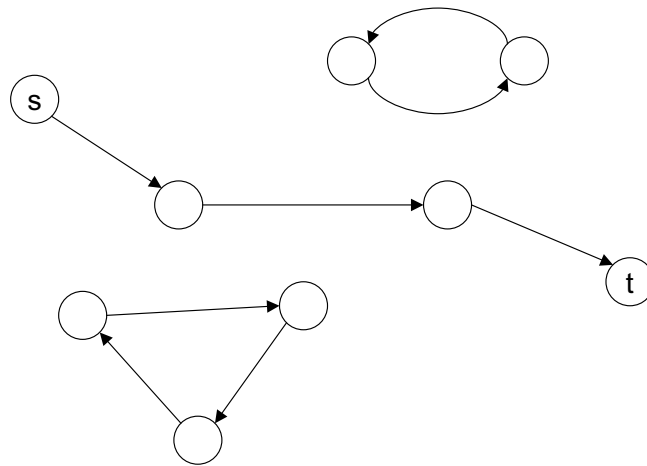


Abbildung 4.2: Beispiel für das Kreisproblem.

#### 4.1.4 Relaxationen

Relaxationen, also das Aufweichen von Nebenbedingungen, sind ein gängiges Vorgehen, beispielsweise um Schranken zu berechnen. Bei den im Rahmen dieser Arbeit durchgeführten Berechnungen wird zum Teil Gebrauch von der sogenannten linearen Relaxation gemacht. Dies bedeutet, dass die Ganzzahligkeitsbedingungen bestimmter Variablen aufgehoben werden. Damit wird das ganzzahlige Basismodell dann zu einem gemischt ganzzahligen Modell. Die dadurch erhaltene Lösung ist normalerweise keine zulässige Lösung für das Ausgangsproblem, bildet jedoch eine untere Schranke. Das heißt, eine optimale Lösung des Ausgangsproblems kann nicht besser sein als eine optimale Lösung des linear relaxierten Problems. Der große Vorteil der Relaxation liegt in der deutlich geringeren Rechenzeit, die für die Lösung benötigt wird.

Da das Basismodell eine ganze Reihe von ganzzahligen Variablen enthält, stellt sich die Frage, welche davon sinnvoller Weise relaxiert werden sollten, damit einerseits die Rechenzeit möglichst gering wird und andererseits eine möglichst gute Schranke gefunden wird. Mit anderen Worten die optimale Lösung des relaxierten Modells soll möglichst nah an der (häufig nicht bekannten) optimalen Lösung des ganzzahligen Modells liegen.

Die beiden Entscheidungsvariablen, die sich direkt für diesen Zweck anbieten, sind die Anzahl der Cross-Connects  $y_n^S$  und die Anzahl der WDM-Systeme  $w_e$ . Die Tatsache, dass diese Geräte jeweils nur in fixen Modulgrößen installiert werden können, trägt wesentlich zur Komplexität des Modells bei und folglich führt eine Relaxierung zu einer deutlichen Reduzierung des Rechenaufwandes. Anschaulich bedeutet eine lineare Relaxierung dieser Variablen, dass ein Basisgerät nur anteilig mit seiner tatsächlich belegten Kapazität berechnet wird. Ein WDM-System, welches theoretisch 40 Kanäle führen kann, in dem aber nur 10 belegt sind, würde also nur mit einem Viertel seiner Kosten veranschlagt. Die mögliche Abweichung von der ganzzahligen Lösung kann also umso größer sein, je größer die Modulgröße ist. Gerade bei sehr teuren Systemen mit großer Kapazität kann diese Abweichung beträchtlich werden. In den meisten von uns berechneten Probleminstanzen ist die Abweichung zur ganzzahligen Lösung bei einer Relaxierung der Cross-Connects deutlich geringer als bei einer Relaxierung der WDM-Systeme, da letztere in der Regel wesentlich teurer sind.

Die anderen ganzzahligen Entscheidungsvariablen sind aus verschiedenen Gründen nicht besonders gut für lineare Relaxationen geeignet. Oft ist die Basiseinheit der Bedarfe identisch zur Basiseinheit der Übertragungskapazität. In diesem Fall ist eine Relaxation der Anzahl der Kanäle mit Basiskapazität wirkungslos, die Ganzzahligkeit ist implizit trotzdem sichergestellt. Als Relaxation zur Berechnung von Schranken sind die  $f_e$  und  $v_e$  Variablen also ungeeignet. Jedoch verschiebt sich dadurch das Verhältnis zwischen Branch&Bound- und Simplex-Operationen, die CPLEX ausführt, was ebenfalls Einfluss auf die Rechenzeit haben kann, siehe Abschnitt 5.1. Eine Relaxation der Binärbedingung für die Routingvariablen  $z_{ij}^{st}$  ist ebenfalls eine Möglichkeit. Anschaulich würde ein zulässiges Intervall zwischen Null und Eins bedeuten, dass der Bedarf sich beliebig auf mehrere Wege aufteilen kann. Bei bestimmten Netzen im IP-Bereich ist dies ohnehin technologisch zulässig, bei SDH nicht so ohne weiteres. Eine solche Relaxation der Routingvariablen hat jedoch nur geringen (und nicht immer positiven) Einfluss auf die Laufzeit und ist für die Berechnung von Schranken deshalb ebenfalls nicht gut geeignet.

Darüber hinaus gibt es weitere Verfahren zur Relaxation, die ebenfalls eine weite Verbreitung gefunden haben, wie z.B. die sogenannte Lagrange-

Relaxation. Hier werden Nebenbedingungen aufgeweicht, indem sie gewichtet in die Zielfunktion aufgenommen werden. Die hier gewählte lineare Relaxation hat aber neben ihrer Einfachheit den großen Vorteil, dass durch einfaches Aufrunden der relaxierten Variablen sofort eine zulässige Lösung zur Verfügung steht. Weiterhin hat sie sich von der Güte der Schranken und der Ersparnis an Rechenzeit her als zufriedenstellende Variante für die hier betrachteten Problemstellungen erwiesen. Deshalb werden andere Arten der Relaxation hier nicht weiter betrachtet.

Aus bestimmten Relaxationen kann sich auch die Möglichkeit ergeben, die Problemstellung in mehrere unabhängige Teilprobleme zerlegen zu können. Die so erzeugten Teilprobleme können dann unabhängig voneinander gelöst werden, was den Rechenaufwand normalerweise sehr stark reduziert. Die hier gewählten linearen Relaxationen erlauben eine solche Dekomposition jedoch nicht.

#### 4.1.5 Schnitte

Schnitte sollen hier als das Gegenteil von Relaxationen verstanden werden, also als eine Verstärkung der Modellformulierung durch zusätzliche Nebenbedingungen, die einen Teil des ursprünglichen Lösungsraumes abschneiden. Idealerweise sollte dabei nachweisbar keine optimale Lösung abgeschnitten werden. Bestimmte Bedingungen werden implizit durch die Minimierung des Zielfunktionswertes sichergestellt. Trotzdem können diese natürlich auch explizit als Nebenbedingungen in das Modell aufgenommen werden. Diese eigentlich redundanten Formulierungen tragen teilweise erheblich zu einer Vereinfachung der Problemstellung bei, weil der Solver nur noch einen deutlich kleineren Lösungsraum durchsuchen muss. Teilweise überwiegt aber auch der Nachteil, dass diese zusätzlichen Bedingungen das Modell natürlich größer machen.

Einige einfach verständliche, zusätzliche Nebenbedingungen lassen sich aus den Skaleneffekten der Gerätepreise ableiten. Wenn eine 10 Gbit/s Verbindung grundsätzlich günstiger ist als drei einzelne 2,5 Gbit/s Verbindungen, dann kann für die Anzahl der 2,5 Gbit/s Verbindungen auf jeder Kante eine obere Schranke von 2 angegeben werden. Ähnlich lässt sich die Gesamtzahl der diskreten SDH-Verbindungen oft beschränken, wenn ab einer bestimmten Kanalzahl ein WDM-System stets kostengünstiger ist. Auch die Längen der Übertragungsleitungen spielen hier eine Rolle. Teilweise kann SDH für bestimmte, sehr lange Kanten explizit ausgeschlossen werden. Dies alles sind Beispiele dafür, dass Teile des Lösungsraumes abgeschnitten werden können, für die man von vornherein nachweisen kann, dass sie keine optimale

Lösung enthalten können.

Ein anderes Beispiel für nur implizit enthaltene Nebenbedingungen sind die bereits in Abschnitt 4.1.3 bei den diversifizierten Lösungen angesprochenen Kreisbedingungen. Einen Bedarf auf einem Teilstück im Kreis zu führen, macht offensichtlich keinen Sinn. Die gewählte Flussformulierung verbietet dies jedoch nicht explizit. Allerdings wird eine gute Lösung trotzdem keine Kreise enthalten, da diese Kapazitäten belegen würden, die in der Zielfunktion Kosten induzieren. Durch die Minimierung des Zielfunktionswertes sind solche Lösungen also implizit ausgeschlossen. Trotzdem kann man natürlich zusätzliche Flussbedingungen formulieren, die explizit solche Kreise verbieten.

#### 4.1.6 Praktischer Einsatz der ganzzahligen Modelle

Zur Lösung konkreter Probleminstanzen wird das passende mathematische Modell in der Modellierungssprache AMPL formuliert. Gegenüber den oben vorgestellten Modellen können bzw. müssen ggf. Änderungen vorgenommen werden, um bestimmte spezielle Randbedingungen der jeweiligen Problemstellung zu berücksichtigen. Weiterhin werden alle notwendigen Eingangsdaten, wie z.B. die Bedarfe und die Netztopologie, für das Einlesen in AMPL aufbereitet. Ein Beispiel für ein einfaches AMPL-Modell und den Datensatz einer kleinen Probleminstanz findet sich in Anhang B. Alternativ zu dem dort gewählten ASCII-Datenformat können die Daten auch über die ODBC-Schnittstelle aus Excel eingelesen werden. Eine entsprechende Anleitung dazu findet sich im Internet auf den AMPL-Webseiten (AMPL Optimization LLC (2000)). Umfassende Informationen zu AMPL bietet das offizielle AMPL-Buch von Fourer *et al.* (2003).

Mit dem eigentlichen Solver, in unserem Fall CPLEX, wird dabei nicht direkt kommuniziert. Dieser wird von AMPL im Hintergrund aufgerufen und wenn die Berechnung beendet ist, dann gibt auch AMPL das Ergebnis wieder an den Nutzer zurück. Das Ergebnis enthält mindestens die ermittelten Werte für alle Entscheidungsvariablen und den Zielfunktionswert. Zusätzlich können noch diverse Statusinformationen zur Berechnung angezeigt werden, z.B. Rechenzeit, Anzahl der ausgeführten Simplex-Operationen, Größe des Branch&Bound-Baumes usw.

Wenn im ersten Schritt eine Relaxation berechnet wird, so kann das Ergebnis als Eingangsdaten für die ganzzahlige Berechnung übergeben werden. Mit diesem zweistufigen Ansatz ist es teilweise möglich, den Speicherbedarf stark zu reduzieren. Die relaxierte Lösung stellt eine untere Schranke dar, aus der durch Aufrunden eine zulässige Lösung ermittelt werden kann, die

wiederum eine obere Schranke darstellt. Somit können bei der ganzzahligen Berechnung Teile des Branch&Bound-Baumes, die schlechter sind als die beste bekannte obere Schranke, von vornherein abgeschnitten werden. Je nach Anwendungszweck müssen die als ASCII-Datei ausgegebenen Daten dann ggf. noch manuell in Form von Grafiken oder Tabellen aufbereitet werden. Das gesamte Vorgehen ist in Abbildung 4.3 noch einmal zusammenfassend dargestellt.

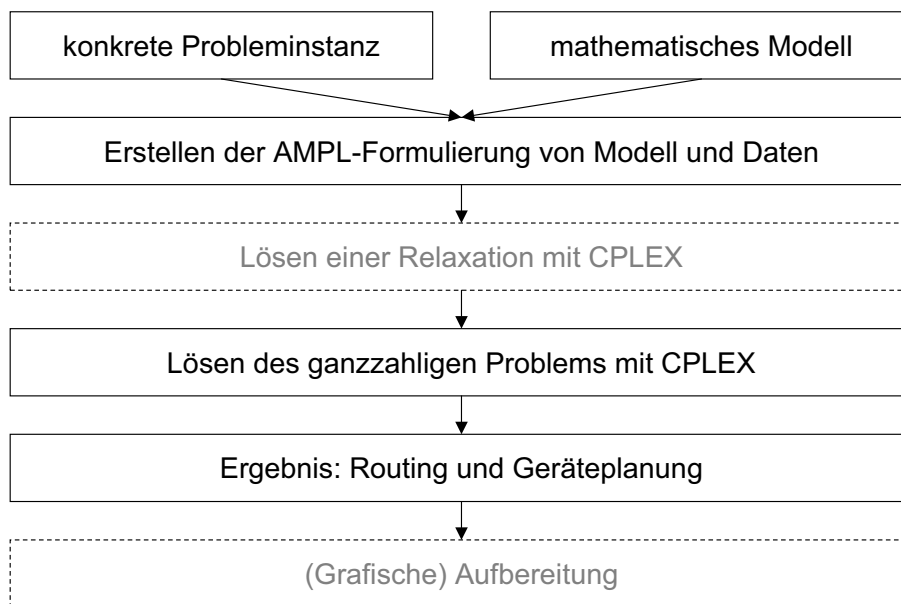


Abbildung 4.3: Grundsätzlicher Ablauf der Planung beim Einsatz von AMPL und CPLEX.

Mit den hier vorgestellten (gemischt) ganzzahligen Modellen ist es möglich, für kleine und mittlere Netze optimale Lösungen zu finden. Dabei ist das Modell sehr flexibel an vielfältige Randbedingungen realer Netze anpassbar, z.B. an Schutzverfahren, Gerätespezifikationen usw. Detaillierte Rechenergebnisse für verschiedene Problem Instanzen werden in Abschnitt 5.1 vorgestellt. Neben der exakten Lösung bietet sich dieser Ansatz auch an, um im Zusammenspiel mit Heuristiken die Lösungsgüte bewerten zu können oder sogar zu verbessern. Auf die Verknüpfung der ganzzahligen Modelle mit Heuristiken wird in Abschnitt 4.3 eingegangen.

## 4.2 Heuristiken

Wie bereits in Abschnitt 4.1 erwähnt, sind die hier betrachteten Problemstellungen so komplex, dass viele reale Probleminstanzen nicht mehr in akzeptabler Zeit mit exakten Verfahren gelöst werden können. Aus diesem Grund bietet sich der Einsatz von sogenannten Heuristiken bzw. Metaheuristiken an.

Der eigentliche Begriff „Heuristik“ leitet sich aus einem altgriechischen Verb ab: „heurisko = ich finde“. Im Allgemeinen wird ein Verfahren als Heuristik bezeichnet, welches versucht anhand bestimmter Regeln eine möglichst gute Lösung für ein Problem in relativ kurzer Zeit zu finden. Eine Heuristik kann also umgangssprachlich auch als Daumenregel bezeichnet werden. Polya (1957) (S. 113) benutzt in seinem wegweisenden Werk über Heuristiken in diesem Zusammenhang die Worte: „... a more or less plausible guess...“.

Eine Heuristik ist somit per Definition ein Näherungsverfahren und kann keine optimale Lösung des Problems garantieren. Die Betonung liegt dabei auf „garantieren“. In der Tat sind gute Heuristiken in der Praxis sehr wohl teilweise in der Lage, optimale Lösungen zu finden. Einen Beweis, dass eine Lösung optimal ist, kann eine Heuristik jedoch nicht erbringen. Darüber hinaus wird als Metaheuristik ein übergeordneter Mechanismus verstanden, der eine einfache Heuristik steuert.

Eine Beispiel für einen sehr einfache Heuristik ist eine gierige, lokale Suche. Es wird immer der Schritt als nächster gewählt, der unmittelbar den größten Profit verspricht. Für sich alleine genommen ist ein solches Vorgehen allerdings oft zu kurzfristig, als dass damit gute Resultate erzielt werden könnten. Es eignet sich jedoch sehr wohl als Baustein komplexerer Metaheuristiken.

Detaillierte Beschreibungen vieler aktueller Metaheuristiken, z.B. Greedy Randomized Adaptive Search Procedure (GRASP), Tabu Search und Simulated Annealing, finden sich in den einzelnen Beiträgen zu Glover und Kochenberger (2003). Auch Michalewicz und Fogel (2004) liefern konkrete Beschreibungen einzelner Metaheuristiken. Es handelt sich dabei aber vom Fokus her eher um eine generelle Einführung in Strategien zum heuristischen Lösen praktischer Probleme und liegt sowohl vom Inhalt als auch von der Intention her ganz in der Tradition von Polya (1957).

Die hier entwickelte Heuristik (Wenn im Folgenden von „Heuristik“ gesprochen wird, so ist in der Regel die spezielle, für diese Problemstellung entwickelte damit gemeint und nicht der Oberbegriff.) deckt weitestgehend die gleichen Funktionen ab wie das ganzzahlige Modell, also auch die funktionalen Erweiterungen wie 1+1 Path Protection, Grooming usw. Lediglich Hop-Beschränkungen und Stellplatzbegrenzungen sind nicht implemen-

tiert. Sie ermöglicht also ebenfalls eine gleichzeitige Planung der Hardware und des Routings, allerdings auch für große Netze mit bis zu 1000 Knoten. Auch das Format der Eingabedaten ist folglich dem in Abschnitt 4.1.1 für das ganzzahlige Modell beschriebenen sehr ähnlich. Darüber hinaus ist die Heuristik in der Lage, weitergehende Restorationverfahren zu unterstützen sowie das Vorbeischleifen voll ausgelasteter Glasfasern an Netzknoten zu berücksichtigen. Bei direkten Vergleichsrechnungen mit dem MIP-Modell werden diese zusätzlichen Funktionen natürlich abgeschaltet; bei den Fallstudien werden sie selbstverständlich genutzt, soweit sie der realen Aufgabenstellung nach erlaubt sind.

### 4.2.1 Greedy Randomized Adaptive Search Procedure

Die aktuelle Version der Heuristik hat sich Stück für Stück aus der in Höller und Voß (2006) als „random start“ bezeichneten Heuristik entwickelt. Diese enthält bereits einige zentrale Ideen, wie die grundlegende Suchstruktur und die Kostenfunktionen, ist aber noch sehr stark zufallsbasiert. Vor allem bei großen Probleminstanzen und den daraus resultierenden umfangreichen Lösungsräumen ist es jedoch unwahrscheinlich, dass ein deterministisches Startverfahren in Verbindung mit einer primär zufallsbasierten lokalen Suche eine gute oder sogar optimale Lösung in akzeptabler Zeit findet. Aus diesem Grund setzen wir eine sogenannte Greedy Randomized Adaptive Search Procedure ein, um die Suche in vielversprechende Bereiche des Lösungsraumes zu leiten. Gegenüber anderen metaheuristischen Strategien hat GRASP hier vor allem den Vorteil, dass es sich nahtlos in die alte Variante der Heuristik integrieren lässt. Es bietet somit die Möglichkeit, die Lösungsgüte erheblich zu verbessern, ohne das prinzipiell bewährte Grundkonzept zu verwerfen.

Eine GRASP-Metaheuristik besteht in der Regel aus vier Teilen, die wiederholt ausgeführt werden: einer gierigen Konstruktionsphase mit einer Zufallskomponente, einer lokalen Suche und einer adaptive Komponente, welche die gierige Phase im Laufe des Verfahrens modifiziert. Damit ist GRASP ein klassisches Multi-Start-Verfahren, vgl. Marti (2003). Eine breit gefächerte Einführung in die Grundprinzipien und in viele Varianten von GRASP findet sich in Pitsoulis und Resende (2001) oder Resende und Ribeiro (2003). Resende (2001) beinhaltet eine umfangreiche Bibliographie zu GRASP. Diese Quellen liefern auch vielfältige Hinweise auf Anwendung und weiterführende Literatur zu GRASP im Telekommunikationsbereich.

In der Literatur ist der Ablauf des GRASP meistens genau vertauscht gegenüber dem von uns gewählten Vorgehen. Die Zufallskomponente wird bei der Generierung der Startlösung eingesetzt, während die lokale Suche

dann oft deterministisch ist. Die Motivation, bei unserem Ansatz andersherum vorzugehen, begründet sich hauptsächlich aus der guten Qualität der Startlösung, die mit einem einfachen, deterministischen Kürzeste-Wege-Verfahren erzielt werden kann.

Das Ablaufdiagramm unserer an GRASP angelehnten Heuristik ist in Abbildung 4.4 zu sehen. Im ersten Schritt wird eine zulässige Startlösung erzeugt. Dies geschieht mit Hilfe von Dijkstras Kürzestem-Wege-Algorithmus (Dijkstra (1959)) auf Basis des gegebenen Glasfasergraphen. Suurballes Erweiterung für knotendisjunkte Pfade (vgl. Suurballe (1974) und Suurballe und Tarjan (1984)) wird genutzt, falls Ersatzwege, beispielsweise für 1+1 Protection, benötigt werden. Dieses Verfahren ermittelt das Wegepaar mit der kleinsten Summe der beiden Einzelwege. Aus Sicht der Anwendung ist dieses Vorgehen sinnvoll, da bei 1+1 Protection immer beide Pfade gleichzeitig genutzt werden. Somit würde es keinen Sinn machen, einen kürzeren Erstweg zu bevorzugen, wenn damit die Summe der beiden Wege länger würde. Für die Startlösung dient im Normalfall die geographische Länge der Glasfasern als Metrik. Zwei etwas abweichende Startverfahren, „Minimale Zahl an Hops“ und „Minimal-spannender-Baum“ (Engl. Minimum Spanning Tree, MST) sind ebenfalls implementiert, finden jedoch nur in Sonderfällen Anwendung. Wenn die Zahl der Hops minimiert werden soll, dann kann ebenfalls der Dijkstra-Algorithmus eingesetzt werden, jedoch mit uniformen Kantengewichten.

Für den Minimal-spannenden-Baum können die bekannten Algorithmen wie Kruskal, Prim-Dijkstra oder Sollin genutzt werden. Ersterer wird hier verwendet, ist allerdings nicht für Probleminstanzen mit geschützten Bedarfen geeignet, da in der aktuellen Implementierung nur ein einziger Baum aufgespannt wird. Obwohl die MST-Startlösung den kürzesten Wegen in den meisten uns bekannten Szenarien unterlegen ist, gibt es einige Spezialfälle, in denen sie zu guten Ergebnissen führt, insbesondere wenn sich gute Lösungen durch ein sehr weitmaschiges Glasfasernetz auszeichnen. Dies kann z.B. der Fall sein, wenn die Basiskosten, um eine Kante in die Lösung aufzunehmen, sehr hoch sind und deshalb relativ lange Umwege lohnenswert werden.

Zunächst wird, wie beschrieben, eine zulässige Startlösung für das Routing gesucht. Dazu werden dann ein detaillierter Geräteplan aufgestellt und die Gesamtkosten des resultierenden Netzes berechnet. Die Eingangsdaten für Kosten und Kapazitäten entsprechen genau denen, die in Abschnitt 4.1.1 für das ganzzahlige Modell beschrieben werden. Die so erzielte Startlösung (mit Ausnahme der MST-Variante) hat tendenziell relativ viele, eher gering ausgelastete Kanten. Sie dient als Ausgangspunkt für ein iteratives Verbesserungsverfahren, welches darauf abzielt, die in Abschnitt 3.2 dargestellten Economies of Scale auszunutzen, indem es die Wegführung auf immer weniger



---

*Wiederhole die folgende Sequenz n Mal:*

- (1) Berechne eine Startlösung;
  - kürzeste Wege (Distanz);
  - oder kürzeste Wege (Hops);
  - oder Minimal-spannender-Baum.
- (2) Berechne die resultierenden Gesamtkosten des Netzwerkes.
- (3) *Wiederhole die folgende Sequenz bis keine Verbesserung des Ergebnisses mehr erzielt wird:*
  - (a) Gierige Sortierung der Kandidatenliste.
  - (b) Zufällige Permutation des ersten Teiles der Kandidatenliste.
  - (c) Wiederhole für alle Bedarfe (bzw. Bedarfsbündel):
    - Rerouting des Bedarfes, falls dies laut Kostenfunktion vorteilhaft ist.*
  - (d) Berechne die resultierenden Gesamtkosten des Netzwerkes.

---

Abbildung 4.4: GRASP-Heuristik

Kanten bündelt.

Diese Bündelung erfolgt, indem die Bedarfe der Reihe nach neu geroutet werden. Dabei wird die genaue Reihenfolge durch ein zweistufiges Sortierverfahren, mit einer gierigen und einer zufallsgesteuerten Phase, festgelegt. Zuerst werden alle Bedarfe in absteigender Folge der Kosten, die ihr Routing anteilmäßig verursacht, sortiert. Die Idee hinter dieser Sortierung ist die folgende: Große Bedarfe sind, nicht bezogen auf das einzelne Bit, aber bezogen auf die Gesamtkosten, sehr teuer. Deshalb erscheint es normalerweise sinnvoller, diese Bedarfe so direkt wie möglich zu führen und bei kleineren, billigeren Bedarfen Umwege in Kauf zu nehmen, als andersherum.

Eine strikt gierige Strategie ist jedoch zu sehr eingeengt und wird alleine nur selten zu einem sehr guten Gesamtergebnis führen. Deshalb wird diese deterministische Sortierung der Kandidatenliste durch zufällige Vertauschungen im vorderen Bereich der Liste aufgelockert. Das Verhältnis zwischen der gierigen Sortierung und den zufälligen Vertauschungen kann in einem wei-

ten Bereich eingestellt werden. Die daraus resultierenden Effekte werden in Abschnitt 5.2 im Detail erläutert. Für alle Berechnungen im Rahmen dieser Arbeit wird jedoch eine feste Einstellung für dieses Verhältnis gewählt, die sich universell bewährt hat.

Ein Grundproblem solch eines sequentiellen Reroutingverfahrens folgt aus der Stufigkeit der realen Gerätekosten. Beispielsweise kann ein WDM-Terminal eine Kapazität von 40 Kanälen besitzen. Wenn 41 Kanäle benötigt werden, müsste ein komplettes, zusätzliches Terminal aufgestellt werden, was sich für einen einzelnen Kanal wohl nur in den seltensten Fällen lohnen würde. Ein sequentieller Algorithmus besitzt jedoch keine Kenntnis darüber, ob ggf. später weitere Bedarfe dieses Gerät nutzen würden. Deshalb wäre dieses Gerät in dieser kurzsichtigen Betrachtung unwirtschaftlich und der einzelne Bedarf würde wahrscheinlich auf einem anderen, längeren Weg geführt werden. Im schlechtesten Fall wiederholt sich dieser Vorgang bei einer großen Zahl jeweils für sich genommen kleiner Bedarfe, die zwar zusammen ein solches 40-Kanal-System wirtschaftlich auslasten könnten, aber nicht alleine, und deshalb durch das sequentielle Rerouting alle auf Umwegen geführt würden. Dieses Problem kann stark abgemildert werden, indem für das sequentielle Rerouting mit geglätteten Kostenfunktionen gerechnet wird. Erst bei der Bewertung am Ende der Sequenz wird dann wieder auf die realen, stufigen Gerätekosten zurückgegriffen.

Während des sequentiellen Reroutings arbeitet der Kürzeste-Wege-Algorithmus also auf einer neuen Metrik. Dabei handelt es sich um eine geglättete Funktion der Quotienten aus Kantenkosten und Kantenlasten. Für die Modellierung dieser Funktion steht ein Polynom der Form  $a \cdot x^n + b \cdot x^m + c$  zur Verfügung, wobei  $x$  die jeweilige Kantenlast ist. Die Koeffizienten und Exponenten werden normalerweise so gewählt, dass die Funktion in etwa eine Glättung der stufigen Gerätekosten darstellt. Die Kosten für die Technik in den Netzknoten wird ebenfalls durch ein solches Polynom beschrieben. Für das Routing werden die Knotenkosten jeweils auf die angrenzenden Kanten verteilt. Näherungsweise sind diese Kostenfunktionen oft Quadratwurzelfunktionen. Mit anderen Worten, eine Vervierfachung der Kapazität führt lediglich zu einer Verdopplung der Gerätekosten. Darin spiegeln sich die Economies of Scale wider. Eine Ausnahme von dieser Regel bildet häufig die jeweils aktuellste Technologiegeneration. Diese ist unter anderem auf Grund der hohen Investitionen für die Entwicklung und der noch geringen Stückzahlen oft überproportional teuer.

Die Definition der Kostenfunktionen als Polynom mit frei wählbaren Koeffizienten und Exponenten erlaubt es, diese an ein breites Spektrum verschiedener Geräte anzupassen. Unabhängig davon kann jedoch der Fall auftreten, dass eine Kante jedes Mal marginal teurer ist als eine andere. In Abhängigkeit

von der Kostenfunktion kann dieser Unterschied etwas größer oder etwas kleiner ausfallen, aber er bleibt prinzipiell immer bestehen. Selbst wenn dieser mathematische Unterschied für reale Anwendungen irrelevant klein sein sollte, wählt ein deterministischer Algorithmus folglich immer dieselbe dieser beiden Kanten aus. Um dies zu verhindern, wird ein zufälliger Multiplikator, z.B. mit einem Wertebereich zwischen 0.8 und 1.2, auf die Kantenkosten angewandt. Dieser hat keinen Einfluss auf den prinzipiellen Verlauf des Algorithmus, er erlaubt jedoch oft, dass der kleine letzte Schritt gefunden wird, der von einer nahezu optimalen Lösung zu einer optimalen führt. Ein negativer Effekt dieses Multiplikators konnte unter keinen Umständen festgestellt werden; entweder verbessert sein Einsatz das Endergebnis oder er hat keinen Einfluss.

Nach jedem Rerouting erfolgt eine Zuordnung der tatsächlich zur Verfügung stehenden Geräte und es werden die realen (im Gegensatz zu den geglätteten und damit virtuellen) Gesamtkosten errechnet. Wenn das Rerouting zu einer günstigeren Lösung führt, wird die neue Lösung gespeichert, anderenfalls wird sie verworfen. Danach wiederholt sich der Prozess, indem aufbauend auf der neuen Lösung eine neue Kandidatenliste berechnet wird usw., solange, bis keine Verbesserung mehr erzielt werden kann. Der Algorithmus hat nun ein lokales Optimum gefunden. Bezogen auf die Idee der Bündelung, um die Economies of Scale auszunutzen, bedeutet dies, dass eine weitere Konzentration auf noch weniger Kanten zu so vielen Umwegen führen würde, dass der Vorteil der Skaleneffekte überkompensiert werden würde.

Das Rerouting kann auf der Basis zweier verschiedener Nachbarschaften unterschiedlicher Granularität durchgeführt werden; entweder für einzelne Bedarfe oder für Bedarfsbündel, also alle Bedarfe, die vom gleichen Knoten ausgehen. Der Algorithmus kann während des Ablaufes von der gröberen Granularität (Bedarfsbündel) auf die feinere Granularität (Einzelbedarfe) wechseln. Standardvorgehen ist, auf Basis der Bündel ein lokales Optimum zu suchen und dann auf Einzelbedarfe umzuschalten, um die Lösung durch die verfeinerte Nachbarschaft weiter zu verbessern. Dieser zweistufige Ansatz ist für die meisten im Rahmen dieser Arbeit untersuchten Probleminstanzen am besten. Der Lösungsraum hat normalerweise viel zu viele lokale Optima, um alleine mit dem Einzelrerouting eine global gesehen gute Lösung zu finden. Die vorgeschaltete Bündelrerouting-Phase dient dazu, das Verfahren schnell in einen vielversprechenden Bereich des Lösungsraumes zu führen und dann lediglich diesen mit dem verfeinerten Verfahren genauer zu untersuchen.

Im Prinzip kann dieses Vorgehen als ein Variable Neighborhood Search (VNS) aufgefasst werden; die Nachbarschaft wird im Laufe des Verfahrens verändert. Allerdings unterscheidet es sich doch wesentlich von der in der Literatur gebräuchlichen Definition von VNS (vgl. z.B. Hansen und Mlade-

novic (2003)). Trotzdem sind einige Grundideen damit verwandt. In einigen Fällen ist es denkbar, dass eine noch bessere Lösung gefunden werden könnte, wenn die verfeinerte Suche mit Einzelbedarfen nicht nur von der besten Bündellösung, sondern von z.B. den besten fünf oder sogar allen zwischendurch gefundenen Bündellösungen durchgeführt werden würde. Dieser Ansatz wäre jedoch sehr rechenzeitintensiv. Wenn ohne Rücksicht auf die erhöhte Laufzeit eine möglichst gute Lösung für eine einzelne Problem Instanz gesucht wird, ist der Algorithmus jedoch leicht entsprechend zu modifizieren.

Nachdem ein lokales Optimum erreicht worden ist, wird die dazugehörige Lösung gespeichert und das Verfahren startet wieder von vorne. Damit fällt das Verfahren in die Klasse der sogenannten Multi-Start-Heuristiken. Die Anzahl dieser Neustarts ist ein externer, manuell festzulegender Parameter. Die sinnvolle Größenordnung hängt von der Größe und Struktur der jeweiligen Problem Instanz ab, vgl. Abschnitt 5.2. Die Startlösung und die erste Kandidatenliste vor der zufälligen Permutierung sind dabei immer identisch. Danach führen die Zufallskomponenten die Suche jedoch in unterschiedliche Teile des Lösungsraumes. Da jede äußere Iteration ein eigenständiger Prozess ist, eignet sich diese Heuristik prinzipiell sehr gut für den Einsatz auf Multiprozessorsystemen. Sie ist ohne nennenswerte Modifikationen direkt parallelisierbar. Weiterhin bewegt sie sich ausschließlich im zulässigen Lösungsraum, kann also jederzeit abgebrochen werden, und die aktuell beste Lösung ist immer eine gültige Lösung für das Ausgangsproblem.

Das Ergebnis eines vollständigen Laufs dieser Heuristik ist eine detaillierte Liste, welche Geräte an welchen Stellen im Netzwerk aufgebaut werden müssen und wie die einzelnen Bedarfe geroutet werden. Diese Gerätelisten enthalten für jeden Netzknoten und jede Glasfaserstrecke Angaben über Typ und Anzahl der benötigten Cross-Connects, Multiplexer, Verstärker, Transponder usw. Darüber hinaus werden natürlich die Gesamtkosten des Netzes, also der eigentliche Zielfunktionswert der Optimierung, ermittelt. Zur Visualisierung des Netzes kann das Ergebnis der Planung in einem zu Graphviz (<http://www.research.att.com/~north/graphviz>) kompatiblen Format ausgegeben werden. Abbildung 4.5 zeigt beispielhaft die Darstellung eines Glasfasergraphen mittels Graphviz.

Für die Implementierung eines Verfahrens mit einer starken Zufallskomponente ist die Verfügbarkeit eines guten Zufallszahlengenerators besonders wichtig. Der bei Visual Studio 6 integrierte Generator hat im Laufe der Arbeit immer wieder zu Ergebnissen geführt, die ernsthafte Zweifel an der Zufälligkeit der Zahlen aufgeworfen haben. Deshalb wird in der aktuellen Version der Software, mit der auch die meisten der in Kapitel 5 vorgestellten Rechenergebnisse erzielt wurden, der sogenannte „Mersenne twister“ von Matsumoto und Nishimura (1998) eingesetzt. Konkret wird die frei verfügbare

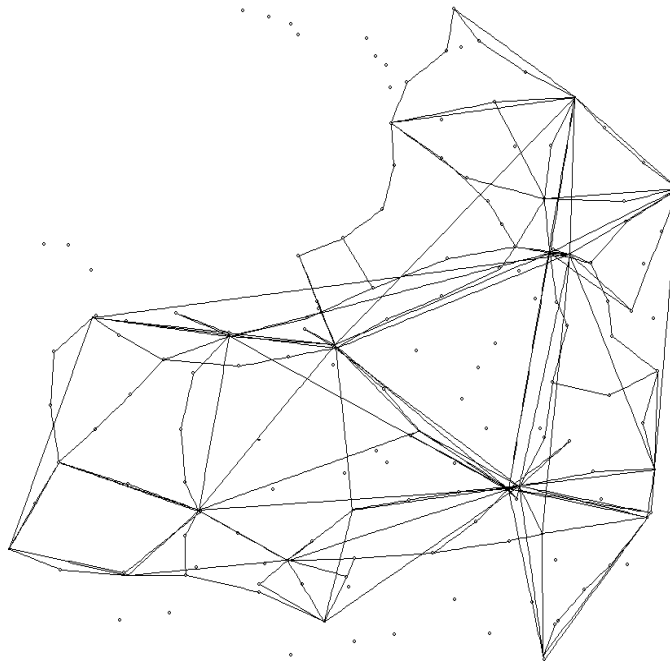


Abbildung 4.5: Glasfasergraph als Ergebnis einer Netzplanung.

Implementierung des „Mersenne twister“ von Wagner (2003) genutzt.

Vorläufer unseres GRASP-Algorithmus war ein stärker zufallsbasierter Algorithmus, der prinzipiell ebenfalls dem in Abbildung 4.4 gezeigten Ablauf folgt. Allerdings fand dort noch keine Bewertung der Bedarfe und dementsprechend auch noch keine Sortierung der Liste statt. Das Verhalten dieser Vorläuferversion lässt sich durch den aktuellen GRASP weitgehend simulieren, indem der Parameter, der das Verhältnis zwischen Zufälligkeit und Gierigkeit einstellt, maximal in Richtung Zufälligkeit verschoben wird. In Abschnitt 5 wird ein Vergleich vorgenommen, der aufzeigt, welche Verbesserungen durch die Einführung des GRASP erzielt werden konnten.

### 4.2.2 Pilotmethode

Viele lokale Suchstrategien basieren auf einer gierigen Auswahl der nächsten Lösung. Dabei orientiert sich die Bewertung der Lösungen oft nur am unmittelbar nächsten Schritt. Dieses kurzsichtige Verhalten kann leicht dazu führen, dass auf lange Sicht gesehen suboptimale Entscheidungen getroffen werden. Die grundlegende Idee der Pilotmethode ist es, die Auswahl des nächsten Schrittes nicht nur auf Basis einer Iteration zu treffen, sondern eine bestimmte Anzahl an Iterationen vorzuschauen. Es wird dann entspre-

chend der nächste Schritt gewählt, der am Ende des sichtbaren Horizonts zur besten Lösung führt, und nicht einfach der nächstbeste. Dieses grundsätzliche Vorgehen kann sowohl bei der Konstruktion einer Startlösung als auch bei einem nachfolgenden Verbesserungsverfahren eingesetzt werden. Hier wird es für die lokale Suche des Verbesserungsverfahrens genutzt. Eine umfassende Beschreibung der Pilotmethode und verwandter Ideen findet sich in Duin und Voß (1999) und Voß *et al.* (2005). In Höller *et al.* (2006) wird der hier vorgestellte GRASP+Pilot-Ansatz mit einer VNS+Pilot-Metaheuristik verglichen. Die Details des dort eingesetzten VNS sind nicht Bestandteil dieser Arbeit. Einige interessante Erkenntnisse der Vergleichsrechnungen aus Höller *et al.* (2006) werden jedoch in Abschnitt 5.2 diskutiert.

Die Pilotmethode wird als nachgelagertes Verfahren am Ende der GRASP-Suche durchgeführt. Ausgehend von der besten gefundenen Lösung wird damit untersucht, welche Auswirkungen das Umleiten von Transitverkehr hat. Dazu werden der Reihe nach alle Knoten des Netzes für Transitverkehr, also Verkehr für den dieser Knoten weder Quelle noch Senke ist, gesperrt. Formal kann dies durch einen binären Vektor  $t[1, \dots, N]$  dargestellt werden, dessen Elemente am Anfang alle den Wert 1 (= Transit zugelassen) enthalten. Ein Verbot des Transitverkehrs in Knoten  $n$  entspricht dann einer Invertierung des entsprechenden Elements  $t_n$  von 1 auf 0. Die vom Piloten zu durchsuchende Nachbarschaft ist die Menge aller Vektoren  $t'$ , die sich in genau einem Bit vom aktuellen Vektor  $t$  unterscheiden.

Der ursprüngliche GRASP-Algorithmus überprüft dann für jedes Element der so erzeugten Nachbarschaft mit einer begrenzten Suchtiefe, also z.B. ein Zehntel der normalen Zahl an Iterationen, ob die Modifikation zu einem guten Ergebnis führt. Die Veränderung, die bei dieser begrenzten Suche zum besten Zielfunktionswert führt, wird dauerhaft übernommen und beschreibt die neue Nachbarschaft, auf der nun ein vollständiger Durchlauf des GRASP-Algorithmus durchgeführt wird. Wenn der Zielfunktionswert des so gefundenen lokalen Optimums besser ist als der beste bisher bekannte, wird die Lösung gespeichert und der Pilot startet von neuem. Die Pilotsuche ist also eine wiederholte Anwendung des ursprünglichen Suchverfahrens mit verringerter Suchtiefe.

Spätere Läufe der Pilotsuche können dabei ursprüngliche Modifikationen wieder rückgängig machen. Wenn also z.B. in der ersten Pilotphase Knoten 3 gesperrt wurde und in der zweiten Phase zusätzlich Knoten 7, dann kann in der dritten Knoten 3 wieder freigegeben werden. Im Prinzip kann ein Tabu-Verfahren eingesetzt werden, um zu verhindern, dass bereits untersuchte Nachbarn erneut geprüft werden. Alternativ oder zusätzlich zum Sperren von Knoten können auch Kanten aus der Netzstruktur entfernt werden. Dies kann analog mit einem entsprechenden Vektor  $e[1, \dots, E]$  abgebildet werden.

Der Ablauf der Pilotsuche ist in Abbildung 4.6 dargestellt.

---

Führe den normalen GRASP-Algorithmus aus.

*Wiederhole bis keine Verbesserung mehr erzielt wird:*

- (1) *Wiederhole für alle Knoten [und Kanten]:*
    - a) Verbiete vorübergehend den Transitverkehr im jeweiligen Knoten [den Verkehr auf der jeweiligen Kante].
    - b) Führe den GRASP-Algorithmus mit einem Zehntel der normalen Zahl an Iterationen auf der so veränderten Nachbarschaft aus.
  - (2) Verändere die Nachbarschaft dauerhaft entsprechend der, die in Schritt (1) zur besten Lösung geführt hat.
  - (3) Führe den GRASP-Algorithmus mit der vollen Anzahl an Iterationen aus.
- 

Abbildung 4.6: Pilotsuche auf Basis des GRASP-Algorithmus.

Theoretisch könnte neben der oben angesprochenen Tabuliste eine weitere Beschränkung der untersuchten Lösungen vorgenommen werden. In der aktuellen Version wird auch das Blockieren von Knoten untersucht, die in der bekannten Lösung gar keinen Transitverkehr führen. Allerdings ist es möglich, dass diese Knoten in einer Zwischenlösung auf dem Weg zum lokalen Optimum Transitverkehr geführt haben. Deshalb kann es durchaus Sinn machen, diese Knoten in die Pilotsuche mit einzubeziehen.

Grundsätzlich skaliert die Laufzeit der Pilotsuche mit der Anzahl der Knoten im Netz. Deshalb ist dieser Ansatz vorteilhaft für eng vermaschte Netze mit verhältnismäßig wenig Knoten, aber nachteilig für weitmaschige Netze mit vielen Knoten. Mit der optionalen Einbeziehung von Kanten verhält es sich genau entgegengesetzt. In der Regel erfordert die Kantensuche deutlich mehr Rechenzeit, sie hat sich jedoch speziell für Probleminstanzen mit 1+1 Protection als wirkungsvoll erwiesen, vgl. Abschnitt 5.2.

Unabhängig davon, ob die Nutzung von Knoten oder von Kanten blockiert wird, ist zu beachten, dass diese Elemente nicht vollständig aus dem Netz entfernt werden dürfen. Anderenfalls könnte der Zusammenhang des Net-

zes verloren gehen und eine zulässige Lösung unmöglich werden. Aus diesem Grund werden die entsprechenden Elemente nur als extrem teuer gekennzeichnet, stehen aber im Notfall prinzipiell zur Verfügung, wenn es keinen anderen Weg im Netz gibt. Einziger Nachteil dieses Vorgehens ist, dass die Rechenzeit kürzer sein könnte, wenn das entsprechende Element ganz aus dem Netz entfernt worden wäre. Jedoch wäre dieser Zeitgewinn ohnehin nur marginal, da jeweils immer nur ein einziges Element wegfallen würde.

Die vom Pilotverfahren genutzte Funktion, Transitverkehr in Knoten zu blockieren, ist auch für operative Randbedingungen nutzbar. Beispielsweise können bestimmte Knoten im Netz großen Kunden gehören, die an ihrem Standort eigene Hardware aufbauen. Besonders bei mehreren unabhängigen Anbindungen an das eigentliche Netz des Telekommunikationsunternehmens müssen diese Kundenknoten direkt in die Planung einbezogen werden. Allerdings möchte der Kunde ggf. nicht, dass fremder Verkehr seine Geräte nutzt. Deshalb muss für das Planungsverfahren eine Möglichkeit bestehen, dies für einzelne Knoten gezielt zu verbieten. Genau dieser Mechanismus wird nun vom Pilotverfahren für die Modifikation der Nachbarschaft genutzt.

Eine explizite Veränderung der Nachbarschaft über Beschränkungen bei der Geräteplanung, beispielsweise keine WDM-Technologie auf einer Kante oder kein Technologiewechsel im Vergleich zur bisherigen Lösung, würde nicht mit der Grundphilosophie der Heuristik kompatibel sein. Der Routingalgorithmus arbeitet ausschließlich auf der Kostenmatrix und ist durch diese von der Hardware entkoppelt. Die Möglichkeiten zu Veränderungen der Nachbarschaft beschränken sich also lediglich auf Dinge, die durch eine Veränderung der Kostenmatrix abbildbar sind.

### 4.3 Zusammenspiel zwischen ganzzahligem Modell und Heuristiken

Abbildung 4.7 zeigt die Zusammenhänge zwischen den verschiedenen Softwariemodulen im Überblick. Ausgehend von den Datenbanken des Netzbetreibers und den Gerätespezifikationen werden die erforderlichen Eingabedaten für die Netzplanung generiert. Die Vorverarbeitung stellt dabei sicher, dass die Daten bereinigt und konsistent sind. Beispielsweise werden hier der (zweifache) Zusammenhang der Glasfaserstruktur geprüft oder doppelte Netzelemente entfernt. Weil es es sich dabei nur um standardisierte Datenbehandlung mit bekannten Verfahren handelt, wird hierauf nicht im Detail eingegangen. Der eigentliche Kern ist die Hauptverarbeitung, also die Netzplanung mit Hilfe der Heuristik oder AMPL/CPLEX. Die Ergebnisse werden abschlie-



ßend als Tabellen und Grafiken in verbreiteten Dateiformaten zur Verfügung gestellt. Details zu den unterstützenden Funktionen, wie z.B. der grafischen Ausgabe oder den Excel Tabellen, finden sich in Anhang A.

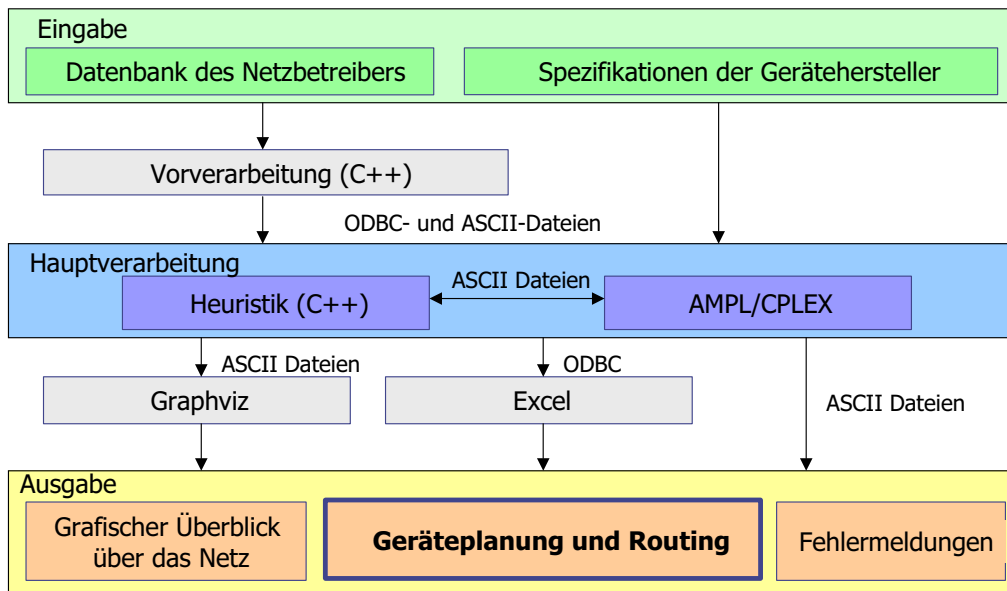


Abbildung 4.7: Zusammenspiel der Komponenten.

Es bieten sich verschiedene Möglichkeiten, durch ein direktes Zusammenwirken von Heuristik und AMPL-Modell weitere Verbesserungen zu erzielen. Das Beispiel der individuellen Hop-Limitierung wurde bereits in Abschnitt 4.1.2 kurz erwähnt. Hier dient die Heuristik als Hilfe, um die minimale Zahl an Hops für jeden Bedarf zu ermitteln. Weiterhin kann die Heuristik Schranken liefern, die an das MIP-Modell übergeben werden. Damit können Speicherplatz und Rechenzeit gespart werden, ähnlich wie bei einer vorherigen Berechnung einer Relaxation. Es ist ebenfalls denkbar, dass bestimmte Variablen des MIP-Modells auf feste Werte fixiert werden, die einer heuristischen Lösung entstammen. Durch die so reduzierte Problemgröße kann es möglich werden, auch größere Probleminstanzen mit dem exakten Verfahren zu lösen.

Andersherum ist es natürlich auch denkbar, dass MIP-Modelle zur Unterstützung der Heuristik eingesetzt werden. Beispielsweise kann eine (ggf. nicht beweisbar optimale) Lösung des MIP-Modells an Stelle des Kürzeste-Wege-Verfahrens als Startlösung für die Heuristik genutzt werden. Die Heuristik würde dann als lokales Verbesserungsverfahren auf Basis einer MIP-Lösung eingesetzt. Außerdem können Teilprobleme einer Heuristik, wie z.B.

eine lokale Suche in einem eng begrenzten Teil des Lösungsraumes, mit Hilfe exakter Verfahren durchgeführt werden. Hier würde also ein MIP-Modell als Unterproblem des übergeordneten heuristischen Ansatzes gelöst. Diese Möglichkeiten werden im Rahmen dieser Arbeit erst teilweise ausgenutzt und bieten noch Raum für weitere Forschung.

## 4.4 Ergänzende Literaturhinweise

Neben den in den vorhergehenden Abschnitten direkt im Text diskutierten Quellen gibt es eine Reihe weitere Publikationen mit verwandten Fragestellungen. Für die meisten der unten genannten Autoren(gruppen) gilt, dass sie diverse Publikationen auf eng angrenzenden Gebieten erstellt haben. Hier sind nur exemplarisch Beispiele genannt, die einen guten Überblick bieten oder besonders eng mit unserem Vorgehen verwandt sind.

Für den Bereich der ganzzahligen Optimierung gibt es nur wenige Quellen, die WDM, SDH und IP in einem Modell verbinden. Eine Ausnahme bilden hier Mobiot *et al.* (2003), die ein ganzzahliges Modell für ein vorberechnetes Multilayer Restorationverfahren entwickeln und auch (nach Einführung einer Lagrange-Relaxation) auf Testnetze bis zu 19 Knoten anwenden. Chigan *et al.* (2002b) und Chigan *et al.* (2002a) erwähnen explizit den Einsatz von gemischt ganzzahliger Optimierung bei einem koordinierten IP/WDM-Schutzverfahren und geben einige numerische Ergebnisse ihrer Berechnungen an, bleiben eine Beschreibung ihres Modells jedoch in beiden Publikationen schuldig. Sowohl Mobiot *et al.* (2003) als auch Chigan *et al.* (2002b) stellen dabei explizit den Bezug zu GMPLS her.

Eine dreischichtige Planung (Glasfaser, WDM und SDH) ist durchaus üblich. Hier sind z.B. Melian *et al.* (2005a) und Melian *et al.* (2004) zu nennen, bei denen ein Modellierungsansatz genutzt wird, der direkt mit dem von uns eingesetzten vergleichbar ist. Auch Melian (2003) greift verwandte Fragestellungen auf. Ähnliche Ansätze werden auch von Kennington *et al.* (2003) und Birkan *et al.* (2002) verwendet. Alle diese Quellen nutzen neben den ganzzahligen Optimierungsmodellen auch verschiedene heuristische Ansätze für größere Probleminstanzen. Auch Grover und Doucette (2001) verwenden ganzzahlige Modelle für ähnliche Problemstellungen. Fumagalli und Valcarenghi (2000) beschreiben gängige Schutzverfahren für IP over WDM und arbeiten mit einer Simulated Annealing-Heuristik zur Lösung beispielhafter Probleminstanzen mit 16 Knoten. Dabei wird das Maschennetz in verschiedene Ringe aufgeteilt. Allerdings ist die Modellierung der einzelnen Netzelemente bei den genannten Quellen in der Regel nicht so detailliert wie in dem in dieser Arbeit vorgestellten Modell.

Im deutschsprachigen Bereich wird am Zuse-Institut in Berlin viel mit ganzzahliger Optimierung für die Planung von SDH/WDM-Netzen gearbeitet. Entsprechend gibt es dort auch eine große Zahl an Veröffentlichungen zu diesem Thema, so beispielsweise in Wessály (2000). Einen sehr umfassenden Überblick dazu bieten Bley *et al.* (2003) und Orłowsky und Wessály (2004), ohne jedoch fertige Lösungsverfahren für komplexere Problemstellungen explizit darzulegen. Im Unterschied zu unserem eigenen Vorgehen wird maßgeblich ein selbst entwickelter Branch&Cut-Algorithmus an Stelle des kommerziellen CPLEX-Solvers eingesetzt. Unter Wessály und Pióro (2006) steht auch eine umfangreiche, aktuelle Literaturdatenbank zur Verfügung. Köhn und Gauger (2003) sind ebenfalls im Bereich SDH/WDM-Planung aktiv, allerdings mit Fokus auf einen hybriden Knoten mit elektrischer SDH- und optischer WDM-Schaltmatrix und bei sehr kleinen Netzgrößen (8 Knoten).

Ähnliche Modelle wie die hier genannten werden in vielen anderen Bereichen eingesetzt, die zugrunde liegenden mathematischen (Fluss-)Formulierungen sind relativ universell einsetzbar. Allerdings sind entsprechende Veröffentlichungen nur für die direkte Arbeit an den Modellen relevant, nicht aber für die Planung von Telekommunikationsnetzen im Allgemeinen. Für die Heuristiken gilt dies ebenso.

# Kapitel 5

## Rechenergebnisse und Fallstudien

Im Folgenden werden die Rechenergebnisse und Fallstudien diskutiert. Während die Rechenergebnisse aus Berechnungen mit realen oder generierten Probleminstanzen gewonnen werden, um die Leistungsfähigkeit der Verfahren unter unterschiedlichen Rahmenbedingungen zu testen, gehen die Fallstudien einen Schritt weiter. Hier werden die Verfahren auf reale Problemstellung aus dem Kontext strategischer Netzplanung angewandt.

Alle Berechnungen, unabhängig ob mit den Heuristiken oder mit CPLEX, werden auf einem handelsüblichen PC mit einem 2,4 GHz Pentium 4 Prozessor und 512 MB Hauptspeicher durchgeführt. Die eingesetzte CPLEX-Version ist 8.1. Die Heuristik hat einen von der Problemgröße abhängigen, aber während der Laufzeit konstanten Speicherbedarf. Bei CPLEX dagegen steigt der Speicherbedarf mit fortschreitender Rechenzeit durch den wachsenden Branch&Bound-Baum an. Somit könnte CPLEX theoretisch von mehr Speicher profitieren, die Heuristik jedoch nicht. In der Praxis ist es aber so, dass Probleme, die mit 512 MB Hauptspeicher nicht gelöst werden können, auch mit mehr Speicher nicht sinnvoll gelöst werden könnten, weil die Rechenzeiten dann extrem lang werden würden. Für die Heuristiken ist noch anzumerken, dass die Kostenfunktionen und andere einstellbare Parameter nur an die realen Randbedingungen der jeweiligen Gruppe von Probleminstanzen angepasst werden. Eine für die Praxis wenig praktikable, probleminstanzspezifische, manuelle Anpassung findet also nicht statt und ist auf Grund der Robustheit des Ansatzes auch nicht notwendig.

## 5.1 Möglichkeiten und Grenzen der MIP-Modelle

Die (gemischt) ganzzahligen Modelle aus den vorangehenden Abschnitten werden nun für verschiedene Arten von Probleminstanzen gelöst. Die Kostendaten spiegeln dabei reale Gegebenheiten (Gerätepreise) wider, sind jedoch aus Gründen der Übersichtlichkeit auf kleine ganzzahlige Werte normalisiert. Wenn die Lösung einer Probleminstanz nicht nachweislich optimal ist, dann bedeutet das, dass die Berechnung wegen Speichermangel vorzeitig beendet worden ist. Konkret wird die CPLEX-Berechnung immer dann abgebrochen, wenn die Baumgröße über 400 MB ansteigt. Die Einstellungen der Parameter für AMPL und CPLEX sind für alle Berechnungen identisch und entsprechen den herstellerseitigen Voreinstellungen. Die in den Tabellen benutzten Bezeichner  $|N|$ ,  $|E|$  und  $|D|$  entsprechen in ihrer Bedeutung den in Abschnitt 4.1.1 definierten Mengen aller Knoten, Kanten und Bedarfe.

**Vergleichsrechnungen für die verschiedenen Modellvarianten** Die erste Testreihe in Tabelle 5.1 zeigt die Streuung der Laufzeit für Probleminstanzen gleicher Größe (11 Knoten, 16 Kanten und 19 Bedarfe). Die Struktur des Netzes, also die Art, wie die 11 Knoten durch die 16 Kanten verbunden werden, und die Größe der Bedarfe, also die zu übertragende Datenmenge, die jedem der 19 Bedarfe zugrunde liegt, unterscheiden sich jeweils wie in der zweiten Spalte angegeben. Als Modell wird das in Abschnitt 4.1.1 eingeführte Basismodell mit einem Limit von acht Hops genutzt. Dieses Hoplimit hat aber keinen Einfluss auf das Routing, sondern dient nur der Beschleunigung der Berechnung. Alle gefundenen optimalen Lösungen haben ausschließlich Pfade mit weniger als 8 Hops, was bei lediglich 11 Knoten auch nicht verwunderlich ist. Die Basiseinheit der Bedarfe ist 2,5 Gbit/s. Dies gilt auch für alle folgenden Tabellen, soweit nicht explizit etwas anderes angegeben wird.

Alle Probleminstanzen können in weniger als zwei Sekunden optimal gelöst werden. Ein eindeutiger positiver oder negativer Einfluss der Netzstruktur, der Kostenstruktur oder der Bedarfsgröße ist bei fest vorgegebener Anzahl an Knoten, Kanten und Bedarfen nicht festzustellen. Gleichzeitig deutet sich aber bereits an, dass die zufällig aus der Problemstruktur resultierenden Laufzeitunterschiede durchaus deutlich ausfallen können. Auch wenn der absolute Unterschied bei solch kleinen Probleminstanzen noch nicht ins Gewicht fällt, so wird für Netz Nummer 5 annähernd fünfmal so viel Rechenzeit benötigt wie für Netz Nummer 4.

Tabelle 5.2 zeigt einen Vergleich der Laufzeiten des Basismodells für Probleminstanzen verschiedener Größen. Wenn innerhalb des verfügbaren Spei-

Nr.	Probleminstanz	Kosten	Laufzeit [s]
1	Test Netzwerk 1	1983	0,4
2	Nr. 1 mit Restkapazitäten	1924	0,3
3	Nr. 1 mit anderen Kantenkosten	1905	0,3
4	Nr. 2 mit anderen Kantenkosten	1848	0,2
5	Nr. 1 mit größeren Bedarfen	10442	1,9
6	Nr. 5 mit anderen Kantenkosten	10257	0,7
7	Test Netzwerk 2	1967	0,6
8	Nr. 7 mit größeren Bedarfen	10201	0,3
9	Test Netzwerk 3	1678	1,2
10	Nr. 9 mit größeren Bedarfen	8393	0,5
11	Test Netzwerk 4	1776	0,6
12	Nr. 11 mit größeren Bedarfen	8742	0,9

Tabelle 5.1: Vergleich verschiedener Probleminstanzen mit 11 Knoten, 16 Kanten und 19 Bedarfen.

chers von 400 MB keine beweisbar optimale Lösung gefunden wird, dann werden die beste gefundene Lösung und die Differenz zur besten bekannten Schranke angegeben.

Auch wenn die Laufzeiten in Tabelle 5.2 ebenfalls eine starke Streuung aufweisen, so lassen sich hier doch einige grundlegende Zusammenhänge aufzeigen, die auch in den anderen Berechnungen immer wiederzufinden sind. Natürlich skaliert die Rechenzeit (und auch der Speicherbedarf) mit der Größe der Probleminstanz. Dabei ist jedoch zu beachten, dass diese Größe kein einfach zu definierender Wert ist. Vielmehr kommt es auf die Kombination aus der Anzahl an Knoten, Kanten und Bedarfen an.

Bei einem Netz mit nur sehr wenigen möglichen Wegen, wie zum Beispiel Nummer 1, ist eine große Zahl an Bedarfen nicht kritisch. Bei einem höheren Vermaschungsgrad und somit mehr Wegen kann eine große Zahl an Bedarfen die Rechenzeit stark negativ beeinflussen, wie der Vergleich zwischen den Netzen Nummer 2 und 3 deutlich zeigt. Ein ähnlicher Effekt ist beim Vergleich der Netze Nummer 4, 5 und 6 zu sehen. Obwohl Netz Nummer 5 in jeder Dimension größer ist als Netz Nummer 4, so ist die Rechenzeit doch länger. Durch die deutlich größere Zahl an Knoten, aber eine nur etwas größere Zahl an Kanten ist das Netz Nummer 5 wesentlich weitmaschiger,

Nr.	$ N $	$ E $	$ D $	Kosten	Laufzeit [s]	Max. Differenz
1	10	10	45	5019	1,7	-
2	11	16	19	1678	1,6	-
3	11	16	54	4445	10,7	-
4	11	22	27	2758	38,1	-
5	17	26	28	2995	15,6	-
6	20	22	46	7959	4,6	-
7	20	33	84	9426	9362	1,6%

Tabelle 5.2: Vergleich der Laufzeiten für verschiedene Netzgrößen.

bietet also weniger Alternativen für das Routing. Folglich ist es tendenziell einfacher zu berechnen. Noch deutlicher wird dieser Effekt beim Vergleich mit Netz Nummer 6. Eine Struktur aus 20 Knoten und 22 Kanten besteht effektiv nur aus sehr wenigen, großen, gekoppelten Ringen. Folglich gibt es nur wenige mögliche Routen für die einzelnen Bedarfe, so dass diese vermeintlich große Probleminstanz sehr schnell lösbar ist.

Die gleiche Anzahl an Knoten in Netz Nummer 7, diesmal jedoch kombiniert mit einer dichteren Maschenstruktur und noch mehr Bedarfen, ist dann bereits nicht mehr beweisbar optimal lösbar. Die für Probleminstanz Nummer 7 gefundene Lösung wird bereits nach weniger als einer Minute und mit lediglich 20 MB Speicherbedarf gefunden, kann jedoch auch mit 400 MB Speicher nicht abschließend als optimal verifiziert werden.

Tabelle 5.3 zeigt die Ergebnisse für Probleminstanzen mit 1+1 geschützten Bedarfen, vgl. Abschnitt 1.2.4. Von den bisher eingeführten kleinen Probleminstanzen weisen lediglich zwei eine zweifach zusammenhängende Netztopologie auf. Deshalb beschränkt sich der Vergleich in Tabelle 5.3 lediglich auf diese beiden Probleminstanzen. Weitere 1+1 geschützte Netze finden sich in Abschnitt 5.2 beim Vergleich mit den Heuristiken. Das allgemeine Bild bei Aufnahme von 1+1 Protection in das Modell wird jedoch von diesen beiden Beispielen ideal widergegeben.

Die Laufzeit liegt in etwa im Bereich der identischen Probleminstanz mit ungeschützten Bedarfen, teilweise darunter, teilweise darüber. 1+1 Protection ist also mit Hilfe der vorgestellten Modellierung ohne negative Folgen nutzbar. Dies ist auch von der Art der Modellformulierung her bereits zu vermuten. Die Anzahl der Entscheidungsvariablen bleibt identisch, es wird lediglich für jeden geschützten Bedarf eine neue Nebenbedingung eingeführt. Bezogen auf die Gesamtgröße des Modells ist diese Veränderung vernachlässigbar. Die

Nr.	N	E	D	1+1 geschützt		ungeschützt	
				Kosten	Laufzeit [s]	Kosten	Laufzeit [s]
1	11	16	19	4082	2,6	1678	1,2
2	11	22	27	5724	21,7	2758	38,1

Tabelle 5.3: Vergleich der Laufzeiten zwischen 1+1 geschützten und ungeschützten Bedarfen.

Gesamtkosten des geschützten Netzes liegen jeweils knapp doppelt so hoch wie die des ungeschützten Netzes. Dies ist für so kleine Netze ebenfalls repräsentativ.

Die Auswirkungen der Modellerweiterung um mehrere Typen an Cross-Connects (vgl. Abschnitt 4.1.2) sind in Tabelle 5.4 dargestellt. Jede zusätzliche Schicht in den Knoten erhöht den Rechenaufwand beträchtlich. Die Angaben in Klammern bei den Probleminstanzen Nummer 3 und 6 geben die Laufzeit für den Fall an, dass für alle Knoten die zusätzlichen Nebenbedingungen  $y_n^S, y_n^M \leq 1$  gelten. Anschaulich bedeutet dies, dass zwei kleine Cross-Connects immer teurer sind als ein Cross-Connect der nächst größeren Stufe (der mindestens so viel Kapazität hat wie die beiden kleinen zusammen). Damit wird effektiv von allen Typen bis auf den jeweils größten maximal ein Gerät in jedem Knoten installiert. Bei den beiden Probleminstanzen ist dies implizit durch die Gerätespezifikationen sichergestellt, so dass der Zielfunktionswert folglich unverändert bleibt. Die Ersparnis an Rechenzeit erfolgt also lediglich durch die explizite Formulierung dieser ohnehin implizit enthaltenen Nebenbedingungen.

Die mit zunehmender Zahl an XC-Schichten sinkenden Gesamtkosten der beiden Netze spiegeln die Skaleneffekte wider. Stehen nur kleine Geräte zur Verfügung, so sind diese bezogen auf ihre Kapazität relativ teuer. Können dagegen ein oder mehrere kleine Geräte an einem Standort durch ein größeres ersetzt werden, sinken die Gesamtkosten des Netzes.

Die Einführung von VC4-Bedarfen (vgl. Nebenbedingung 4.1.2) in das Modell führt alleine noch nicht zu einer Veränderung des Laufzeitverhaltens oder des Speicherbedarfs. Erst wenn tatsächlich in einer Probleminstanz Bedarfe auftreten, die kein Vielfaches von 16 sind, wird die Lösung des Problems erschwert. Je mehr es von diesen Bedarfen gibt, desto schwieriger ist die Probleminstanz zu lösen. Tabelle 5.5 zeigt die Laufzeiten für die bereits aus Tabelle 5.2 bekannten Probleminstanzen, diesmal jedoch durchgängig mit VC4-Bedarfen. Die benötigten Laufzeiten sind offensichtlich erheblich länger, im Extremfall von Netz Nummer 3 werden aus gut zehn Sekunden über ein-



Nr.	$ N $	$ E $	$ D $	XC-Typen	Kosten	Laufzeit [s]
1	11	16	19	1	1678	1,5
2	11	16	19	2	1618	5
3	11	16	19	3	1588	16 (5)
4	11	22	27	1	2758	40
5	11	22	27	2	2658	76
6	11	22	27	3	2579	121 (85)

Tabelle 5.4: Vergleich der Laufzeiten bei der Einführung verschiedener XC-Typen.

einhalb Stunden. Wenn innerhalb des verfügbaren Speichers von 400 MB keine beweisbar optimale Lösung gefunden wird, dann werden wieder die beste gefundene Lösung und die Differenz zur besten bekannten Schranke angegeben.

Nr.	$ N $	$ E $	$ D $	Kosten	Laufzeit [s]	Max. Differenz
1	10	10	45	3143	6,4	-
2	11	16	19	1231	11	-
3	11	16	54	2687	6558	-
4	11	22	27	1840	1643	-
5	17	26	28	2119	556	-
6	20	22	46	4951	347	-
7	20	33	84	1207	54900	61,6%

Tabelle 5.5: Vergleich verschiedener Probleminstanzen mit VC4-Grooming.

Das umfassende Modell mit zwei SDH-Geschwindigkeiten (2,5 Gbit/s und 10 Gbit/s) und zwei verschiedenen XC-Typen ist ohne problemspezifische Anpassungen nur noch für sehr kleine Probleminstanzen optimal lösbar. Die entsprechenden Ergebnisse sind in Tabelle 5.6 zusammengefasst.

Das Modell für das Restoration-Schutzverfahren ist nur als Test für die prinzipielle Machbarkeit gedacht, nicht jedoch als Basis für die Berechnung praktischer Probleminstanzen. Folglich werden in Tabelle 5.7 nur zwei kleine Netze aufgeführt, an denen das korrekte Routing und die Einhaltung der Reservekapazitäten überprüft werden kann. Weiterhin ist natürlich auch Restoration genau wie 1+1 Protection auf einen zweifachen Zusammenhang der

Nr.	$ N $	$ E $	$ D $	Kosten	Laufzeit [s]
1	5	8	10	289	0,28
2	10	10	45	1298	0,7
3	11	16	19	768	56,8

Tabelle 5.6: Laufzeiten für zwei SDH-Schichten und zwei XC-Typen.

Netzstruktur angewiesen. Der Vergleich mit Tabelle 5.3 zeigt wie erwartet, dass die Laufzeiten deutlich über denen für ungeschützte bzw. 1+1 geschützte Netze liegen. Die Gesamtkosten sind in etwa mittig zwischen den Kosten für ein ungeschütztes und ein 1+1 geschütztes Netz.

Nr.	$ N $	$ E $	$ D $	Kosten	Laufzeit [s]
1	11	16	19	2774	4,2
2	11	22	27	4102	190

Tabelle 5.7: Laufzeiten für Netze mit Restoration als Schutz.

**Betrachtung technischer Veränderungen der Modellierung** Bei Tests mit linearen Relaxationen hat sich wie erwartet gezeigt, dass die Ganzzahligkeitsbedingungen entscheidend zur Schwere des Problems beitragen. Die Auswirkungen der einzelnen Variablen sind jedoch höchst unterschiedlich. Die Relaxierung der Gerätevariablen  $w_e$  und  $y_k$  vereinfacht die Problemstellung deutlich (bis zu Faktor 10 geringere Rechenzeit), führt jedoch zu Zielfunktionswerten, die etwa 10-20% unter der besten zulässigen Lösung liegen. Hierbei ist der Einfluss der WDM-Multiplexer tendenziell deutlich größer, da diese aufgrund ihrer hohen Kosten meistens wesentlich mehr zu den Gesamtkosten beitragen als die XC-Basisgeräte.

Die Relaxierung der Ganzzahligkeit der Routingvariablen  $z_{ij}$  hat dagegen keinen eindeutig positiven Einfluss auf die Rechenzeit. Die Berechnung der 12 verschiedenen 11-Knoten-Netze aus Tabelle 5.1 mit relaxierten  $z_{ij}$  zeigt, dass die Laufzeit nur in vier Fällen kürzer ist, während sie bei den anderen acht Probleminstanzen zum Teil bis zu zehnmal länger ist. Eine genaue Betrachtung der jeweiligen Lösungen offenbart, dass nur sehr wenige Bedarfe tatsächlich auf zwei oder gar drei Pfade aufgespaltet werden. Drei der Probleminstanzen hatten auch mit einem relaxierten  $z_{ij}$  einen unveränderten Zielfunktionswert. Hier bleiben die Variablen also implizit ohnehin binär. Für die  $z_{ij}$  Variablen erscheint eine Relaxation somit nicht zielführend. Das

Ergebnis ist zum Teil im Sinne der eigentlichen Problemstellung nicht mehr zulässig, im Gegenzug wird aber kein (Laufzeit-) Vorteil erzielt. Die Restkapazitäten  $g_e$  und  $h_e$  haben keinen Einfluss auf die Laufzeit. Hierbei handelt es sich lediglich um Konstanten, die ohne Nebenwirkungen nach Belieben aufgenommen oder weggelassen werden können.

Abbildung 5.1 zeigt den Einfluss des Hoplimits  $H$  beispielhaft für eine Probleminstanz. Für geringe Werte ( $H = 1 \dots 3$ ) existiert in diesem Fall keine zulässige Lösung. Für  $H = 4$  wird eine Lösung gefunden, allerdings ist diese schlechter als ohne Hoplimit. Für ein Hoplimit von  $H \geq 5$  wird immer eine optimale Lösung des unbeschränkten Problems gefunden; hier liegen die Unterschiede nur in der benötigten Rechenzeit. Auf den Zusammenhang zwischen Hoplimit und Rechenzeit soll jetzt genauer eingegangen werden.

Während das genaue Verhalten der Laufzeit in Abhängigkeit vom Hoplimit natürlich von Probleminstanz zu Probleminstanz variiert, so sind in der Abbildung 5.1 doch einige Effekte erkennbar, die sich als typisch herausgestellt haben. Ohne jegliche Beschränkung und nah an der unteren Grenze des Hoplimits sind die Laufzeiten gewöhnlich relativ lang. Dazwischen existiert ein Tal mit merklich reduzierten Laufzeiten. Eine genaue Vorhersage, für welchen Wert sich die geringste Laufzeit ergibt, ist nicht möglich. Werte etwa zwischen acht und zehn ergeben jedoch fast durchgehend sehr gute Resultate. Dieser Bereich ist auch groß genug, um in Bezug auf Zulässigkeit und Optimalität der Lösung bei den meisten realistischen Probleminstanzen auf der sicheren Seite zu sein. Sowohl aus ökonomischen als auch aus technischen Gründen wird ein Pfad mit mehr als zehn Hops bei der hier gewählten Planungsgranularität kaum vorkommen. Somit hat sich diese Einstellung als Standardwert für fast alle Berechnungen bewährt, um die Rechenzeit zu reduzieren.

Darüber hinaus ist ein variables Hoplimit denkbar. Hier wird das Limit individuell für jeden Bedarf festgelegt, und zwar auf die Anzahl der Hops des kürzesten möglichen Weges zuzüglich einer kleinen Zahl, beispielsweise drei Hops. Allerdings hat dies keinen Vorteil im Vergleich mit dem pauschalen Hoplimit erbracht. Somit wird diese Idee hier nicht weiter verfolgt.

Die Auswirkungen anderer, ähnlicher Nebenbedingungen, z.B. eine Begrenzung für die Anzahl bestimmter XC-Typen in bestimmten Knoten (Stellplatzbegrenzung), sind nicht so einfach zu beurteilen wie die des Hoplimits. Teilweise wird die Berechnung dadurch beschleunigt, weil ein Teil des Lösungsraumes abgeschnitten wird und folglich auch nicht durchsucht werden muss. Teilweise erhöht sich die Rechenzeit aber auch stark, da die zusätzlichen Nebenbedingungen die Lösung erschweren.

Die Variablen  $f_e$  und  $v_e$  sind im Basismodell technologiebedingt grundsätzlich ganzzahlig. Ein Kanal kann immer nur ganz oder gar nicht

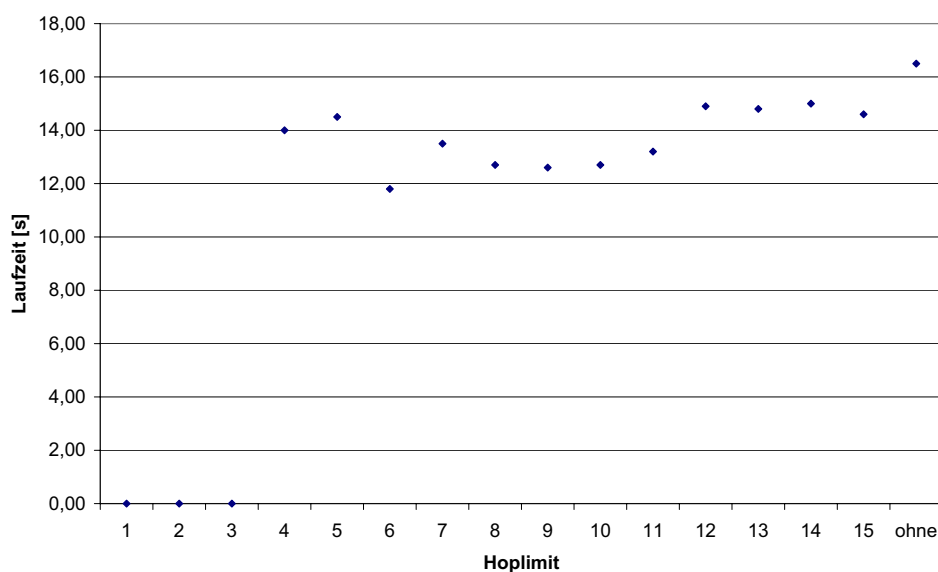


Abbildung 5.1: Beispiel für die Auswirkung des Hoplimits auf die Laufzeit.

installiert werden. Jedoch sind diese Variablen im Modell bereits implizit ganzzahlig, da die Bedarfe nur ganzzahlig sein dürfen. Somit können die Variablen  $f_e$  und  $v_e$  ohne weiteres auch als Fließkomma deklariert werden. Ursprünglich stand die Idee dahinter, dass CPLEX dann theoretisch weniger Branch&Bound-Operationen und dafür mehr Simplex-Operationen ausführen sollte, um die Rechenzeit zu verkürzen. Für einige Probleminstanzen konnte dies auch nachvollzogen werden. Für viele andere zeigte sich dabei jedoch ein daraus resultierender, bemerkenswerter Nebeneffekt. Durch die Deklaration von  $f_e$  und  $v_e$  als Fließkomma ist CPLEX oft in der Lage, intern wesentlich mehr Schnitte auszuführen (laut Logfile hauptsächlich „Gomory fractional cuts“ und „Mixed integer rounding cuts“) und damit deutlich schneller gute Lösungen zu finden. Die beobachteten Laufzeitveränderungen sind in Tabelle 5.8 dargestellt. Von Netz Nummer 1 abgesehen ist immer ein Vorteil zu beobachten.

Der Vollständigkeit halber soll noch erwähnt werden, dass es einen theoretischen Spezialfall gibt, in dem  $f_e$  und  $v_e$  in der Lösung doch die Ganzzahligkeit verlieren. Dies kann genau dann eintreten, wenn ein WDM-Kanal genau so viel kostet wie ein SDH-Kanal und WDM-Multiplexer kostenlos zur Verfügung stehen. Allerdings ist dieser Fall für die Praxis nicht relevant und kann zudem einfach geheilt werden. Von den entsprechenden  $f_e$ - und  $v_e$ -Paaren zu jeder Kante muss in der Lösung jeweils einer der beiden Werte auf

den nächst größeren aufgerundet und der andere auf den nächst niedrigeren abgerundet werden. Es spielt dabei wegen der identischen Kosten keine Rolle welcher Wert auf- und welcher abgerundet wird.

Netz	$f, v$ Integer	$f, v$ Fließkomma
Nr.	Laufzeit [s]	Laufzeit [s]
1	2,8	4,3
2	11,5	3,7
3	12,6	9,4
4	43,7	13,0
5	210,4	129,8

Tabelle 5.8: Einfluss der Variablendeklaration von  $f$  und  $v$ .

**Ergebnisse für diversifizierte Lösungen** Aufgrund der in Abschnitt 4.1.3 diskutierten prinzipiellen Probleme mit den Flussformulierungen bei diversifizierten Lösungen ist die entsprechende Modellierung nur als Prototyp zu verstehen. Folglich wird auch nur die prinzipielle Funktionsweise anhand eines Beispiels diskutiert. Dabei handelt es sich um eines der bereits eingeführten künstlich generierten Netze mit 11 Knoten, 16 Bedarfen und 19 Kanten. Eine Untersuchung des Verhaltens bei verschiedenen Probleminstanzen macht erst Sinn, wenn die Modellierung entsprechend der in Abschnitt 4.1.3 dargelegten Ansatzpunkte weiterentwickelt wird. Die Zielfunktionswerte und Laufzeiten für ein kleines Netz sind in Tabelle 5.9 gegenübergestellt. Dabei symbolisieren  $a$  und  $b$  die beiden unterschiedlichen Lösungen. Für den Gewichtungsfaktor  $\omega$  wird der Wert 10 gewählt. Der unter „Abstand“ angegebene Wert ist die gewichtete Summe aller Unterschiede in den Routingvariablen. Ein Abstand von 770 bedeutet bei  $\omega = 10$  also, dass sich die beiden Lösungen in 77 Routingvariablen unterscheiden.

	Zielfunktionswert $a$	Zielfunktionswert $b$	Abstand	Laufzeit [s]
sequentiell	1983	2247	770	1.6
parallel	2098	2123	860	10568

Tabelle 5.9: Diversifizierung der Lösungen für eine kleine Probleminstanz.

Der Vergleich macht sofort den großen Nachteil des parallelen Verfahrens deutlich; die Laufzeit ist bereits für sehr kleine Probleminstanzen extrem hoch. Während das ursprüngliche Problem noch in weniger als einer Sekunde gelöst werden kann und der Rechenaufwand für das sequentielle Verfahren in etwa doppelt so hoch ist, benötigt das parallele Verfahren fast drei Stunden. Wie bereits für das Basismodell beobachtet, steigt die Rechenzeit mit zunehmender Problemgröße extrem an. Während beim sequentiellen Verfahren lediglich zwei Probleme kleiner Größe unabhängig voneinander gelöst werden müssen, wird beim parallelen Verfahren ein einziges, aber deutlich größeres, Problem gelöst; die Anzahl an Entscheidungsvariablen ist fast doppelt so hoch. Dieses sehr unvorteilhafte Laufzeitverhalten führt dazu, dass das parallele Verfahren zumindest in dieser Form nicht für reale Planungsprobleme geeignet ist.

Tendenziell sollte das parallele Verfahren allerdings den Vorteil bieten, dass die Lösungen ausgewogener sind, also der Unterschied zwischen den beiden Zielfunktionswerten kleiner ist als beim sequentiellen Ansatz. Im Idealfall gibt es sogar eine Mehrzahl an Lösungen mit gleichem Zielfunktionswert, die eine stark unterschiedliche Lösungsstruktur aufweisen. Bildlich gesprochen kann man sich dem Optimum des Lösungsraumes aus sehr verschiedenen Richtungen nähern. Das sequentielle Verfahren erhält jedoch normalerweise als erste Lösung immer eine optimale Lösung. Hierzu wird dann in gewisser Distanz eine zweite Lösung gesucht, die, vom Spezialfall mehrerer verschiedener optimaler Lösungen abgesehen, zwingend schlechter ist. In der Tat wurde für keines der in dieser Arbeit insgesamt untersuchten Netze ein Indiz für die Existenz mehrerer optimaler Lösungen gefunden. Natürlich handelt es sich hierbei nur um eine aus dem grundsätzlichen Vorgehen abgeleitete Aussage. Um diese fundiert belegen zu können, wären zusätzliche Tests notwendig.

**Abschließende Bemerkungen** Zu beachten ist, dass in diesem Abschnitt nur die allgemeinen Formulierungen ohne problemspezifische Anpassungen verglichen werden. Damit sollen die Grenzen für das MIP-Modell aufgezeigt werden, ohne dass speziell an spezifische Probleminstanzen angepasste Formulierungen des Modells einen zu optimistischen Eindruck von der allgemeinen Leistungsfähigkeit hinterlassen. Oft kann jedoch durch kleine problemspezifische Anpassungen der Modellformulierung ein erheblicher Vorteil erzielt werden. Im Extremfall kann beispielsweise eine obere Schranke für bestimmte ganzzahlige Variablen oder eine leichte Modifikation einer Nebenbedingung aus einem unlösbaren Problem ein optimal lösbares machen. Es ist immer ratsam über die Möglichkeit solcher Verstärkungen der Modellformulierung nachzudenken, bevor eine Probleminstanz als zu komplex für den

MIP-Ansatz angesehen wird. Solche problemspezifischen Anpassungen werden dann beim Vergleich mit den Heuristiken in Abschnitt 5.2 berücksichtigt, so dass dort teilweise entsprechend größere Probleminstanzen gelöst werden können.

Die angegebenen Laufzeiten der Berechnungen verdienen eine genauere Betrachtung. Obwohl die Zeiten oft recht lang sind, wird die beste Lösung teilweise schon deutlich eher gefunden. Die verbleibende Zeit wird dann häufig lediglich benötigt, um zu beweisen, dass die bisher gefundene Lösung optimal ist. Tabelle 5.10 zeigt für einige Probleminstanzen, wie viele Branch&Bound-Knoten insgesamt besucht werden und nach wie vielen die beste Lösung bereits gefunden ist. Eine direkte Angabe der bis dahin verstrichenen Zeit ist leider nicht verfügbar. Die Rechenzeit korreliert jedoch in etwa linear mit der Anzahl der besuchten Branch&Bound-Knoten.

In diesem Zusammenhang kann der CPLEX-Parameter „MIPGAP“ in der Praxis hilfreich sein. Er führt zu einem Abbruch der Berechnung, wenn die beste bekannte Lösung lediglich um einen einstellbaren Prozentsatz schlechter als die beste bekannte Schranke ist. Beispielsweise legt der Befehl „MIPGAP = 0.03“ diese Differenz auf 3% fest. Damit kann der beschriebene Effekt, dass viel Rechenzeit nur auf das Beweisen der Optimalität verwendet wird, abgemildert werden. Um bei den in dieser Arbeit durchgeführten Berechnungen, die auch als Referenz für die Heuristik dienen sollen, absolut sicher zu gehen, dass keine optimale Lösung übersehen wird, wird von dieser Möglichkeit jedoch kein Gebrauch gemacht.

Netz Nr.	Laufzeit [s]	besuchte Knoten Gesamtanzahl	besuchte Knoten bis zur besten Lösung
1	14	1.759	539
2	176	29.241	13.564
3	6148	382.862	369.300
4	11194	704.170	26.830
5	47200	1.339.073	628.321

Tabelle 5.10: Anzahl der besuchten Branch&Bound-Knoten.

CPLEX bietet eine große Zahl an Parametern, mit denen die interne Suchstrategie verändert werden kann, wie z.B. unterschiedliche Branching-Strategien. Allerdings hat keiner der diversen Parameter bei einer Veränderung gegenüber den Standardeinstellungen eine konsistente Verbesserung der Ergebnisse erbracht. Ebenso ist der spezielle „network optimizer“

von CPLEX für diese Problemstellung nicht hilfreich. Obwohl von der Idee her wahrscheinlich für ähnliche Modelle gedacht, führt er durchgehend zu schlechteren Ergebnissen. Folglich werden alle Einstellungsmöglichkeiten von CPLEX bei allen hier aufgeführten Berechnungen auf den Standardwerten belassen.

Ein wichtiger Aspekt im Zusammenhang mit den Ergebnissen für die MIP-Modelle ist der große Einfluss des Solvers. Es ist erstaunlich, welche großen Fortschritte auf diesem Gebiet in nur vier Jahren erzielt werden konnten. Die ersten Tests dieser MIP-Formulierungen wurden ursprünglich noch mit CPLEX-Version 6 (erschien 1998) durchgeführt. Mit dem Erscheinen von Version 8 im Jahr 2002 wurden dann alle Berechnungen mit dieser Version durchgeführt und alte Berechnungen ggf. wiederholt. Tabelle 5.11 zeigt exemplarisch für einige Probleminstanzen die Unterschiede in der Laufzeit, die durch den Versionswechsel zu beobachten sind. Tabelle 5.12 stellt die Zielfunktionswerte für Probleminstanzen gegenüber, bei denen Version 8 im Rahmen der Speichergrenze von 400 MB für den Branch&Bound-Baum eine optimale Lösung findet, Version 6 jedoch nicht.

$ N $	$ E $	$ D $	Laufzeit [s] V6	Laufzeit [s] V8
11	16	19	2.8	0.46
11	16	54	130	11
11	22	27	610	76

Tabelle 5.11: Vergleich der Laufzeiten zwischen CPLEX 6 und CPLEX 8.

$ N $	$ E $	$ D $	Zielfunktionswert V6	Zielfunktionswert V8
11	22	27	1365	1231
20	22	46	6978	4951
20	33	84	2584	2579

Tabelle 5.12: Vergleich der Lösungsgüte zwischen CPLEX 6 und CPLEX 8.

Die aktuelle Version 10 von CPLEX ist möglicherweise noch leistungsfähiger, sie stand für diese Berechnungen jedoch noch nicht zur Verfügung. Zusammen mit der stetig steigenden Rechenleistung der Computer ermöglicht diese Entwicklung das Lösen immer größerer Probleminstanzen. Insofern sind einige Probleme, die vor wenigen Jahren nur heuristisch gelöst werden konnten, mittlerweile problemlos mit einem exakten Verfahren



optimal lösbar. Andererseits sind viele reale Problemstellungen hinsichtlich der Lösbarkeit momentan noch weit außer Reichweite und auch die zu planenden Netze selber werden immer größer und komplexer. Insofern erscheint es trotz aller Fortschritte doch zweifelhaft, dass in absehbarer Zeit auf den Einsatz von Heuristiken verzichtet werden kann.

Die in diesem Abschnitt vorgestellten Ergebnisse für das ganzzahlige Modell zeigen jedoch, dass detaillierte Modelle ohne künstliche Einschränkung des Lösungsraumes für in der Praxis relevante Netzgrößen lösbar sind. Gerade die Backbonenetze bestehen oft nur aus weniger als 20 Knoten und sind somit prädestiniert für diesen Lösungsansatz. Das weit verbreitete Schutzverfahren 1+1 Protection hat zudem keinen eindeutig negativen Einfluss auf die benötigte Rechenzeit. Allerdings wird auch deutlich, dass Netze mit einer wesentlich größeren Zahl an Knoten nicht mehr exakt gelöst werden können, insbesondere wenn eine sehr detaillierte Modellierung der Gerätetypen gefordert ist. Je detaillierter das Netz modelliert wird, desto kleiner ist die optimal lösbare Netzgröße. Wenn die exakten Verfahren an ihre Grenze stoßen, kann auf Heuristiken zurückgegriffen werden, deren Leistungsfähigkeit im folgenden Abschnitt untersucht wird.

## 5.2 Leistungsfähigkeit der Heuristiken

Heuristiken bieten keine Garantie für gute oder gar optimale Ergebnisse. Trotzdem sind sie für große Probleminstanzen aufgrund der Komplexität der Problemstellung oft der sinnvollste Lösungsweg. In diesem Abschnitt werden die Ergebnisse der Heuristiken dargestellt und, wo möglich, mit den Ergebnissen oder Schranken des MIP-Ansatzes verglichen. Dabei werden drei verschiedene Arten von Probleminstanzen betrachtet: kleine zufällig generierte Instanzen, wie sie bereits aus Abschnitt 5.1 bekannt sind, Instanzen mittlerer Größe aus Höller *et al.* (2006), die an reale Netztopologien angelehnt sind, und einige sehr große Probleminstanzen, die direkt Praxisproblemen entnommen sind.

Auf Grund der Zufallskomponenten der Heuristik kann streng genommen im Gegensatz zu den CPLEX-Berechnungen kein einzelner Wert für die Laufzeit bzw. den Zielfunktionswert angegeben werden. Eine solche Heuristik bietet per Definition keine Garantie, dass bei jedem Rechenlauf das gleiche Ergebnis in der gleichen Zeit erzielt wird. Bei einer ausreichend hohen Einstellung des Iterationszählers ist die Konvergenz jedoch recht gut. Die in den Tabellen angegebenen Werte sind das jeweils beste Ergebnis aus zehn unabhängigen Rechenläufen.

Der Iterationszähler wird für die meisten Probleminstanzen auf dem Standardwert von 100 belassen. Dies erklärt die vergleichsweise hohen (wenn auch absolut sehr geringen) Rechenzeiten für die sehr kleinen Probleminstanzen; dort hätte ein geringerer Wert ausgereicht. Lediglich bei einigen der größeren Instanzen wird der Zähler auf 1000 oder sogar 10000 eingestellt. Die Strategie „kürzester Weg“ dient fast immer als Ausgangspunkt der GRASP-Suche. Im Gegensatz zum Abschnitt 5.1, wo die allgemeinen Modelle genutzt werden, wird diesmal das AMPL-Modell für viele der Probleminstanzen individuell angepasst, um noch bessere Ergebnisse zu erhalten. Das universelle Modell wäre nicht mehr in der Lage, für die bis zu 50 Knoten großen Netze so gute Ergebnisse zu erzielen. Nur so können entsprechend gute Referenzwerte für die Heuristiken erzeugt werden.

**Einfluss der Pilotmethode** Zunächst sollen die Vorteile des Pilotansatzes dargelegt werden, da fast alle der nachfolgend aufgeführten Vergleichsrechnungen mit der Kombination aus GRASP und Pilot durchgeführt werden. Tabelle 5.13 zeigt die Fortschritte, die durch den Einsatz des Piloten erzielt werden. Die Abweichung  $D(G+P,G)$  in der letzten Spalte gibt dabei an, um wie viel Prozent das Pilotergebnis besser ist als das des normalen GRASP. Die Piloterweiterung führt häufig zu einer Verbesserung des Zielfunktionswertes im niedrigen einstelligen Prozentbereich, mindestens aber zu gleich guten Ergebnissen wie der normale GRASP. Die Laufzeit ist meistens etwas länger, da der Pilot eine nachgelagerte Stufe zum normalen GRASP darstellt. In einigen Fällen kann sie jedoch kürzer sein, wenn die Kombination aus GRASP und Pilot schneller konvergiert als der GRASP alleine und somit die Zahl der Iterationen mitunter deutlich niedriger eingestellt werden kann.

Auf den ersten Blick sieht der Vorteil gering aus. Dabei sind jedoch zwei Dinge zu beachten. Einerseits erreicht bereits der normale GRASP-Algorithmus bei fünf der Probleminstanzen eine optimale Lösung (jeweils fett hervorgehoben, vgl. auch Tabelle 5.16), hier kann der Pilot also ohnehin keine Verbesserung mehr erzielen. Andererseits findet der Pilot bei vier weiteren Probleminstanzen eine optimale Lösung, die der normale GRASP-Algorithmus nicht findet. Insgesamt können jetzt also neun von 21 Probleminstanzen optimal gelöst werden. Zusammen mit der Tatsache, dass der Pilot in keinem Fall zu einem schlechteren Ergebnis führt und die Rechenzeit weiterhin sehr moderat ist, in einigen Fällen sogar geringer als beim einfachen GRASP, stellt die Piloterweiterung einen deutlichen Fortschritt dar.

Nicht nur in Verbindung mit GRASP, sondern auch mit anderen Metaheuristiken lassen sich ähnlich positive Effekte erzielen. Ein gutes Beispiel hierfür ist die bereits angesprochene Kombination aus Pilot und VNS. Ta-

belle 5.14 (vgl. Höller *et al.* (2006)) zeigt die Ergebnisse für die gleichen Probleminstanzen, die Tabelle 5.16 für den GRASP darstellt. Ausgehend von einem bereits gut funktionierenden VNS wird durch den Hybrid mit der Pilotmethode ebenfalls ein spürbarer Fortschritt erzielt. Bemerkenswert ist dabei auch die große Ähnlichkeit der Ergebnisse zwischen VNS+Pilot und GRASP+Pilot. Die VNS-Implementierung ist völlig unabhängig von der GRASP-Implementierung entstanden. Und auch die Pilotidee wurde auf eine deutlich andere Weise umgesetzt. Im Endeffekt jedoch führen beide Heuristiken auf ganz unterschiedlichen Wegen zu sehr ähnlichen, guten Ergebnissen.

**Vergleich mit dem MIP-Modell** In Tabelle 5.15 sind einige Ergebnisse der Heuristik für bereits aus Abschnitt 5.1 bekannte, künstlich generierte Netze aufgeführt. Für die sehr kleinen 11-Knoten-Netze findet die Heuristik immer eine optimale Lösung. Dieses Ergebnis ist auch repräsentativ für andere Netze dieser Größe und so stabil, dass die Heuristik bereits bei einer geringen Zahl an Iterationen nahezu immer auf den optimalen Wert konvergiert. Beim 20 Knoten Netz ist sie CPLEX sogar leicht überlegen.

Die Spalte „SP“ zeigt die Kosten für eine Kürzeste-Wege-Lösung (shortest path). Diese Lösung wird erzielt, indem nur die ersten beiden Schritte des Algorithmus in Abbildung 4.4 ausgeführt werden. Somit wird jeder Bedarf auf dem Weg geführt, der von den Kilometern der Glasfaserstrecke her der kürzeste ist. Die bereits ordentlichen Ergebnisse der Kürzeste-Wege-Lösung zeigen, dass sie ein guter Ausgangspunkt für die GRASP-Heuristik ist.

Für die gleichen Probleminstanzen, die bereits aus dem Vergleich mit und ohne Pilot bekannt sind, werden in Tabelle 5.16 nun ebenfalls die Ergebnisse der Pilot-Heuristik mit den Referenzwerten von CPLEX verglichen.

Hier zeigt sich erneut, dass CPLEX für die kleinen Probleminstanzen immer und die Heuristik zumindest sehr häufig ein optimales Ergebnis findet. Die Laufzeit der Heuristik liegt dabei im Bereich von Sekunden, während CPLEX teilweise schon mehrere Stunden dafür benötigt. Diese Diskrepanz in der Laufzeit nimmt mit steigender Netzgröße weiter zu. Für die größten Netze benötigt CPLEX teilweise mehr als einen Tag Rechenzeit, während die Heuristik maximal etwas mehr als zwanzig Minuten benötigt. Zusammen mit der Rechenzeit steigt bei CPLEX auch der Speicherbedarf entsprechend an. Die Lösungsgüte bleibt jedoch weiterhin auf hohem Niveau und ist bei beiden Ansätzen nur wenige Prozent von den bekannten Schranken entfernt.

Die hier aufgeführten 50-Knoten-Netze sind in etwa die Grenze dessen, was noch sinnvoll mit CPLEX zu lösen ist. Darüber hinaus steigt der Speicherbedarf so stark an, dass innerhalb der gewählten 400 MB Baumgröße oft keine brauchbaren Lösungen mehr gefunden werden. Beispielsweise wird

$ N $	$ E $	$ D $	GRASP		GRASP+Pilot		D(G+P,G)
			Kosten	Laufzeit [s]	Kosten	Laufzeit [s]	
12	33	15	<b>3,69</b>	1	<b>3,69</b>	2	0
		21	6,20	13	<b>6,03</b>	12	2,74
		44	13,90	97	13,88	120	0,14
		66	11,83	14	11,83	23	0
12	46	15	<b>3,69</b>	1	<b>3,69</b>	2	0
		21	<b>6,03</b>	1	<b>6,03</b>	2	0
		44	13,90	108	<b>13,62</b>	33	2,01
		66	<b>11,77</b>	145	<b>11,77</b>	48	0
17	26	27	22,70	83	<b>22,47</b>	56	1,01
		36	82,09	69	82,09	35	0
		81	98,42	21	97,04	139	1,40
		135	173,52	33	171,95	240	0,90
17	68	27	<b>19,27</b>	37	<b>19,27</b>	12	0
		36	64,50	51	64,50	133	0
		81	79,60	78	79,60	29	0
		135	141,29	61	141,29	61	0
50	63	45	34,40	6	<b>34,07</b>	62	0,96
		65	49,52	8	49,28	216	0,48
		91	56,83	11	55,67	208	2,04
		112	43,0	7	42,88	90	0,28
		267	111,92	33	110,24	1274	1,50

Tabelle 5.13: Beispiele für den Fortschritt durch das Pilotverfahren (GRASP).

$ N $	$ E $	$ D $	VNS		VNS+Pilot	
			Kosten	Laufzeit [s]	Kosten	Laufzeit [s]
12	33	15	3.69	0.09	3.69	0.48
		21	6.03	0.46	6.03	22.02
		44	14.15	0.05	13.79	85.07
		66	15.48	0.58	12.49	19.91
12	46	15	3.69	0.02	3.69	0.53
		21	6.15	5.54	6.03	11.87
		44	14.79	0.04	14.39	36.71
		66	13.03	1.49	12.41	1.43
17	26	27	23.81	1.33	22.47	40.97
		36	81.84	8.32	81.84	2.17
		81	97.47	36.62	97.03	104.72
		135	173.11	97.19	172.55	510.11
17	68	27	19.45	26.83	19.27	137.24
		36	64.72	10.54	64.02	189.25
		81	80.75	30.01	78.46	305.08
		135	145.07	53.96	141.75	513.54
50	63	45	36.38	37.07	35.37	357.78
		65	46.74	2.86	46.74	565.10
		91	59.65	58.08	57.00	637.05
		112	41.63	0.27	41.63	756.31
		267	111.70	14.64	110.40	1326.83

Tabelle 5.14: Beispiele für den Fortschritt durch das Pilotverfahren (VNS).

$ N $	$ E $	$ D $	SP	GRASP+Pilot		CPLEX		Schranke
			Kosten	Kosten	Laufzeit [s]	Kosten	Laufzeit [s]	
11	16	19	2079	1983	2	1983	0,4	opt
11	16	19	10577	10442	4	10442	1,9	opt
11	16	19	1347	1231	5	1231	2,7	opt
20	33	63	9728	9405	258	9426	9362	9281

Tabelle 5.15: Vergleich für kleine, künstlich generierte Netze.

$ N $	$ E $	$ D $	GRASP+Pilot		CPLEX			D(G,C)
			Kosten	Laufzeit [s]	Kosten	Laufzeit [s]	Schranke	
12	33	15	3,69	2	3,69	7	opt	0
		21	6,03	12	6,03	67	opt	0
		44	13,88	120	13,66	4228	opt	1,59
		66	11,83	23	11,77	9024	opt	0,51
	46	15	3,69	2	3,69	12	opt	0
		21	6,03	2	6,03	165	opt	0
		44	13,62	33	13,16	6148	opt	3,38
		66	11,77	48	11,77	11194	opt	0
17	26	27	22,47	56	22,47	3	opt	0
		36	82,09	35	81,84	31	opt	0,30
		81	97,04	139	96,65	470	opt	0,40
		135	171,95	240	170,15	29255	opt	1,05
	68	27	19,27	12	19,27	110	opt	0
		36	64,50	133	63,75	28841	opt	1,16
		81	79,60	29	76,66	216806	75,21	3,69
		135	141,29	61	138,40	81098	130,98	2,05
50	63	45	34,07	62	34,07	854	opt	0
		65	49,28	216	48,65	77493	45,56	1,28
		91	55,67	208	55,64	95755	53,19	0,05
		112	42,88	90	42,88	39544	38,67	0
		267	110,24	1274	103,64	90768	102,74	5,97

Tabelle 5.16: Vergleich zwischen CPLEX und der Heuristik für mittelgroße Probleminstanzen.

für das 111-Knoten-Netz aus Tabelle 5.17 nicht einmal mehr eine zulässige Lösung gefunden. Eine Erhöhung des Speichers führt jedoch zu unpraktikablen Rechenzeiten, so dass hier nur die Heuristik als sinnvolles Verfahren verbleibt. Die guten Ergebnisse im Vergleich zu den optimalen Lösungen bzw. Schranken für die kleinen und mittleren Netzgrößen legen die Schlussfolgerung nahe, dass die Heuristik auch bei noch größeren Probleminstanzen konstant gute Ergebnisse liefert. Eine formale Garantie dafür kann jedoch prinzipbedingt nicht gegeben werden.

Tabelle 5.17 zeigt Ergebnisse für sehr große, reale Probleminstanzen. Im Gegensatz zu den ganz oder teilweise künstlich generierten oder zumindest modifizierten Probleminstanzen, die bisher betrachtet wurden, stehen hier

$ N $	$ E $	$ D $	CPLEX	SP	GRASP
39	81	163	2270	2652	2206
111	160	243	-	117336	108661
111	160	243	-	333909	313145
321	vollvermascht	1896	-	52371	51298

Tabelle 5.17: Vergleich für große, reale Problem instanzen.

naturgemäß nur sehr wenige Datensätze zur Verfügung. Beim 39-Knoten-Netz handelt es sich um ein typisches Regionalnetz einer Großstadt mit Umland, dies wird im Detail in Abschnitt 5.3.2 beschrieben. Das 111-Knoten-Netz ist ein eher grobmaschiges landesweites Netz, einmal mit ungeschützten Bedarfen und ein Mal mit 1+1 geschützten Bedarfen gerechnet. Das 321-Knoten-Netz basiert auf den Standorten eines landesweiten Netzes, hier ist die Glasfasertopologie jedoch nicht vorgegeben.

Während CPLEX für das 39-Knoten-Netz noch eine Lösung findet, wenn auch eine schlechtere als die Heuristik, sind für die größeren Netze keine brauchbaren Resultate mehr zu erreichen. Hier dient ein Routing aller Bedarfe auf dem geographisch kürzesten Weg als Referenz. Beim 111-Knoten-Netz wird eine deutliche Verbesserung gegenüber der Kürzeste-Wege-Lösung erzielt, die in der Regel bereits ein guter Ausgangspunkt ist (vgl. auch Tabelle 5.17). Somit ist die Heuristik auch bei großen, realen Netzen gut einsetzbar. Die Rechenzeiten für diese großen Netze sind stark schwankend. Sie liegen aber noch im Minutenbereich und sind somit praxistauglich.

Wenn von einer Vollvermaschung ausgegangen wird, wie beim 321-Knoten-Netz, dann zeigen sich jedoch die Grenzen des hier gewählten Ansatzes. Obwohl der Algorithmus in der Lage ist, bestimmte Kanten aus der Gesamtmenge für die Lösung auszuwählen und andere nicht zu nutzen (in diesem Fall werden 1872 der 51360 theoretisch möglichen Kanten ausgewählt), liegen die Stärken des Verfahrens doch eindeutig bei der Berechnung mit vorgegebener Glasfaserstruktur. Ohne eine grundlegende Vorgabe der Netztopologie ist eine zielgerichtete Suche mit dem gewählten Ansatz nicht möglich. Dies war jedoch auch kein Entwicklungsziel, insofern ist dieses Ergebnis nicht überraschend. Über eine Modifizierung der Kostenfunktionen könnte hier noch eine Verbesserung erzielt werden, beispielsweise indem die Kosten für die Aufnahme einer Kante in die Lösung künstlich erhöht werden. Dies ist jedoch eher als Notlösung zu betrachten und entspricht nicht dem ursprünglichen Zweck des Planungsverfahrens.

**1+1 geschützte Netze** Tabelle 5.18 zeigt die Ergebnisse des GRASP+Pilot-Verfahrens im Vergleich zu CPLEX für 1+1 geschützte Netze. Nicht alle der untersuchten Probleminstanzen verfügen über eine zweifach zusammenhängende Glasfaserstruktur. Deshalb sind die Berechnungen für 1+1 geschützte Netze wiederum auf eine Teilmenge beschränkt. Im Gegensatz zu den ungeschützten Netzen hat sich hier eine Pilotsuche über Knoten- und Kantenmodifikationen bewährt und wird durchgehend genutzt. Die Ergebnisse sind wiederum sehr gut; bei insbesondere für die größeren Probleminstanzen deutlich geringeren Rechenzeiten erzielt die Heuristik durchgehend Werte, die nur wenige Prozent von der CPLEX-Lösung entfernt sind. In einem Fall ist sie sogar besser.

$ N $	$ E $	$ D $	GRASP+Pilot		CPLEX			
			Cost	Time [s]	Cost	Time [s]	Bound	D(G,C)
11	16	19	4082	5	4082	2,7	opt	0
12	33	15	7,43	3	7,30	50	opt	1,75
		21	10,76	38	10,11	52	opt	6,04
		44	25,99	54	25,82	10771	opt	0,65
	46	66	22,80	55	22,38	15047	22,08	1,84
		15	7,26	21	7,26	282	opt	0
		21	10,83	10	10,11	440	opt	6,65
		44	26,31	54	25,39	17345	22,67	3,50
	66	22,50	28	22,52	23398	18,72	-0,09	
17	26	27	52,58	5	52,14	4	opt	0,84
		36	203,87	9	203,26	118	opt	0,30
	135	81	249,22	28	248,26	514	opt	0,39
		135	437,07	46	435,25	455	opt	0,42

Tabelle 5.18: Vergleich zwischen CPLEX und der Heuristik für 1+1 geschützte Probleminstanzen.

**Abschließende Bemerkungen** Wie in Abschnitt 4.2.1 erwähnt, kann das Verhalten des GRASP-Algorithmus in weiten Grenzen eingestellt werden. Insbesondere das Verhältnis zwischen Gierigkeit und Zufälligkeit der Suche ist hier interessant. Abbildung 5.2 verdeutlicht diesen Effekt für ein kleines



Beispielnetz. Die Kurven zeigen die jeweils beste aktuelle Lösung für 100 Iterationen. Um trotz der Zufallskomponente einen aussagekräftigen Graphen zu erhalten, zeigen die Kurven über 100 unabhängige Neustarts gemittelte Werte.

Ein stark gieriges Verhalten entspricht der Einstellung „g10“. Dies führt zu einem schnellen Fortschritt am Anfang, aber die Suche ist zu eingeschränkt, um langfristig ein gutes Ergebnis erzielen zu können. Auf der anderen Seite dauert es bei einer deutlich weniger gierigen Strategie „g2“ etwas länger, gute Ergebnisse zu erzielen, dafür ist das Endergebnis jedoch deutlich besser. Die Einstellung „g4“ liegt erwartungsgemäß dazwischen. Weitgehend unabhängig von der konkreten Probleminstanz hat sich die Einstellung „g2“ als universell guter Wert herausgestellt. Sie wird folglich für alle in dieser Arbeit aufgeführten Berechnungen genutzt.

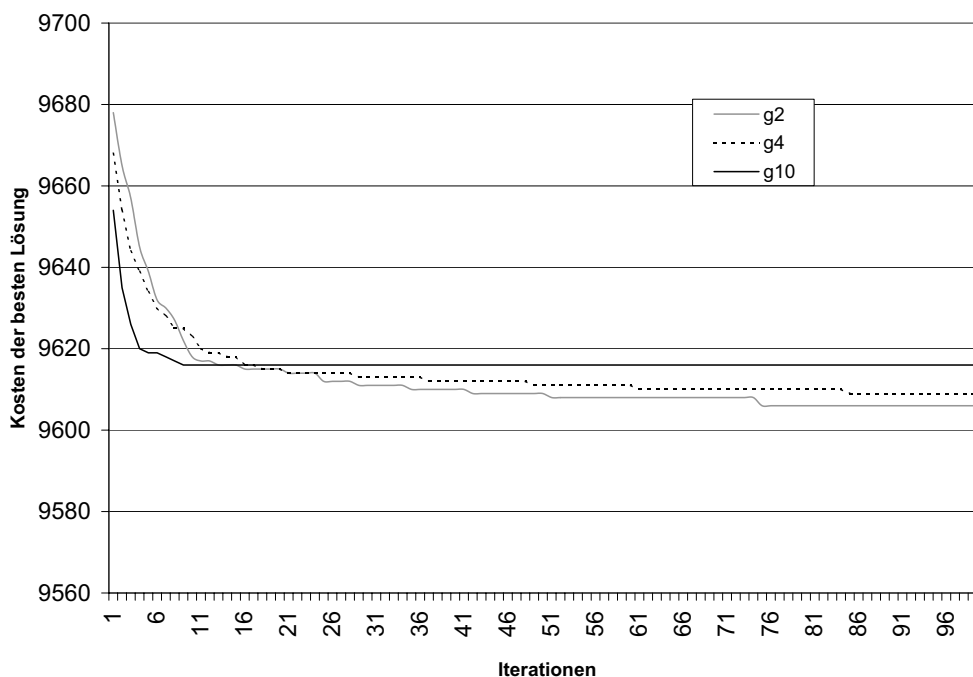


Abbildung 5.2: Einfluss der Gierigkeit der Lösungsstrategie.

Daraus könnte man den Schluss ziehen: Je zufälliger, desto besser. Dass dies nicht so ist und die gezielte Suche des GRASP erfolgreich funktioniert, zeigt ein Vergleich mit einer alten Version des Algorithmus (hier „random start“ genannt), der eine ungerichtete, vollständig zufallsgesteuerte lokale Suche ausführt. In Abbildung 5.3 ist zu erkennen, dass der GRASP deutlich schneller in gute Bereiche des Lösungsraumes findet. Bei ausreichender

Laufzeit konvergieren beide Verfahren zum gleichen Ergebnis. Dieses Verhalten kann allgemein beobachtet werden. Der GRASP-Algorithmus ist deutlich vorteilhaft, wenn die Rechenzeit begrenzt ist, und damit hauptsächlich wenn sehr viele Szenarien durchgespielt werden sollen oder sehr große Netze geplant werden. Vor allem hat er sich bei der oben beschriebenen Einstellung „g2“ auch bei sehr hoher Zahl an Iterationen bei keiner getesteten Probleminstanz als nachteilig gegenüber der zufälligen Suche erwiesen. Die Pilotphase setzt erst nach dem hier visualisierten Teil der Suche ein. Ausgehend von dem durch den GRASP-Algorithmus erreichten Plateau würde der Zielfunktionswert ab dem Start des Piloten, also hier ab Iteration 100, ggf. weiter absinken.

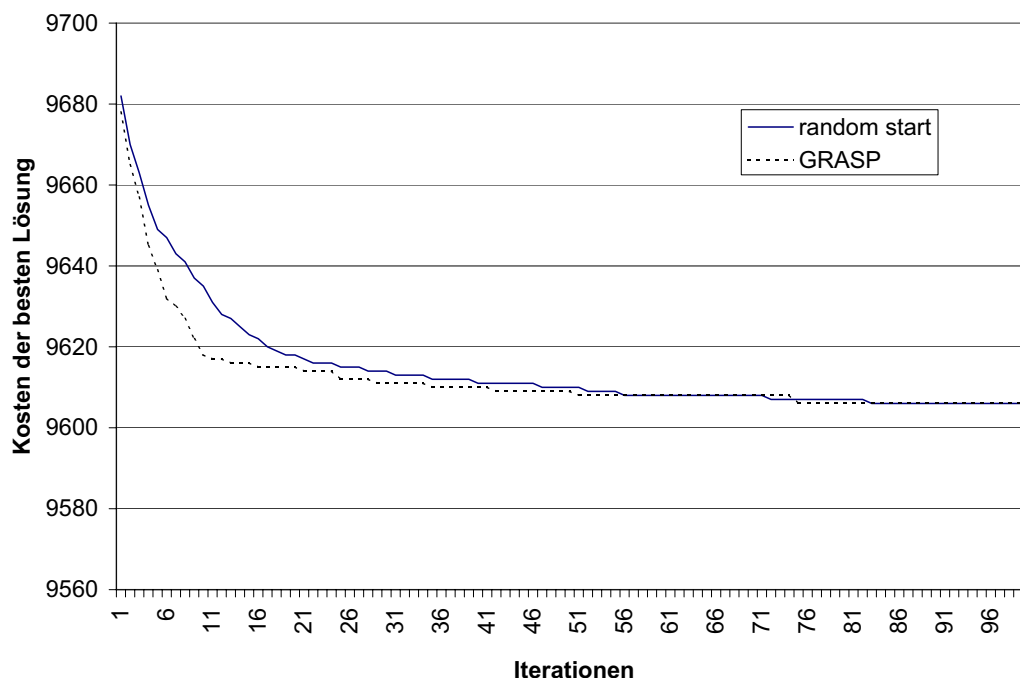


Abbildung 5.3: Vorteil von GRASP gegenüber einem Verfahren ohne gesteuerte lokale Suche.

Zusammenfassend lassen sich folgende Erkenntnisse festhalten. Der Einsatz der Pilotmethode bringt einen deutlichen Fortschritt für die Qualität der Lösungen der GRASP-Heuristik. Das eigentlich relativ einfach strukturierte Vorausschauen als Kern der Pilotmethode ist also eine wirkungsvolle Ergänzung ohne negative Nebeneffekte. Dies deckt sich mit den in Höller *et al.* (2006) gewonnenen Ergebnissen zu einer Kombination aus VNS und Pilot für die gleiche Problemstellung. Der Vergleich mit dem VNS kann als Indiz

dafür gewertet werden, dass eine zielgerichtete Adaption der jeweiligen heuristischen Grundideen an die praktische Problemstellung entscheidend sein kann. Möglicherweise sogar viel entscheidender als die in der Literatur oft diskutierte Frage, ob eine bestimmte Heuristik (sei es GRASP, VNS, Tabu Search, usw.) einer anderen grundsätzlich überlegen ist.

Weiterhin zeigt sich deutlich, dass es trotz großer Fortschritte bei der ganzzahligen Optimierung sinnvoll ist, auch heuristische Konzepte einzusetzen. Auch wenn CPLEX für sehr viele der hier betrachteten Probleminstanzen gute oder sogar optimale Lösungen erzielt, so zeigen die extrem langen Rechenzeiten bei großen Problemstellungen doch eindeutig die aktuelle Grenze dieses Ansatzes auf. Für die größten Probleminstanzen kann nicht einmal eine zulässige Lösung gefunden werden.

Im Gegensatz dazu skaliert die Heuristik deutlich besser mit der Problemgröße und ist für alle in der Praxis auftretenden Problemgrößen geeignet. Zusammen mit der nachweislich sehr guten Ergebnisgüte, die im Normalfall höchstens wenige Prozent vom Optimum abweicht, stellt das heuristische Vorgehen ein leistungsfähiges und schnelles Verfahren als Hilfsmittel für die praktische Netzplanung bereit.

### 5.3 Fallstudien

Auch wenn eine Reihe der Probleminstanzen aus den Abschnitten 5.1 und 5.2 auf realen Datensätzen basieren, so ist die Absicht hinter den dort durchgeführten Berechnungen eher die Evaluierung der Verfahren als die Lösung einer praktischen Problemstellung. Im Rahmen dieses Abschnittes werden zwei Beispiele vorgestellt, bei denen die Verfahren in Zusammenarbeit mit Praxispartnern gezielt für die Bearbeitung von netzplanerischen Fragestellungen aus der Praxis genutzt werden. Hierzu waren mitunter Modifikationen der vorgestellten Basisalgorithmen notwendig, um bestimmte zusätzliche Randbedingungen aus dem Bereich der Geräte oder des Netzbetriebs abbilden zu können. Trotzdem handelt es sich auch bei den Fallstudien vom Zeithorizont und Abstraktionsniveau her eher um strategische (bzw. mit Einschränkungen taktische) als um operative Netzplanung. Darüber hinaus haben sich die Verfahren auch bei verschiedenen weiteren Planspielen zusammen mit Praxispartnern bewährt. Auf diese kann hier jedoch auf Grund des vertraulichen Charakters der dabei genutzten Datenbasis nicht eingegangen werden.

### 5.3.1 Fallstudie zum Einsatz von Carrier Ethernet

Ethernet etabliert sich zunehmend auch im Bereich großer Carriernetze als Konkurrenz zur bestehenden Technologie. In dieser Fallstudie werden drei Szenarien für die technologische Realisierung eines großen, nationalen Backbonenetzes verglichen. Insbesondere werden Ansätze zur Integration von leitungs- und paketvermittelnder Technologie in Betracht gezogen. Ziel ist es, die grundsätzlichen Unterschiede der Designmöglichkeiten herauszuarbeiten und das jeweils daraus entstehende, unterschiedliche Investitionsvolumen für die Realisierung eines solchen Netzes zu vergleichen.

Carrier Ethernet unterscheidet sich vom in Abschnitt 1.2.2 vorgestellten „herkömmlichen“ Ethernet im Wesentlichen durch zusätzliche Funktionalitäten aus den Bereichen Skalierbarkeit und Monitoring. Große Netzbetreiber haben deutlich höhere Anforderungen an das Netzmanagement und die Überwachung der ordnungsgemäßen Funktion als dies bei lokalen, meistens firmeninternen Netzen, für die Ethernet ursprünglich entwickelt wurde, der Fall ist. Details zu Carrier Ethernet im Allgemeinen und der hier eingesetzten Hardware im Speziellen sind aus Fricke *et al.* (2005) zu entnehmen.

**Problemstellung** Die der Problemstellung zugrunde liegende Datenbasis ist ein deutschlandweites Netz mit 142 Knoten. Die dazu gehörige Bedarfsmatrix enthält 6700 Bedarfe zwischen 142 Standorten. Diese Bedarfe sind im Gegensatz zu den meisten anderen Berechnungen in dieser Arbeit unidirektional. Das Datenaufkommen zwischen zwei Knoten  $A$  und  $B$  kann also in der Richtung  $A \rightarrow B$  ein anderes Volumen besitzen und einen anderen Weg nehmen als in der Gegenrichtung  $B \rightarrow A$ .

Dieses Ausgangsnetz wird durch Zusammenfassen von Standorten mit mehreren Einzelknoten zu einem Gesamtknoten auf ein Netz aus 74 Knoten und 208 Kanten reduziert. An diesen 74 Knoten stehen Label Edge Router (LER), die den Verkehr aus dem Access-Bereich bündeln. Für die Bedarfsmatrix ergibt sich somit eine logische Vollvermaschung; es besteht eine bidirektionale Verkehrsbeziehung zwischen allen Knotenpaaren. Die überwiegende Zahl dieser Bedarfe (93%) ist dabei kleiner als 100 Mbit/s, die verbleibenden 7% der Bedarfe verteilen sich auf den Bereich zwischen 100 Mbit/s und 6,6 Gbit/s. Das anvisierte Netz muss nicht nur in der Lage sein, alle diese Bedarfe zu führen, sondern darüber hinaus auch Reservekapazität für einen beliebigen Kanten- oder Knotenausfall zur Verfügung zu stellen.

Als Referenzlösung dient ein ausschließlich aus Routern aufgebautes Netz, welches 100% Redundanz enthält, um eine hohe Ausfallsicherheit zu gewährleisten. Im Prinzip besteht die Struktur aus zwei Teilnetzen, die jeweils die Hälfte des Verkehrs führen und im Normalbetrieb nur bis maximal 50%

ausgelastet werden. Dieses reine Routernetzwerk soll mit zwei alternativen Szenarien verglichen werden.

In diesem Backbonenetz wird die paketbasierte Routerfunktionalität der Label Switching Router (LSR) unabhängig vom Szenario nicht eingesetzt. Es findet ein reines Labelswitching oder bei den alternativen Szenarien ein äquivalentes Verfahren mit fest geschalteten Verbindungen (z.B. SDH-VC-Pfade) auf a priori festgelegten Pfaden statt. Diese Datenbasis und Netzstruktur sind gegenüber dem zugrunde liegenden realen Netz abstrahiert, aber noch hinreichend ähnlich.

**Szenarien** Das Referenzszenario (I) ist in Abbildung 5.4 dargestellt. Die Router sind direkt untereinander vernetzt. Ob für die Punkt-zu-Punkt-Verbindungen SDH- oder Ethernet-Portkarten eingesetzt werden, ist für die prinzipielle Struktur unerheblich; jedoch sind Ethernet-Karten hier etwas preisgünstiger. Da das Netz physikalisch nicht vollständig vermascht ist, müssen die Router nicht nur ihre jeweiligen Quellverkehre verarbeiten, sondern zusätzlich noch den Transitverkehr anderer Verbindungen. Im Prinzip ist die Unterscheidung in LER und LSR hier nebensächlich. Explizite LSR werden streng genommen nur dann benötigt, wenn der Transitverkehr so groß ist, dass er vom LER nicht bewältigt werden kann. Sonst können alle LER/LSR-Funktionen von einem einzigen, entsprechend leistungsfähigen Gerät übernommen werden.

Die Szenarien (II) und (III) setzen ebenfalls an jedem Standort einen LER voraus. Allerdings sind die Router auf der Core-Seite nicht direkt miteinander verbunden, sondern an ein getrenntes Layer 1/2 Netz angeschlossen (vgl. Abbildung 5.5). Zur Erhöhung der Ausfallsicherheit ist diese Kopplung verdoppelt. Das Core-Netz besteht entweder aus SDH-Cross-Connects (II) oder Ethernet Switches (III). Die grundsätzliche Netzstruktur ist in beiden Fällen identisch, der Unterschied ist primär die Layer 2 Technologie. Beim Einsatz von Ethernet besteht darüber hinaus noch die Wahl zwischen MPLS und Virtual Local Area Network (VLAN) als Traffic-Engineering-Mechanismus. Details dazu finden sich in Fricke *et al.* (2005), sie sind für den hier beschriebenen Planungsprozess nicht relevant. Damit werden die Router in diesen Szenarien vollständig vom Transitverkehr entlastet. Abbildung 5.6 zeigt die Struktur eines solchen Netzknotens mit seinen maßgeblich für die Gesamtkosten verantwortlichen Komponenten.

**Netzplanung** Aufgrund der Größe des Netzes kommt nur die in Abschnitt 4.2 vorgestellte Heuristik und nicht CPLEX für diese Aufgabenstellung in Frage. Alle benötigten Funktionalitäten, inklusive Grooming und

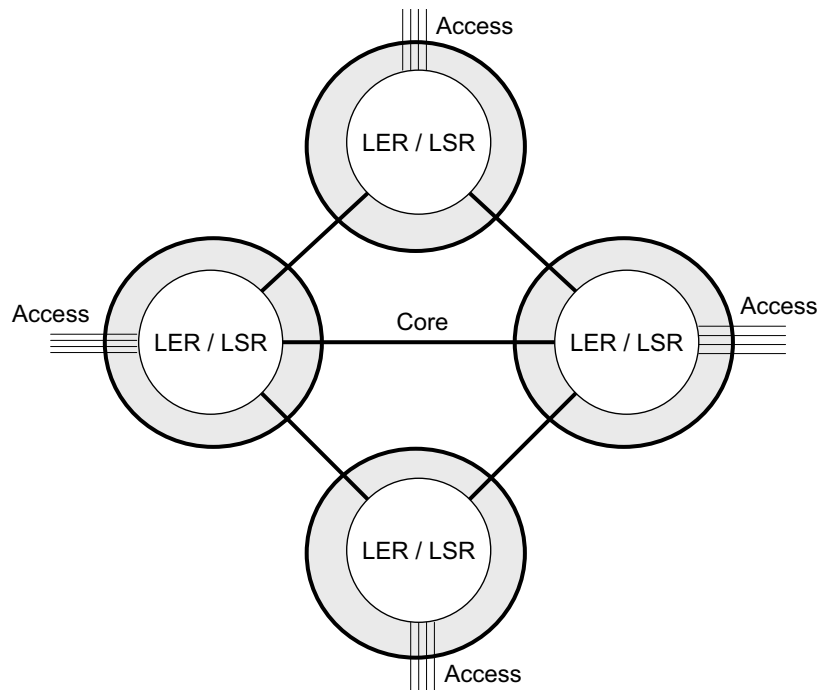


Abbildung 5.4: Netzstruktur bei Szenario (I).

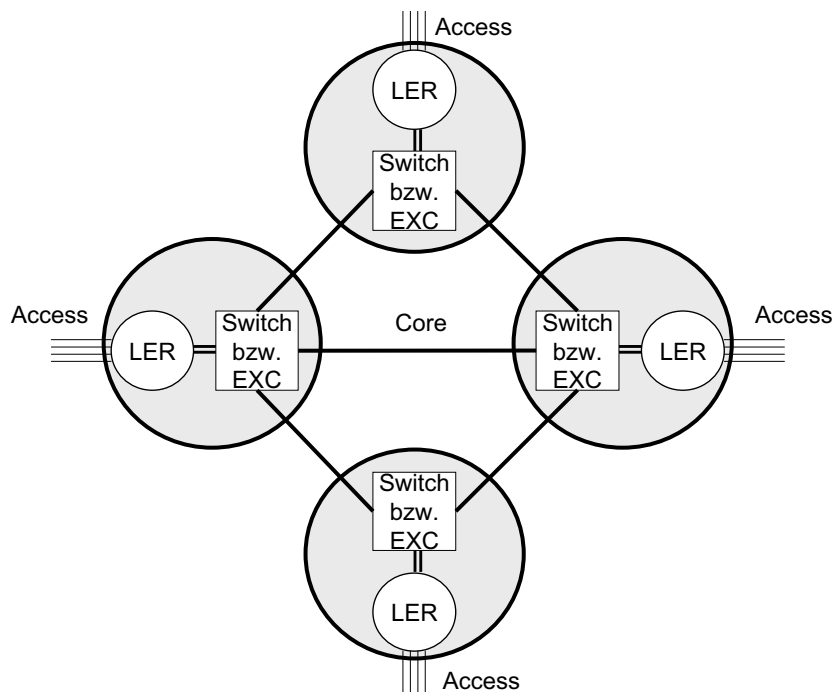


Abbildung 5.5: Netzstruktur bei den Szenarien (II) und (III).

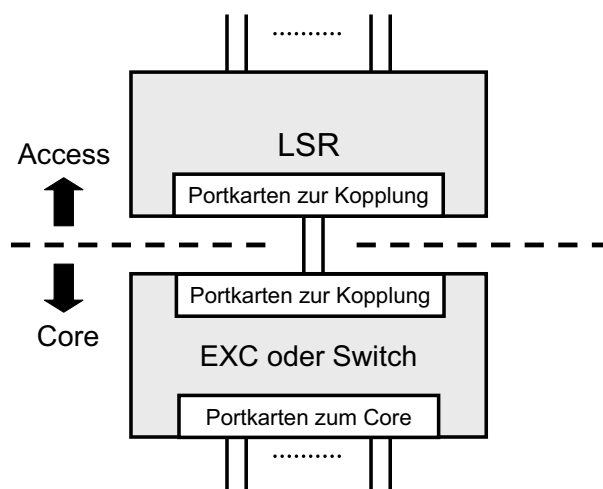


Abbildung 5.6: Der Aufbau eines Netzknotens bei den Szenarien (II) und (III) im Detail.

unidirektionaler Wegführung, werden von der Heuristik unterstützt. Somit verbleibt lediglich die Anpassung an die spezielle Hardwarestruktur als problemspezifische Modifikation. Da es bei dieser Aufgabenstellung eine klar definierte Trennung zwischen den Routern und den Switchen gibt, muss auch das eigentliche Planungsverfahren nicht modifiziert werden. Eine gemischte Führung des Transitverkehrs über diese beiden Technologien hinweg hätte die Problemstellung stark verändert und eine umfassende Modifikation der Heuristik notwendig gemacht. Eine solche Führung war hier jedoch nicht gefordert.

Die Planung wird dabei auf Basis der aus den Listenpreisen abgeleiteten stufigen, also nichtlinearen, Gerätekosten durchgeführt. Die Preise für die Geräte unterschiedlicher Technologien sind alle vom gleichen Hersteller für diese Fallstudie zur Verfügung gestellt. Somit ist sichergestellt, dass die Preisrelationen realistisch sind und nicht die eine oder andere Variante durch die, verglichen mit den tatsächlichen Preisen für Großkunden etwas willkürlichen Listenpreisen, übervorteilt wird.

Die Modellierung des Netzes umfasst im Wesentlichen die Komponenten, die in den einzelnen Szenarien bereits aufgeführt wurden. Den Hauptkostenblock bilden die Portkarten (2,5 Gbit/s und 10 Gbit/s SDH-Karten bzw. 1 Gbit/s und 10 Gbit/s Ethernet-Karten). Alle Karten sind Varianten mit verhältnismäßig kurzer Reichweite, da jeweils nur lokal eine Verbindung zum WDM-Multiplexer hergestellt werden muss. Eine direkte Verbindung ohne WDM ist aufgrund der langen Faserstrecken zwischen den Knoten in

den meisten Fällen ohnehin nicht sinnvoll. Daneben werden die Kosten für die Basisgeräte berücksichtigt (Router oder Cross-Connect oder Ethernet-Switch). Deren Listenpreise sind jedoch gegenüber denen der Portkarten vor allem bei den Routern eher gering.

Es wird vorausgesetzt, dass Glasfasern und WDM-Systeme in ausreichender Zahl vorhanden sind. Diese Kosten sind für alle Szenarien identisch und müssen für den Vergleich deshalb nicht näher spezifiziert werden (szenariofixe Kosten). Allerdings benötigen Ethernet-Verbindungen bei geringen Bandbreiten, also wenn sich 10 Gbit/s noch nicht lohnt, unter Umständen mehr WDM-Kanäle und damit mehr Transponder als SDH (1 Gbit/s gegenüber 2,5 Gbit/s pro Kanal) bzw. aufwändigere Multiplex-Transponder (welche hier nicht betrachtet werden). Diese erhöhten Transponderkosten fließen in den Kostenvergleich ein. Auf Basis der Kosteninformationen, der Netzstruktur und der Bedarfsdaten werden die einzelnen Szenarien jeweils getrennt mit Hilfe des heuristischen Softwaretools berechnet und ausgewertet.

Die Betrachtung der Kosten soll hier unter gleichen Voraussetzungen erfolgen; deshalb wird für alle Szenarien eine disjunkte Wegeführung mit 100% Redundanz zugrunde gelegt. Bei SDH entspricht dies 1+1 Protection, bei Ethernet ist ein äquivalenter Schutz beispielsweise über MPLS-Mechanismen realisierbar. Natürlich ermöglichen alle betrachteten technologischen Varianten kosteneffizientere Schutzverfahren (Restoration, Fast-Reroute usw.), diese haben aber wiederum unterschiedliche Vor- und Nachteile auf anderen Ebenen (betriebliche Aspekte, Reaktionsgeschwindigkeit usw.), so dass ein Vergleich unter gleichen Bedingungen nicht mehr möglich wäre.

Unabhängig davon werden jedoch allgemeine technologiespezifische Vorteile berücksichtigt. So haben die Ethernet- und Routerszenarien den Vorteil, dass das Grooming in Transitknoten nahezu byteweise erfolgen kann, während bei SDH das Grooming auf VC4-Ebene (also in 155 Mbit/s Containern) stattfindet. Weiterhin hat die SDH-Technologie den Nachteil der starr symmetrischen Bidirektionalität. Stark asymmetrische Bedarfe erzeugen eine große Zahl an leeren Containern in einer Richtung. Bei Router- bzw. Ethernet-Netzen besteht dagegen die Möglichkeit, dass sich gegensätzliche Asymmetrien (auch auf Teilstrecken) ausgleichen. Auch dies führt auf einigen Strecken dazu, dass im Gegensatz zur Referenzlösung zusätzliche Transponder, diesmal im SDH-Szenario, benötigt werden. Dieser Effekt kommt konkret dann zum Tragen, wenn durch die Leercontainer bei SDH gerade die Kapazitätsgrenze einer Portkarte überschritten wird, bei Ethernet somit aber noch nicht.



**Ergebnisse** Bei den Ergebnissen ist grundsätzlich noch einmal darauf hinzuweisen, dass nur die szenariovariablen Kosten dargestellt sind. Komponenten, die bei allen Varianten identisch sind, wie z.B. die LERs und die WDM-Systeme, sind nicht in den Zahlen enthalten.

Das Ziel der Planung ist die Minimierung der Gesamtinvestitionen in die Geräte. Abbildung 5.7 zeigt die Differenz der Gesamtinvestitionen für alle Szenarien im Vergleich. Szenario (Ia) dient als Referenzszenario; hier werden die LSRs Punkt-zu-Punkt über SDH-Schnittstellenkarten verbunden (sowie selbstverständlich, wie bei allen anderen Szenarien auch, über WDM-Systeme). Die hohen Kosten der Schnittstellenkarten für Router führen hier zu den höchsten Gesamtkosten aller Szenarien; andere szenariovariable Kosten treten nicht auf. Router-Basisgeräte werden für die LER-Funktionalität in allen Szenarien gleichermaßen benötigt. Die in Szenario (Ia) ggf. auftretenden Mehrkosten für Basisgeräte größerer Kapazität sind gegenüber den Portkartenkosten vernachlässigbar gering und hier deshalb nicht dargestellt.

Die Variation mit Ethernetkarten für die Punkt-zu-Punkt-Verbindungen der LSRs (Ib) zeichnet sich durch niedrigere Kartenkosten aus, welche die zusätzlichen Transponderkosten mehr als kompensieren. Wird darüber hinaus der Transitverkehr aus den Routern herausgenommen und auf dem Layer 1/2 geführt, so sinken die Gesamtinvestitionen nochmals deutlich. Szenario (II) zeigt dies für SDH. Die deutlich geringeren Portkartenkosten wiegen den zusätzlichen Aufwand für die EXC-Basisgeräte, die Kosten für die Kopplung zwischen LER und EXC, die beschriebenen Nachteile bei asymmetrischen Bedarfen - dadurch werden etwa 10% mehr Portkarten benötigt - und die gröbere Grooming-Granularität mehr als auf.

Ein ähnliches Bild ergibt sich für das Ethernet-Szenario (III). Die im Vergleich zu SDH nochmals etwas geringeren Gerätekosten führen zusammen mit den Bidirektionalitätsgewinnen zu weiter reduzierten Gesamtinvestitionen. Der analog zum Szenario (Ib) auftretende Mehraufwand für Transponder reduziert den Vorteil wieder etwas, ändert aber nichts daran, dass Szenario (III) die insgesamt kostengünstigste Lösung ist.

Die Untersuchungen haben deutlich gezeigt, dass die Verlagerung von Transitverkehr aus den Routern heraus auf kostengünstigere Layer 2 Technologie (ggf. in Verbindung mit MPLS) bei den hier betrachteten Randbedingungen zu signifikanten Einsparungen bei den Investitionen führt. Es sollte also bei der Planung solcher IP-Netze sehr genau darauf geachtet werden, ob die zusätzlichen Funktionen, die ein LSR bieten kann, wirklich benötigt werden oder ob eine kostengünstigere Layer 2 Alternative nicht bereits alle Anforderungen erfüllen kann. Dies gilt sowohl für die SDH- als auch für die Carrier-Ethernet-Variante. Zwischen diesen beiden Alternativen bietet Ethernet bei einem völlig neu aufzubauenden Netz das etwas größere Ein-

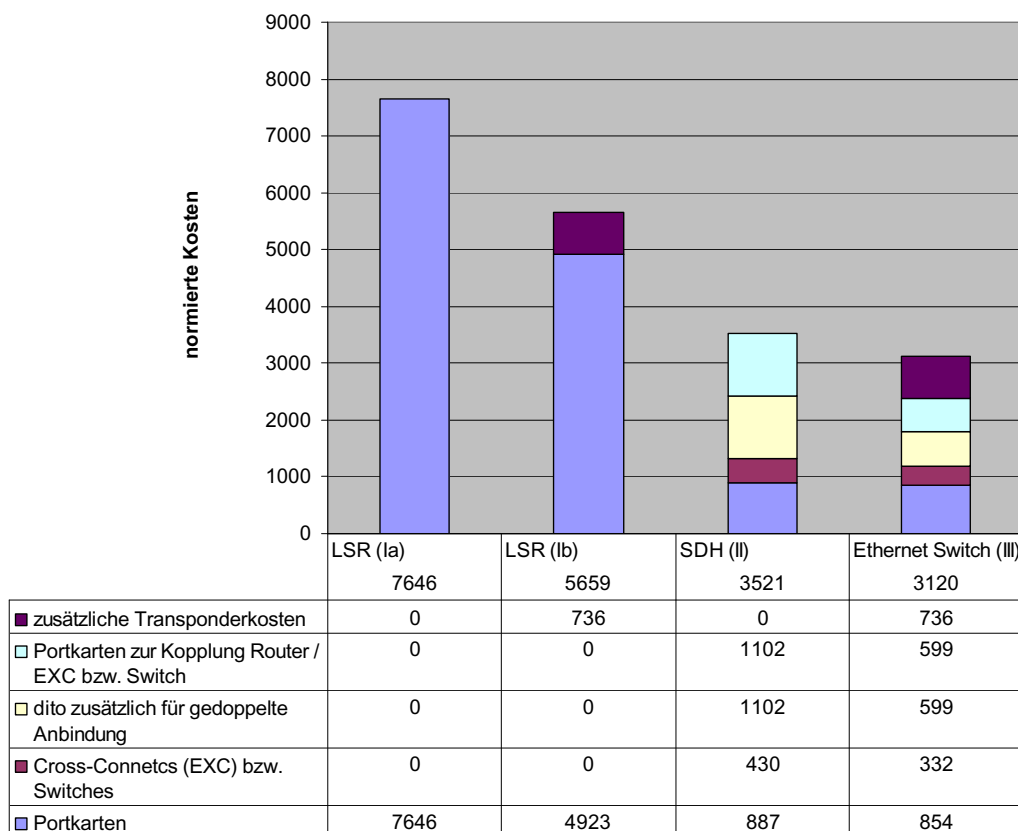


Abbildung 5.7: Normierte szenariovariable Kosten.

sparpotential. Wenn die Möglichkeit besteht, auf schon vorhandene Transporthetzstrukturen zurückzugreifen, so muss dies im konkreten Fall natürlich ebenfalls berücksichtigt werden. Weiterhin sind für den praktischen Einsatz selbstverständlich die betrieblichen Aspekte (Netzwerkmanagement, Personalbedarf bzw. -schulung usw.) zu berücksichtigen und gegen die Kostenersparnis bei der Infrastruktur aufzuwiegen. Eine Entscheidung zu Gunsten einer bestimmten Variante ist alleine auf Basis der hier durchgeführten Geräteplanung sicher nicht sinnvoll.

### 5.3.2 Fallstudie zum Vergleich verschiedener Schutzverfahren

In Abschnitt 1.2.4 werden verschiedene Schutzverfahren mit ihren theoretischen Vor- und Nachteilen diskutiert. Hier wird nun anhand von zwei realen

Netztopologien ermittelt, welche Kosten für die Geräte auftreten, wenn das Netz entweder mit 1+1 Path Protection oder mit statischem Restoration mit stub release realisiert wird. Insbesondere wird auch der Einfluss verschiedener Bedarfsstrukturen untersucht. In der Literatur findet sich eine von der Idee her verwandte Untersuchung speziell mit Bezug auf den Einfluss der Vermaschungsdichte in Doucette und Grover (2001). Die folgenden Abschnitte enthalten im Kern eine Zusammenfassung der entsprechenden Abschnitte aus der ursprünglichen Veröffentlichung der Fallstudie (Höller *et al.* (2004)). Dort finden sich auch Details zu den eingesetzten SDH- und WDM-Geräten.

Die hier untersuchten Schutzverfahren sind zum einen das klassische 1+1 Protection, also zwei disjunkte Wege von der Quelle zur Senke, die beide das Signal führen, und zum anderen ein vorberechnetes Restorationverfahren. Bei diesem Restorationverfahren wird zu jedem primären Pfad ein disjunkter Backuppfad berechnet, aber noch nicht belegt. Erst im Fehlerfall des Erstweges wird auf diesen Pfad umgeschaltet. Somit können sich Erstwege, die disjunkt sind, Backup-Ressourcen teilen.

Es wird eine fehlerabhängige vorberechnete Restoration eingesetzt, der Backuppfad hängt also vom spezifischen Fehler im Primärpfad ab. Abschnitte des Primärpfades, die nicht vom Fehler betroffen sind, können also auch vom Backuppfad genutzt werden. Diese Restoration-Variante hat dadurch gewöhnlich einen geringeren Ressourcenverbrauch als Varianten mit einem einzelnen fehlerunabhängigen Backupweg. Durch die vorberechneten Ersatzwege folgt die Umschaltung im Fehlerfall gerade bei großen Netzen im Allgemeinen deutlich schneller als bei dynamischen Restorationverfahren, die erst im Fehlerfall den Ersatzweg berechnen.

Dieses Restorationverfahren wird nur von der Heuristik unterstützt, nicht jedoch vom AMPL-Modell. Folglich kann für die Ergebnisse der Berechnungen keine Optimalität garantiert werden. Es muss bei der Interpretation der Ergebnisse immer die mögliche Schwankung der Lösungsgüte der Heuristik berücksichtigt werden, die jedoch meistens nur bei wenigen Prozent liegt (vgl. Abschnitt 5.2).

**Netztopologie und Bedarfsprognosen** Bei den beiden untersuchten Netztopologien handelt es sich um ein landesweites Backbonenetz mit 14 Standorten und 23 Glasfaserstrecken, vgl. Abbildung 5.8, sowie um ein Regionalnetz mit 40 Standorten und 82 Glasfaserstrecken, vgl. Abbildung 5.9. Diese beiden Netze können in ihrer Struktur durchaus als typisch für Deutschland angesehen werden.

Die Längen der Übertragungsstrecken im Backbonenetz liegen, mit einer Ausnahme, im Bereich zwischen 100 km und 430 km; folglich kann auf

allen Strecken gleichartige long-haul WDM-Technologie eingesetzt werden. Der durchschnittliche Knotengrad liegt bei 3,2 und der maximale bei 5. Die Standorte in diesem Netz sind ausschließlich Backboneknoten; sie repräsentieren jeweils Sammelpunkte für den überregionalen Verkehr in ihrer Umgebung. Somit ist eine logische Vollvermaschung zwischen diesen Standorten ein realistisches Szenario. Es existiert folglich eine Verkehrsbeziehung zwischen jedem Knotenpaar.

Der durchschnittliche Knotengrad im Regionalnetz beträgt 4,1. Zwei zentrale Knoten der Region treten durch vergleichsweise viele physikalische und auch logische Verbindungen hervor; diese bilden in der Originalnetzstruktur die Ablaufpunkte in die nächst höhere Hierarchieebene. Die Längen der einzelnen Streckenabschnitte im Regionalnetz liegen im Bereich von bis zu 15 km. Das Netz repräsentiert einen städtischen Ballungsraum, welcher etwa ein Oval von 50 km bzw. 30 km Radius umfasst. Hier kommt SDH-Technologie mit 2,5 Gbit/s bzw. 10 Gbit/s zum Einsatz, auf Strecken mit entsprechend hoher Last ggf. auch Metro-WDM-Technologie.

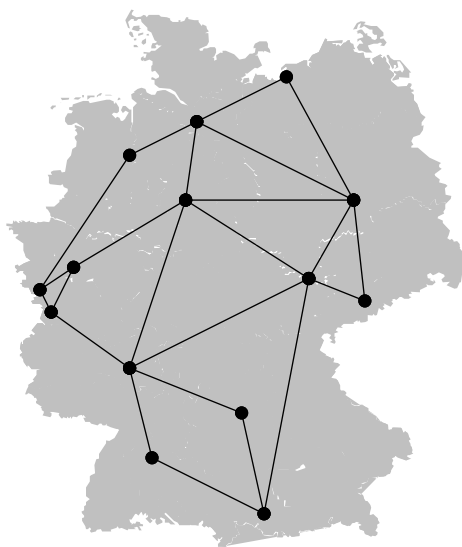


Abbildung 5.8: Deutschlandweiter Backbone.

Für beide Netzwerke steht eine aus realen Prognosedaten abgeleitete Bedarfsverteilung zur Verfügung. Regional liegen die Bedarfe hauptsächlich im Bereich unter 2,5 Gbit/s, nur vereinzelt auch darüber. Im Backbone bewegen sich die meisten Bedarfe im Bereich von 5 Gbit/s bis 15 Gbit/s. Für die Vergleichsrechnungen werden diese Bedarfsmatrizen variiert. Die Größe der Bedarfe wird dabei im Bereich von einem Viertel bis zum Vierfachen der Originalbedarfe betrachtet. Die Spreizung der Bedarfe simuliert eine Veränderung

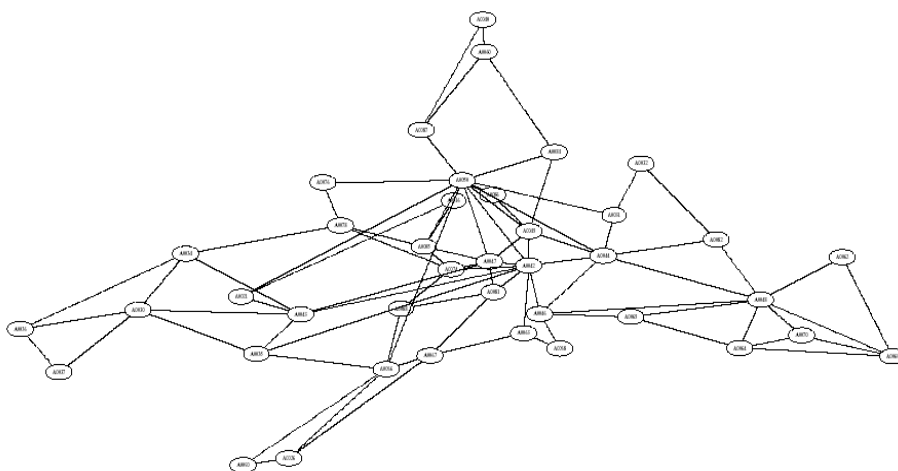


Abbildung 5.9: Regionalnetz.

der Größenverhältnisse der einzelnen Bedarfe zueinander. Bei einer hohen Spreizung ist die Differenz zwischen dem größten und dem kleinsten Bedarf sehr hoch, bei einer geringen Spreizung haben die Bedarfe alle eine ähnliche Größe.

Die Bedarfsmatrix für den Backbone ist logisch vollvermascht, jeder Knoten unterhält also mit jedem anderen eine Verkehrsbeziehung. Die Bedarfsmatrix für das Regionalnetz enthält lediglich 129 Verkehrsbeziehungen. Für die Berechnungen unter Variation der logischen Vermaschung wird deshalb eine künstlich erstellte Matrix genutzt, in der alle Bedarfe die gleiche Höhe haben. So kann die logische Vermaschung bis auf eine fast vollständige Vermaschung angehoben werden, ohne die Vergleichbarkeit einzuschränken.

**Numerische Ergebnisse** Die numerischen Ergebnisse sind in vier Tabellen und Diagrammen dargestellt. Alle Kostenangaben sind jeweils auf die Kosten der ungeschützten Netze normiert. Die Werte spiegeln deshalb die relative Veränderung der Kosten für ein mittels 1+1 Protection bzw. statischem Restoration geschütztes Netz im Vergleich zum ungeschützten Netz wider. Selbstverständlich ist ein Netz mit großen Bedarfen teurer als eines mit kleinen; diese Differenzen sind jedoch nicht Kern der hier vorliegenden Untersuchung und werden deshalb durch die Normierung aus den Ergebnissen ausgeblendet.

Abbildung 5.10 zeigt den Einfluss der Bedarfsgröße auf das Schutzverfahren. Die Bedarfsgröße „1“ entspricht den originalen Bedarfsdaten. Der Mehraufwand für Restoration (65%) bzw. für Protection (111%) liegt in dem Bereich, der auch in der Literatur häufig genannt wird, vgl. z.B. Ellinas

*et al.* (2003). Je größer die Bedarfe sind, desto höher sind die Mehrkosten für beide Schutzverfahren im Vergleich zu einem ungeschützten Netz. Hierin spiegelt sich unter anderem wider, dass ein ungeschütztes Netz noch Reservekapazitäten in Form von Koppelfeldern und Multiplexerkanälen bereit hält. Für einen Teil der Backupwege werden also nur Portkarten und Transponder benötigt. Je größer also die Bedarfe sind, desto geringer ist der Anteil, der noch von diesen Restkapazitäten profitieren kann. Folglich steigen die relativen Kosten bezogen auf das ungeschützte Netz langsam an. Der Unterschied zwischen Protection und Restoration ist hierbei nur marginal; das Kostenverhältnis verbessert sich zu Gunsten von Restoration von 1,22:1 auf 1,3:1. In diesem Szenario hat das Schutzverfahren also einen nahezu konstanten Einfluss.

Um den Einfluss der Verteilung der Bedarfsgrößen zu untersuchen, wird das Bedarfsspektrum mehr oder weniger aufgefächert (gespreizt), vgl. Abbildung 5.11. Der Spreizfaktor gibt an, um wie viele Male der größte Bedarf größer ist als der kleinste. Die Originalmatrix hat einen Spreizfaktor von 10. Eine Vergrößerung geht mit einer Erhöhung der Gesamtbedarfsmenge einher, um eine Verkleinerung der kleinsten Bedarfe auf Null zu vermeiden. Der (geringe) Kostenanstieg beider Schutzverfahren im Vergleich zum ungeschützten Netz ist im Wesentlichen auf diese Erhöhung der Gesamtbedarfsmenge zurückzuführen. Relativ zueinander verändern sich die Schutzverfahren erneut wenig. Selbst eine weite Auffächerung der Bedarfsgrößen führt nur zu einer marginalen Verschlechterung des Restorationverfahrens im Vergleich zum Protectionverfahren. Theoretisch könnte man erwarten, dass das Restorationverfahren bei stark unterschiedlichen Bedarfsgrößen deutlich schlechter abschneidet, da Backupressourcen immer nur im Umfang des kleineren Bedarfes doppelt genutzt werden können. Die verhältnismäßig hohe Zahl an Bedarfen in Verbindung mit der weitmaschigen Netzstruktur führt im vorliegenden Beispiel jedoch dazu, dass die jeweiligen Backupressourcen häufig von mehr als zwei Pfaden genutzt werden und die Größenunterschiede der Bedarfe nicht so stark ins Gewicht fallen.

Das Regionalnetz unterscheidet sich signifikant vom Backbonenetz. Der wichtigste Punkt ist die deutlich geringere logische Vermaschung in Verbindung mit einer höheren physikalischen Vermaschung. Entsprechend unterschiedlich fallen auch die Ergebnisse aus, vgl. Abbildung 5.12. Bei sehr kleinen Bedarfen sind beide Schutzverfahren nur wenig teurer als das ungeschützte Netz. Dieses auf den ersten Blick überraschende Ergebnis kann mit dem sehr hohen Fixkostenblock erklärt werden, der das Ergebnis dominiert. Die Größe der Cross-Connects ist auf die Originalbedarfsverteilung abgestimmt und insbesondere für die Berechnung mit den kleinsten Bedarfen nicht sinnvoll dimensioniert. Entgegen den Beobachtungen im Backbonenetz steigen

die Kosten mit größer werdenden Bedarfen nicht kontinuierlich an, sondern fallen gegen Ende sogar wieder leicht ab. Dies geht einher mit dem verstärkten Einsatz von WDM-Systemen, die auf diesen kurzen Strecken erst bei großen Bedarfen wirtschaftlicher sind als diskrete SDH-Systeme. Im Vergleich zum ungeschützten Netz schneidet 1+1 Protection lediglich etwas schlechter ab als im Backbonenetz, während Restoration im Regionalnetz deutlich weniger Vorteile einbringt als im Backbone. Die im Verhältnis zu den zur Verfügung stehenden Faserstrecken deutlich geringere Zahl an Bedarfen führt dazu, dass Backupkapazitäten nur selten sinnvoll von zwei oder mehr Bedarfen geteilt werden können.

Der Einfluss der logischen Vermaschung im Regionalnetz ist in Abbildung 5.13 dargestellt. Hier wird deutlich, dass das Restorationverfahren unter 200 Verkehrsbeziehungen nur einen geringen Vorteil gegenüber 1+1 Protection hat, während die Ergebnisse für eine sehr hohe logische Vermaschung fast an die Resultate für das Backbonenetz bei hoher Last heranreichen. Wenn nur relativ wenige Bedarfe in der Region vorhanden sind, führt dies zusammen mit der recht engen Vermaschung dazu, dass nur wenig Backupkapazitäten mehrfach genutzt werden können. Zusätzlich verringert die höhere physikalische Vermaschung der Region den Vorteil für den Restorationansatz.

Bei allen Berechnungen wurde die Verfügbarkeit einer ausreichenden Zahl von Glasfasern auf den vorgegebenen Trassen als gegeben angesehen. Wenn Kosten für den Aufbau dieser Strecken entstehen, kann im Regionalnetz ein zusätzlicher Vorteil von Restoration zum Tragen kommen. Die Berechnungen haben gezeigt, dass bei Restoration mit der Originalbedarfsmatrix je nach Bedarfsgröße nur etwa 74 bis 77 Kanten der 82 zur Verfügung stehenden genutzt werden. Erst bei einer sehr hohen Zahl an Bedarfen nutzt auch das Restorationrouting alle verfügbaren Kanten. Im physikalisch weitmaschigeren Backbonenetz mit logischer Vollvermaschung werden immer alle verfügbaren Trassen genutzt; hier entsteht also kein Vorteil.

**Zusammenfassung** Die Fallstudie betrachtet die notwendigen Geräteinvestitionen für zwei typische SDH/WDM-Netzwerke speziell unter der Berücksichtigung verschiedener Bedarfsstrukturen. Dabei hat sich gezeigt, dass die Einsparungsvorteile, die sich durch den Einsatz von statischem Restoration ergeben können, bei logisch vollvermaschten und physikalisch weitmaschigen Netzen relativ unabhängig von der Bedarfsstruktur sind. Weder eine starke Veränderung der durchschnittlichen Größe der Bedarfe, noch eine stark asymmetrische Verteilung der Bedarfsgrößen hat bei dem untersuchten Backbonenetz einen wesentlichen Einfluss auf die Vorteilhaftigkeit von Restoration. Die Entscheidung für oder gegen

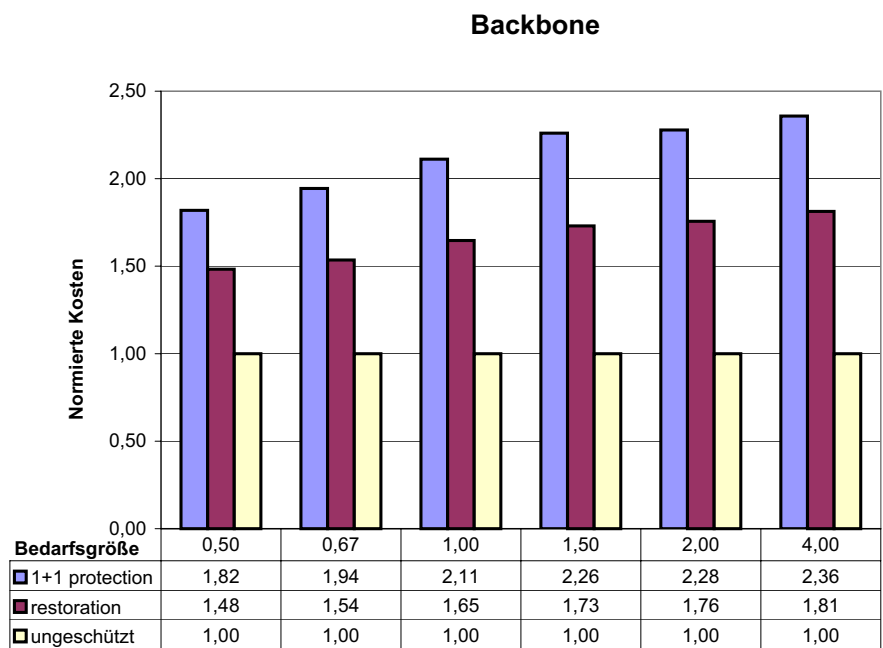


Abbildung 5.10: Einfluss der Bedarfsgröße im Backbone.

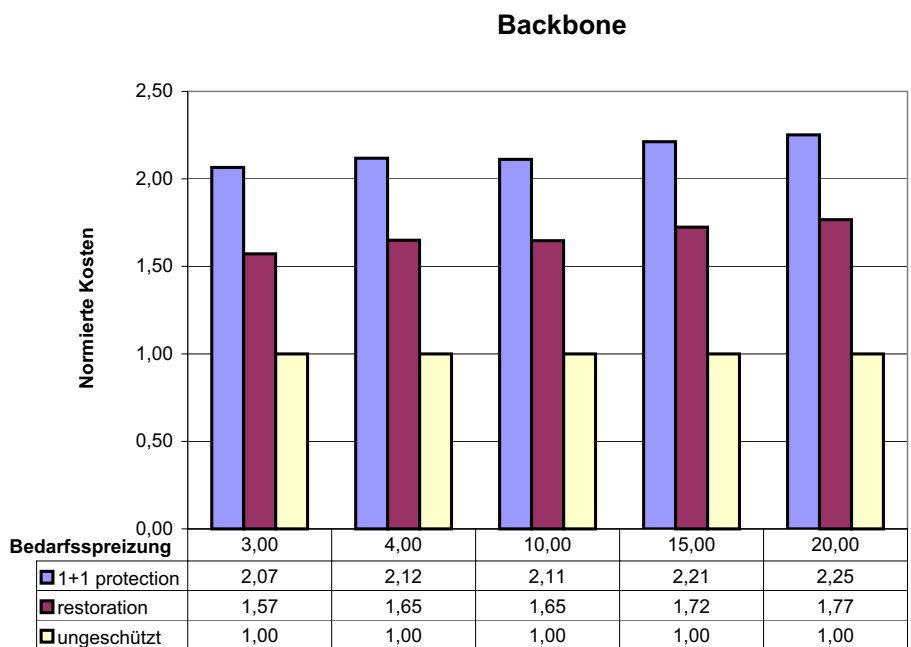


Abbildung 5.11: Einfluss der Bedarfsspreizung im Backbone.



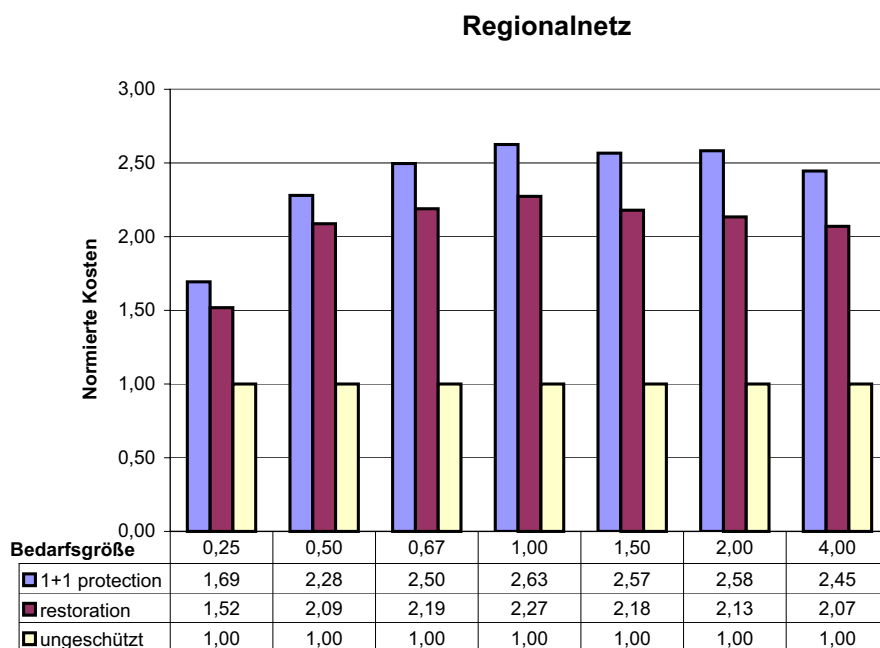


Abbildung 5.12: Einfluss der Bedarfsgröße in der Region.

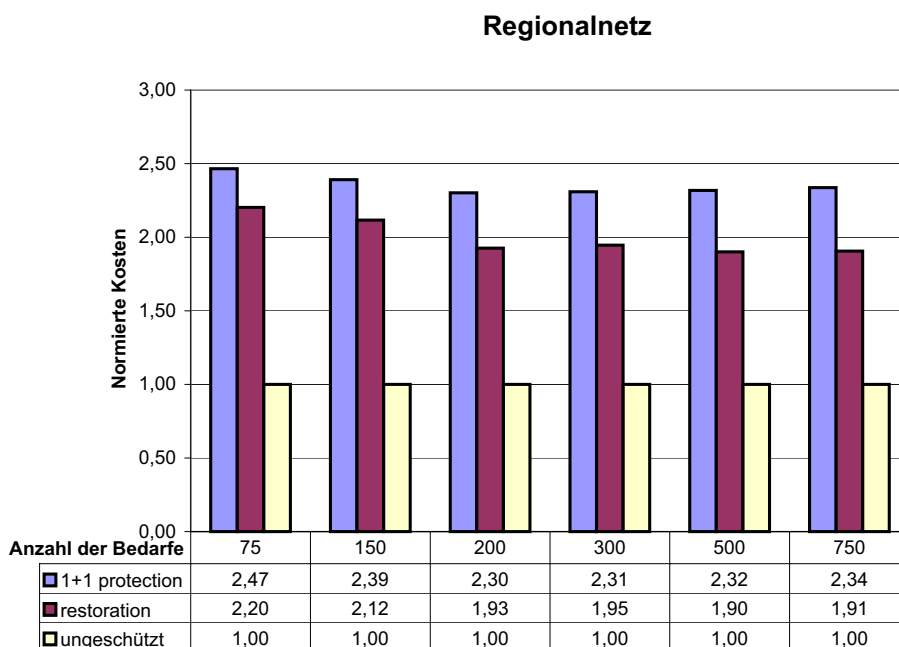


Abbildung 5.13: Einfluss der logischen Vermaschung in der Region.

den Einsatz eines solchen Restorationverfahrens kann also weitestgehend unabhängig von den Unsicherheiten der Bedarfsprognose erfolgen.

Etwas differenzierter muss die Vorteilhaftigkeit bei physikalisch eng vermaschten Netzen und vor allem bei einer Veränderung der logischen Vermaschung gesehen werden. Auch hier ist Restoration natürlich in keinem Fall von den Investitionskosten her teurer als 1+1 Protection; bei relativ wenigen Bedarfen schmilzt der Vorteil jedoch deutlich zusammen. Gleiches gilt, wenn die Infrastruktur von Anfang an große Kapazitätsreserven bereitstellt. Wird bei der Bedarfsprognose also von einer großen Schwankungsbreite in bezug auf Zahl und Größe der Bedarfe ausgegangen, so kann nur eine Betrachtung des Einzelfalles Klarheit darüber schaffen, ob das Restorationverfahren wirklich die erhofften Einsparungen bringt oder ob die ggf. verbleibenden 10% Investitionersparnis nicht an anderer Stelle, z.B. bei den Betriebskosten, wieder verloren gehen.

# Kapitel 6

## Fazit

In dieser Arbeit werden verschiedene Ansätze zur Lösung schichtenübergreifender Netzplanungsprobleme für schnelle Glasfasernetze vorgestellt. Der technologische Fokus liegt dabei auf WDM, SDH und MPLS, jedoch sind die Verfahren auch auf andere Technologien wie beispielsweise OTN übertragbar. Aufbauend auf einer möglichst realistischen Modellierung tatsächlicher Problemstellungen werden zwei Lösungsverfahren implementiert und getestet, ein exaktes Verfahren auf Basis von AMPL/CPLEX und ein heuristisches Verfahren auf Basis der GRASP-Ideen. Ein weiterer Schwerpunkt ist die Berücksichtigung unterschiedlicher Schutzverfahren.

Die wesentlichen Fortschritte, die mit dieser Arbeit erzielt wurden, liegen auf drei Gebieten: weitgehend simultane Planung über mehrere Netzschichten hinweg, eine sehr detaillierte Modellierung der Problemstellung bis hin zu einzelnen Portkarten und dies beides zusammen auch für sehr große Problemstellungen. Der theoretische Erkenntnisgewinn, der aus diesen Modellen und den damit bearbeiteten synthetischen Testinstanzen gewonnen wurde, konnte erfolgreich auf Probleme aus der Praxis übertragen werden. Insbesondere die Heuristik wurde erfolgreich auf verschiedene, teilweise sehr große, Praxisprobleme aus dem Bereich der taktischen und strategischen Netzplanung angewandt. Das exakte Lösen kleiner Probleminstanzen mit Hilfe von CPLEX dient dagegen primär der Kontrolle der Leistungsfähigkeit der heuristischen Verfahren.

Methodisch ist vor allem die gezielte Anpassung der GRASP-Idee auf das vorliegende Planungsproblem sowie die Hybridisierung mit der Pilotmethode hervorzuheben. Darüber hinaus wurden auch wertvolle Erkenntnisse gewonnen, welche mathematischen Formulierungen sich besonders gut eignen, um Flussprobleme mit Ganzzahligkeitsbedingungen effizient mit CPLEX lösen zu können. Sowohl die Heuristik, als auch das mit CPLEX gelöste ganzzah-

lige Modell haben den direkten Vergleich mit konkurrierenden Modellformulierungen oder Heuristiken in verschiedenen Vergleichsrechnungen sehr gut bestanden.

Alle vorgestellten Verfahren und Softwaretools stellen Hilfsmittel für einen erfahrenen Planer dar. Sie ermöglichen es, in kurzer Zeit unterschiedliche Szenarien durchzuspielen und gute Lösungsvorschläge für komplexe Problemstellungen zu generieren. Somit bieten sie eine wertvolle Hilfe für die taktische und strategische Netzplanung. Es handelt sich dabei jedoch keineswegs um eine vollautomatische Netzplanung. Der gesamte Planungsprozess mit seinen vielfältigen Randbedingungen ist viel zu komplex, als dass dies heute oder in naher Zukunft sinnvoll machbar wäre. Darüber hinaus stößt man im praktischen Einsatz immer wieder auf zusätzliche Probleme, die sich ohnehin kaum automatisieren lassen, wie z.B. die Interpretation einer unklaren Datenlage oder Kompetenzverteilungen zwischen einzelnen Abteilungen, letzteres insbesondere bei der Planung über Technologiegrenzen hinweg.

Gerade im Sinne einer noch besseren Entscheidungsunterstützung für den Netzplaner bieten dieses Forschungsgebiet im Allgemeinen und die in dieser Arbeit vorgestellten Ansätze im Speziellen vielfältige Anregungen für weiterführende Forschung. Dabei erscheint es alleine schon aufgrund der Ungenauigkeit der Eingangsdaten wenig sinnvoll, die letzten 0,5% zu suchen, die noch zu einem theoretisch optimalen Ergebnis fehlen, oder grundlegend andere Heuristiken anzuwenden. Der Vergleich mit der VNS-Heuristik in Melian *et al.* (2005b) und Höller *et al.* (2006) zeigt deutlich, dass man aus ganz unterschiedlichen Richtungen letztendlich zu sehr ähnlichen Ergebnissen gelangen kann.

Vielmehr bietet es sich an, den Aspekt der Entscheidungsunterstützung noch stärker zu betonen, beispielsweise in dem nicht eine einzige Lösung generiert wird, sondern ein Pool guter Lösungen, die sich jedoch strukturell möglichst weitgehend voneinander unterscheiden. Erste Ideen dazu sind bereits in Abschnitt 4.1.3 dargestellt. Dies ist aus Sicht der praktischen Anwendung ein Anfang, den es sich lohnt gerade auch auf Basis der in dieser Arbeit gewonnenen Erkenntnisse gezielt weiter zu verfolgen.

# Anhang A

## Programmdokumentation

Für die praktische Nutzung der in Abschnitt 4.2 beschriebenen Heuristik wurde im Laufe der Zeit ein Programm entwickelt, welches viele der Funktionalitäten über definierte Schnittstellen nutzbar macht. Die Bedienung dieses Programms und die entsprechenden Schnittstellen werden in den folgenden Abschnitten beschrieben. Es handelt sich dabei jedoch primär um ein Testwerkzeug und nicht um ein Programm für Endanwender. Dieser Prototypencharakter spiegelt sich auch noch in der aktuellen Versionsbezeichnung 0.9f wider.

### A.1 Programmstart und Systemanforderungen

Die Software wurde ausschließlich auf Windows XP-Computern entwickelt und getestet. Eine Kompatibilität mit anderen Betriebssystemen kann nicht garantiert werden. Eine erfolgreiche Nutzung unter Windows 98 / 2000 / NT 4.0 kann aber unter bestimmten Bedingungen möglich sein. Die minimalen Anforderungen an den zur Verfügung stehenden Arbeitsspeicher zum Starten des Programms betragen 128 MB. Für eine erfolgreiche Berechnung großer Netze wird entsprechend zusätzlicher Speicher benötigt; für eine Netzgröße von 1000 Knoten sind 512 MB jedoch in jedem Fall ausreichend. Zum Start des Programms muss die Datei `Regionalplanung.exe` ausgeführt werden. Die Datei `default.xls` muss sich in demselben Verzeichnis wie die Programmdatei befinden. Andere Komponenten sind zum Starten nicht erforderlich.

## A.2 Die Bedienelemente

Nach dem Start des Programms erscheint das in Abbildung A.1 gezeigte Fenster. Auf der linken Seite befinden sich von oben nach unten gesehen die einzelnen Schritte, die vom Nutzer ausgeführt werden können bzw. müssen. Zwingend erforderliche Planungsschritte sind die linksbündigen Schaltflächen. Die eingerückten Schaltflächen führen optionale Zwischenschritte aus. Auf der rechten Seite befinden sich allgemeingültige Optionen und die Statusanzeigen.

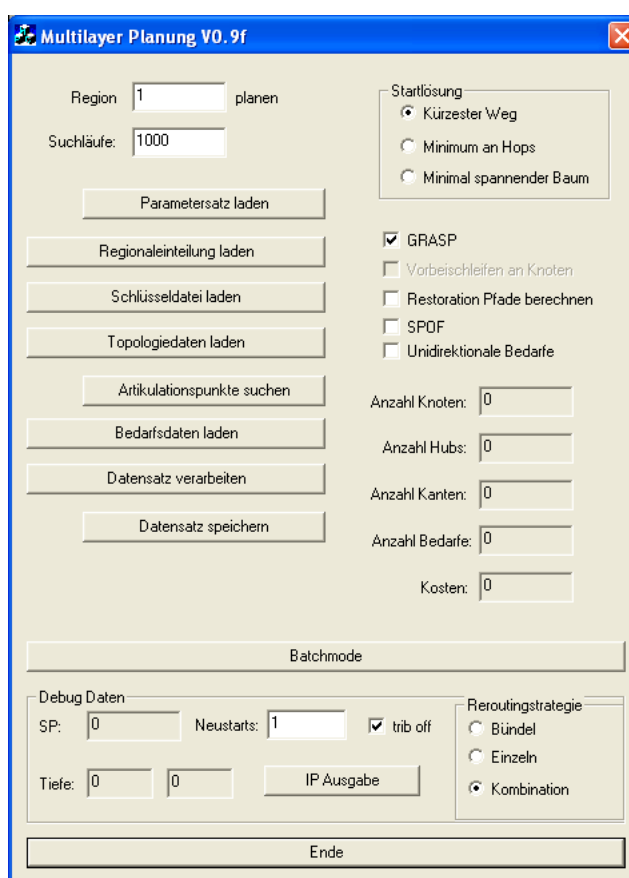


Abbildung A.1: Bedienoberfläche.

Die Funktion der Schaltflächen wird im Folgendem am Beispiel eines vollständigen Planungsdurchlaufes beschrieben. Zuerst muss die zu planende Region ausgewählt werden. Bei nicht regional strukturierten Netzen wird dies normalerweise die Nummer „1“ sein. Details dazu finden sich im Abschnitt A.3. Die Anzahl der Suchläufe gibt an, wie oft die äußere Schleife

des Planungsalgorithmus durchlaufen wird, also wie viele Neustarts des Re-routings durchgeführt werden sollen. Bei einem Wert von „0“ wird nur die Startlösung berechnet. Falls ein anderer als der Standardparametersatz in der Datei `default.xls`, die beim Programmstart automatisch geladen wird, verwendet werden soll, kann dieser im ersten Schritt über die Schaltfläche „Parametersatz laden“ eingelesen werden. Das Format ist in Abschnitt A.3 beschrieben. Bei jeder Schaltfläche, die Dateien einliest oder ausgibt, erscheint der entsprechende normale Windowsdialog.

Der erste verpflichtende Planungsschritt ist das Einlesen der Regionaleinteilung. Hierdurch werden aus der Gesamtmenge der Knoten diejenigen ausgewählt, die sich innerhalb der vorher eingestellten Region befinden. Nach dem Laden der Regionaleinteilung erscheinen rechts im Programmfenster die Werte für die Anzahl der Knoten und die Anzahl der Hubs in der ausgewählten Region. Durch das Laden der Schlüsseldatei (vgl. Abschnitt A.3) werden die geographischen Koordinaten und die Schlüsselziffern für die Knoten eingelesen. Als Nächstes müssen die Topologiedaten eingelesen werden, also die Glasfaserstrecken zwischen den Knoten. Sollte die Topologiedatei Kanten enthalten, deren Start- oder Endknoten nicht innerhalb der bekannten Knotenmenge liegt, so werden diese ignoriert. Die Gesamtzahl der Kanten des Netzes wird rechts im Statusfeld angezeigt. Wenn Zweifel bestehen, ob der Glasfasergraph vollständig (zweifach) zusammenhängend ist, kann eine Suche nach Artikulationspunkten durchgeführt werden. Ein Statusfenster informiert dann über ggf. vorhandene, einbeinig angebundene oder gar nicht angebundene Knoten. Der Text kann innerhalb des Ausgabefensters mit den Cursortasten gescrollt werden.

Der nächste Schritt ist das Laden der Bedarfsdaten. Zum Format der entsprechenden Datei vgl. Abschnitt A.3. Hier ist zu beachten, dass Bedarfe bei denen Start- und Zielknoten außerhalb der gewählten Region liegen, ignoriert werden, während Bedarfe, bei denen nur einer der Knoten außerhalb der Region liegt, auf die zugewiesenen Hubknoten umgelegt werden. Bei der Planung eines Backboneetzes werden dagegen nur die Bedarfe geführt, deren Start- und Zielknoten in unterschiedlichen Regionen liegen, indem sie im Backbone zwischen den beiden Hubknoten geführt werden. Wiederum wird die Gesamtzahl der Bedarfe rechts im Statusfeld angezeigt.

Die Schaltfläche „Datensatz verarbeiten“ startet die eigentliche Netzplanung. Bei mehreren Suchläufen wird der entsprechende Wert im Eingabefeld sukzessive heruntergezählt, und der aktuell beste gefundene Wert für die Gesamtkosten erscheint rechts im Statusfenster. So kann der Fortschritt der Berechnung bei längeren Rechenzeiten jederzeit verfolgt werden. Wenn die Berechnung abgeschlossen ist, wird darauf mittels eines Nachrichtenfensters hingewiesen. Nun können die detaillierten Ergebnisse als Text-, Excel- und

Grafikdatei („Datensatz speichern“) gespeichert werden.

Die Schaltfläche „Batchmode“ führt alle manuellen Schaltflächen und die dahinterstehenden Funktionen – bis auf die Suche nach Artikulationspunkten – automatisch aus. Dazu müssen die Dateinamen und -pfade in einer Textdatei gespeichert sein. Das genaue Format dieser Datei ist in Abschnitt A.3 beschrieben.

Auf der rechten Seite des Programmfensters befinden sich oberhalb der bereits beschriebenen Statusanzeigen eine Reihe von Optionen, die vor dem Start der Berechnung ausgewählt werden können. Die Routingstrategie für die Startlösung kann über das entsprechende Optionsfeld gewählt werden. Für die Berechnung der Startlösung kann jeder Bedarf auf dem geographisch kürzesten Weg geroutet werden. Hier wird die Summe der Kantenlängen als Metrik für den Dijkstra-Algorithmus genutzt. Alternativ können die Längen der Kanten auch ignoriert werden und die Bedarfe werden jeweils auf dem Weg mit den wenigsten Hops geführt. Wege unterschiedlicher Länge, welche die gleiche Anzahl an Hops aufweisen, werden dabei als gleichwertig betrachtet. In der Praxis unterscheidet sich diese Startlösung nur wenig vom geographisch kürzesten Weg, da die Dreiecksungleichung für die Glasfaserstreckenführung in realen Netzen überwiegend erfüllt sein wird. Die dritte Variante nutzt einen Minimal-spannenden-Baum als Ausgangspunkt. Obwohl in den meisten Fällen der kürzeste Weg die besten Resultate erzielt, eignet sich der Minimal-spannende-Baum für bestimmte Sonderfälle, insbesondere wenn eine optimale Lösung ein sehr dünnes Netz mit nur wenigen Kanten darstellt. Dies kann z.B. auftreten, wenn die Basiskosten für die Einrichtung einer Kante sehr hoch sind und somit auch längere Umwege bei der Bedarfsführung lohnen.

Die Kontrollkästchen rechts in der Mitte ermöglichen die Auswahl verschiedener Optionen für das iterative Optimierungsverfahren. „GRASP“ aktiviert die gleichnamige Variante des lokalen Suchverfahrens für das Rerouting. Mit Hilfe dieses Verfahrens können gute Lösungen oft in einer kürzeren Zeit gefunden werden als mit dem Standardverfahren. Wenn die Zahl der Suchläufe sehr groß gewählt werden kann ergibt sich normalerweise kein Vorteil mehr. „Vorbeischleifen an Knoten“ ermöglicht in gewissem Umfang das direkte Führen voll belegter Glasfasern an Knoten vorbei, ohne dass Koppel-feldkapazitäten belegt werden. Diese Funktion ist noch experimental. „Restoration Pfade berechnen“ wählt ein vorberechnetes, statisches Restorationverfahren als Schutz für alle Bedarfe. Die Ersatzwege werden genau wie bei 1+1 Protection im Voraus bestimmt und entsprechend in den Ausgabedateien aufgeführt. Allerdings können die Ressourcen der Ersatzwege von mehreren Erstwegen genutzt werden, die keine gemeinsamen Fehler aufweisen können. „SPOF“ (Single Point of Failure) führt geschützten, überregionalen Verkehr



nur über den ersten Hubknoten in den Backbone, die disjunkte Wegführung innerhalb der Regionalnetze von der Quelle bis zum Hubknoten bleibt davon unberührt. „Unidirektionale Bedarfe“ ermöglicht für ein Knotenpaar  $A - B$  die Größe des Bedarfes von  $A$  nach  $B$  und von  $B$  nach  $A$  jeweils unterschiedlich zu wählen (z.B. bei-IP Verkehr). Solche asymmetrischen Bedarfe belegen alle Ressourcen richtungsgebunden.

Der „Debug Daten“ genannte Kasten im unteren Bereich des Programmfensters enthält neben einigen programminternen Statusinformationen auch die Bedienelemente für neue Funktionen, die noch nicht vollständig integriert sind. In der Version 0.9f ist hier für den Anwender nur der rechte Bereich mit den Routingstrategien interessant. Dort lässt sich auswählen, ob das Verbesserungsverfahren jeweils Bedarfsbündel, Einzelbedarfe oder eine sequentielle Kombination beider Verfahren verwenden soll. Diese Auswahl ist hauptsächlich für Vergleiche der Verfahren nützlich, für praktische Netzrechnungen liefert das Kombinationsverfahren fast immer die beste Lösung. Die Option „trib off“ ermöglicht es, die Anschaltung der Kundenbedarfe (tributary demands) an die Cross-Connects aus der Kapazitäts- und Kostenbetrachtung auszublenden.

### A.3 Eingabedaten

Die Eingabedaten werden mit Ausnahme der Regionaleinteilung alle aus Microsoft-Excel-Dateien eingelesen. Ob diese Daten alle in einer Datei stehen, oder ob diese auf verschiedene Dateien verteilt sind, ist unerheblich. Entscheidend ist, dass die Datenfelder alle die vorgegebenen Bezeichner im „Namenfeld“ besitzen. Sollten die Daten nicht korrekt abgelegt oder bezeichnet sein, zeigt das Datenbankmodul eine entsprechende Fehlermeldung an. Die Daten können von Hand über die entsprechenden Schaltflächen oder mit Hilfe einer Batchdatei eingelesen werden.

#### Batchdatei

Die Batchdatei ist eine ASCII-Datei (American Standard Code for Information Interchange), welche die vollständigen Pfadangaben aller Dateien enthält. Die Ausgabedatei benötigt keinen Suffix, die entsprechenden Endungen `.txt`, `.xls` und `.dot` werden automatisch hinzugefügt. Die Reihenfolge ist:

- Programmparameter
- Regionaleinteilung

- Knotendaten
- Glasfasergraph
- Bedarfsmatrix
- Ausgabedatei

Ein Beispiel für eine Batchdatei könnte wie folgt aussehen:

```
C:\Testnetz\parameter_s1.xls
C:\Testnetz\region_4.txt
C:\Testnetz\knoten.xls
C:\Testnetz\topologie.xls
C:\Testnetz\bedarfe.xls
C:\Testnetz\ausgabe
```

## Programmparameter

Die Programmparameter werden in der Ini-Datei festgelegt, die beim Programmstart automatisch geladen wird. Alternativ dazu können sie auch durch Nachladen einer separaten Datei individuell angepasst werden. Die Parameterdatei ist eine Exceldatei, welche sich in zwei Bereiche gliedert. Kostenwerte und Faktoren der Kostenfunktionen haben Fließkommaformat, alle anderen Werte müssen ganzzahlig angegeben werden.

Das Blatt „Polynom“ enthält die Kostenfunktionen, die vom Reroutingalgorithmus für Bündelung genutzt werden. Es können zwei verschiedene Polynome definiert werden, eines für die Kantenkosten in Abhängigkeit von der Kantenlänge  $x$  ( $ax^m + bx^n + c$ ) und eines für die Knotenkosten inklusive Portkarten in Abhängigkeit von der Verkehrslast  $y$  des jeweiligen Knotens ( $dy^o + ey^p + f$ ). Die Basiskostenwerte definieren interne „Strafkosten“, die für die erstmalige Benutzung leerer Kanten oder Knoten angesetzt werden. Diese Werte können im Normalfall unverändert bleiben.

Das Blatt „Gerätedaten“ enthält die Kapazitäts- und Kostenangaben für alle zulässigen Hardwarekomponenten (Cross-Connects, Portkarten, WDM-Multiplexer, Transponder und Verstärker). Die Datei `default.xls` enthält eine kommentierte Beispielkonfiguration für ein SDH-Netz mit 2,5 Gbit/s und 10 Gbit/s Portkarten. Im Prinzip sind die Größen und Reichweiten der Karten beliebig, die große Karte sollte jedoch ein ganzes Vielfaches der kleinen an

Kapazität besitzen. Die Kapazität der Multiplexer ist als Anzahl der installierbaren kleinen Karten angegeben. Es wird vorausgesetzt, dass eine von der Kapazität her äquivalente Anzahl an großen Karten installiert werden kann. Durch die Verknüpfung über die Mux-Kapazitäten sollten die kleinen Karten kurzer und langer Reichweite die gleiche Kapazität besitzen, bei den großen ist dies nicht notwendig. Z.B. wäre eine Kartenauswahl 2,5 Gbit/s short haul (sh), 2,5 Gbit/s long haul (lh), 10 Gbit/s sh und 40 Gbit/s lh zulässig. Wenn große Karten pro Bit teurer sind als kleine, werden sie nie eingesetzt. Short haul-Karten werden nur für die Verbindung zu WDM-Multiplexern eingesetzt, long haul-Karten nur auf direkten SDH-Kanten. Wenn die kleinen Karten nicht die Kapazität 1 besitzen, dann führt das Programm automatisch ein Grooming durch. Mehrere kleine Bedarfe, welche die gleiche Kante benutzen, werden dann wenn möglich auf einer Karte zusammengefasst. Es wird davon ausgegangen, dass die Cross-Connects in der Granularität schalten können, in der die Bedarfe angegeben werden. Soll dies verhindert werden, müssen die Bedarfe vorher in einem externen Vorverarbeitungsschritt auf die kleinste Kartenkapazität aufgerundet werden.

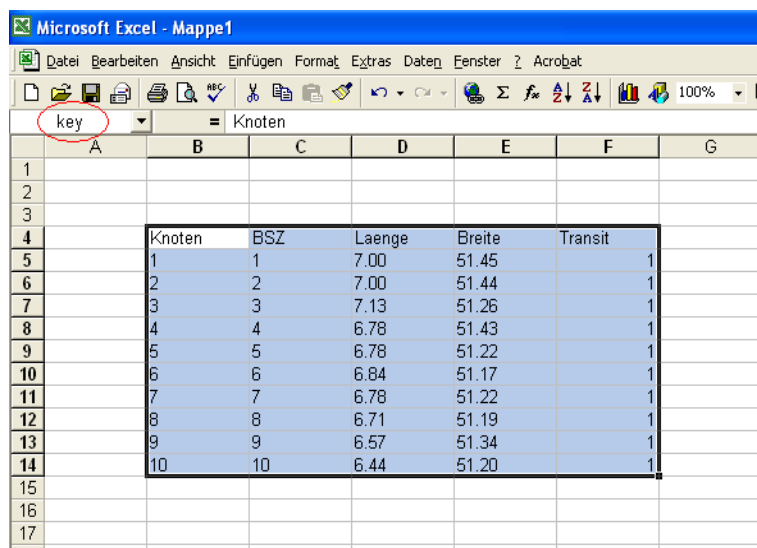
## Regionaleinteilung

Die Regionaleinteilung wird aus Kompatibilitätsgründen zu alten Programmmodulen (Vorverarbeitung, Regionenbildung) als ASCII-Datei eingelesen. Das Datenformat ist unverändert. Am Anfang der Datei stehen diverse Parameter für die weitere Verarbeitung und zusätzliche Informationen für den Nutzer. Die @-Zeichen dienen als Suchmarken für die Einleseroutine und dürfen keinesfalls gelöscht oder verschoben werden. Im Abschnitt „Hubknoten“ steht jede Zeile für eine Region; die aufgelisteten Knoten bilden die Backboneknoten (je 12 Zeichen Feldgröße). Diese gehören von links nach rechts abwechselnd zu den beiden Backbone-Netzen. Im Abschnitt „Knotenzuordnung“ sind alle Knoten des Netzes mit den jeweils zugeordneten Backboneknoten aufgelistet (je zehn Zeichen Feldgröße). An Hand der Backboneknoten lässt sich ermitteln, zu welcher Region der jeweilige Knoten gehört. Für die Planung nicht regional strukturierter Netze kann prinzipiell eine beliebige Regionaldatei genommen werden und alle Knoten bekommen das gleiche Paar Hubknoten zugewiesen. Lediglich die Anzahl der Knoten ist entsprechend zu modifizieren und einzutragen.

## Knotendaten

Die Knotendaten umfassen eine fortlaufende Nummer, die bei 1 beginnen muss, eine frei wählbare Identifikationsnummer (BSZ) und den geographischen Längen- und Breitengrad des Standortes. Zusätzlich können die einzelnen Knoten für Transitverkehr freigegeben (1) oder gesperrt werden (0). Die Knotendatei kann prinzipiell immer alle vorhandenen Knoten enthalten, auch wenn diese nicht für die aktuelle Planung genutzt werden sollen. Die Selektion, welche Knoten bei der Planung berücksichtigt werden sollen, wird über die Regionaleinteilung gesteuert.

Abbildung A.2 zeigt ein Beispiel einer Knotenliste. Wichtig ist, dass die komplette Tabelle inklusive der Kopfzeile das Namensfeld „key“ trägt (in Abbildung A.2 rot eingekreist). Anhand dieses Bezeichners identifiziert die Software den gültigen Datenbereich für die Knotentabelle innerhalb des Excel-Dokumentes. Die Spaltenüberschriften dienen zur Identifikation der einzelnen Datensätze und dürfen somit ebenfalls nicht verändert werden. Diese Grundprinzipien gelten für alle Exceltabellen.



key		= Knoten					
	A	B	C	D	E	F	G
1							
2							
3							
4		Knoten	BSZ	Laenge	Breite	Transit	
5		1	1	7.00	51.45	1	
6		2	2	7.00	51.44	1	
7		3	3	7.13	51.26	1	
8		4	4	6.78	51.43	1	
9		5	5	6.78	51.22	1	
10		6	6	6.84	51.17	1	
11		7	7	6.78	51.22	1	
12		8	8	6.71	51.19	1	
13		9	9	6.57	51.34	1	
14		10	10	6.44	51.20	1	
15							
16							
17							
18							

Abbildung A.2: Knotendaten mit Betriebsstellenschlüssel.

## Glasfasergraph

Die Topologie des Glasfasergraphen wird aus einer Tabelle mit dem Namensfeld „Kanten“ eingelesen, vgl. Abbildung A.3. Jede existierende direkte Glas-

faserverbindung zwischen zwei Knoten wird durch eine Zeile repräsentiert, die den Startknoten, den Endknoten und die Länge der Faserstrecke in Kilometern enthält. Die Knotennummern sind die fortlaufenden Nummern aus der Knotendatei, nicht die jeweilige BSZ.

Knoten_1	Knoten_2	Kantenlaenge
36	39	6
5	8	7
97	98	14
88	89	15
118	119	18
16	18	20
8	10	20

Abbildung A.3: Glasfasertopologie.

## Bedarfsmatrix

Die Bedarfsmatrix (vgl. Abbildung A.4) enthält zu jedem Bedarf den Start- und Endknoten in der fortlaufenden Nummerierung, die Größe des Bedarfes in Basiseinheiten (vgl. die Definition der Basiseinheiten in Abschnitt A.3) und den gewünschten Schutz. Als Basiseinheit wird die kleinste Dateneinheit bezeichnet, die im Netz vollständig separat geführt werden kann bzw. soll. Bei SDH-Netzen könnte dies z.B. ein VC4 sein, bei IP-Netzen z.B. ein Kilobit. Alle Kapazitätsangaben müssen in ganzzahligen Vielfachen dieser Basiseinheit gemacht werden. Eine „1“ kennzeichnet einen ungeschützten Bedarf, eine „2“ einen 1+1 geschützten Bedarf. Geschützte und ungeschützte Bedarfe können beliebig gemischt werden. Ein Sonderfall ist die Nutzung der Planung mit Restoration als Schutzverfahren. Hier müssen die Bedarfe als ungeschützt definiert sein und eine Mischung verschiedener Schutzmechanismen ist nicht möglich. Das Namensfeld der Tabelle muss „Bedarfe“ lauten, die Spaltenüberschriften dürfen auch hier nicht modifiziert werden.

	A	B	C	D	E	F
1						
2						
3		Knoten_1	Knoten_2	Anzahl	Schutz	
4		121	20	11	2	
5		121	59	19	2	
6		121	29	43	2	
7		121	76	9	2	
8		121	15	7	2	
9		121	18	8	2	
10		121	34	2	2	
11		121	78	8	2	
12		121	133	6	2	
13		121	44	3	2	
14		121	49	6	2	
15		121	43	4	2	
16		121	59	3	2	
17		121	5	9	2	
18						
19						

Abbildung A.4: Bedarfsmatrix.

## A.4 Ausgabedaten

Die Ausgabedaten sind, unabhängig vom Format, mit erklärenden Kommentaren versehen. Dabei werden die kleinen Portkarten momentan noch grundsätzlich als 2,5 Gbit/s SDH und die großen als 10 Gbit/s SDH bezeichnet, unabhängig von den tatsächlich in den Parametern angegebenen Kapazitätswerten.

### Textausgabe

Die bei der Textausgabe erzeugte ASCII-Datei enthält alle Daten, die vom Programm zur Verfügung gestellt werden. Die Ausgabe ist mit entsprechenden Kommentaren versehen.

### Excelausgabe

Die Excelausgabe enthält die kompletten Gerätetabellen, das Routing ist hier im Gegensatz zur Textausgabe nicht enthalten. Die Stückzahlen für die

Kanten enthalten die Geräte für beide Enden. Zwei Portkarten entsprechen also einer einzelnen Verbindung.

## Grafikausgabe

Die Grafikausgabe erfolgt über eine ASCII-Datei, die von der frei verfügbaren Software Graphviz<sup>1</sup> als Vektorgrafik dargestellt werden kann. Normale Netzknoten werden durch Ellipsen dargestellt, Hubknoten werden durch Rechtecke repräsentiert. Die BSZ ist jeweils in den Knoten eingetragen. Je nach Größe des Gesamtnetzes können diese Knotennummern unter Umständen nur im Zoom lesbar sein. Reguläre, genutzte Glasfaserstrecken werden durch schwarze Linien gekennzeichnet. Unbenutzte Faserstrecken sind gelb, vom Programm eigenständig eingefügte Glasfasern (bei fehlendem Zusammenhang des Ursprungsnetzes) rot dargestellt. Die Projektion der Längen- und Breitengrade auf die Ebene erfolgt nur näherungsweise, das Bild erscheint also leicht verzerrt. Abbildung A.5 zeigt ein Beispiel eines solchen Netzbildes.

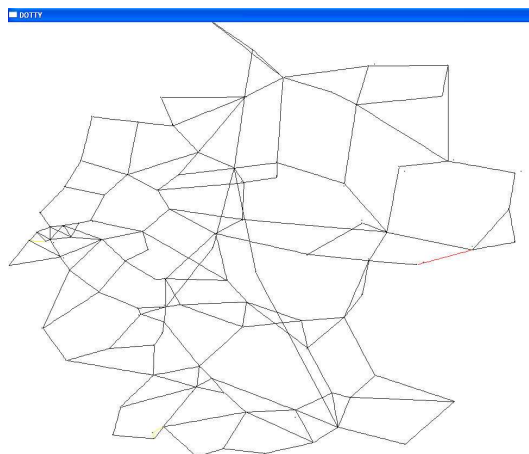


Abbildung A.5: Graphische Ausgabe.

---

<sup>1</sup>(<http://www.research.att.com/~north/graphviz/>)

# Anhang B

## Beispiel für ein AMPL-Modell

In AMPL findet eine Trennung zwischen dem eigentlichen Modell und den Eingangsdaten statt. Das entsprechende AMPL-Modell für das in Abschnitt 4.1.1 vorgestellte ganzzahlige Basismodell sieht dann im Prinzip wie folgt aus (ohne Restkapazitäten):

```
### Parameter

# Infrastruktur

param n integer;          # Knotenanzahl
set N := 1..n;           # Knotenmenge
set E within { N, N };   # gerichtete Kantenmenge
set EE within { N, N };  # gerichtete Kantenmenge
param MW integer;        # WDM-Kapazität
param MOS integer;       # OXC-Kapazität

# Bedarfe und Kosten

set D within { N, N };   # Knotenpaare mit Bedarfen
param d { D } integer;   # Bedarfe in VC4-Einheiten

param CF { E };         # Kosten für eine 2,5G Strecke auf Kante E
param CW { E };         # Kosten für ein WDM-System auf Kante E
param COS;              # Kosten je OXC
param CC;               # Kosten je WDM-Kanal
```



```

param CP;          # Portkosten am Cross-Connect

### Variablen

var f { E } integer, >= 0;      # Anzahl 2,5G-Systeme
var w { E } integer, >= 0;      # Anzahl WDM-Systeme
var v { E } integer, >= 0;      # Anzahl WDM-Kanäle
var ys { N } integer, >= 0;     # Anzahl OCX-Systeme
var z { D, EE } binary;         # Routingvariablen

### Optimierungsmodell

minimize Kosten:
    sum {(i,j) in E} ((CF[i,j] + 2*CP)*f[i,j] + CW[i,j]*w[i,j]
    + (CC + 2*CP)*v[i,j]) + sum {i in N} (COS*ys[i]);

subject to

NB1a { (s,t) in D } :
    sum { (j,i) in EE : i=s } z[s,t,j,i]
    - sum { (i,j) in EE : i=s } z[s,t,i,j] = -1;

NB1b { (s,t) in D, k in N : (k!=s) and (k!=t) } :
    sum { (j,i) in EE : i=k } z[s,t,j,i]
    - sum { (i,j) in EE : i=k } z[s,t,i,j] = 0;

NB1c { (s,t) in D } :
    sum { (j,i) in EE : i=t } z[s,t,j,i]
    - sum { (i,j) in EE : i=t } z[s,t,i,j] = 1;

NB2 { (i,j) in E } :
    sum { (s,t) in D } d[s,t]*(z[s,t,i,j]+z[s,t,j,i])
    <= v[i,j] + f[i,j];

NB3 { (i,j) in E } :
    v[i,j] <= MW*w[i,j];

```

```
NB4 { k in N } :  
  sum { (i,j) in E : (i=k) or (j=k) }  
    ( v[i,j] + f[i,j] + g[i,j] ) <= ( MOS*ys[k] );
```

Ein dazu passender Datensatz für eine Problem­instanz mit 5 Knoten, 8 Kanten und 10 Bedarfen könnte wie folgt aussehen:

```
# Testdatensatz 5 Knoten, 8 Kanten, 10 Bedarfe
```

```
# Knotenanzahl
```

```
param n := 5;
```

```
# Kanten:
```

```
set E :=
```

```
1 2
```

```
1 3
```

```
1 4
```

```
1 5
```

```
2 3
```

```
2 4
```

```
3 4
```

```
3 5
```

```
;
```

```
# gerichtete Kanten:
```

```
set EE :=
```

```
1 2 2 1
```

```
1 3 3 1
```

```
1 4 4 1
```

```
1 5 5 1
```

```
2 3 3 2
```

```
2 4 4 2
```

```
3 4 4 3
```

```
3 5 5 3
```

```
;
```

```
# WDM-Kapazität
```

```
param MW := 32;
```

```
# OXC-Kapazität
param MOS := 16;

# Knotenpaare mit Bedarfen:
set D :=
  1  2
  1  3
  1  4
  1  5
  2  3
  2  4
  2  5
  3  4
  3  5
  4  5
;

# Bedarfe:
param d :=
  1  2  8
  1  3  9
  1  4  2
  1  5  5
  2  3  4
  2  4  9
  2  5  10
  3  4  3
  3  5  8
  4  5  1
;

# Kosten für 2,5G Glasfaser auf Kante:
param CF :=
  1  2    8
  1  3    8
  1  4    8
  1  5    8
```

```
2 3 8
2 4 8
3 4 8
3 5 8
;

# Kosten für 10G Glasfaser auf Kante:
param CG :=
1 2 18
1 3 18
1 4 18
1 5 18
2 3 18
2 4 18
3 4 18
3 5 18
;

# Kosten für WDM auf Kante:
param CW :=
1 2 32
1 3 32
1 4 32
1 5 32
2 3 32
2 4 32
3 4 32
3 5 32
;

# Kosten je OXC
param COS := 10;

# Kosten je WDM-Kanal
param CC := 1;

# Portkosten
param CP := 1;
```

# Literaturverzeichnis

- Ahuja, R. K., Magnanti, T. L., und Orlin, J. B. (1993). *Network Flows*. Prentice-Hall, Englewood Cliffs.
- AMPL Optimization LLC (2000). Relational database access. <http://www.ampl.com/NEW/tables.html>, Stand: 19.1.2008.
- AMPL Optimization LLC (2006). AMPL. <http://www.ampl.com>, Stand: 19.1.2008.
- Barry, M., Bodamer, S., Späth, J., Jäger, M. und Hülsermann, R. (2004). A classification model for network survivability mechanisms. In: *ITG-Fachbericht 182 Photonische Netze, Vorträge der 5. ITG Fachtagung Photonische Netze in Leipzig*, S. 129–136. VDE Verlag, Berlin.
- Berger, L. (2003). Generalized multi-protocol label switching (GMPLS) signaling functional description. RFC3471, <http://www.ietf.org>.
- Bhandari, R. (1999). *Survivable Networks: Algorithms for Diverse Routing*. Kluwer, Boston.
- Birkan, G., Kennington, J., Olinick, E., Ortynski, A. und Spiride, G. (2002). Making a case for using integer programming to design DWDM networks. <http://engr.smu.edu/~jlk/publications.htm>, Stand: 19.1.2008.
- Bley, A., Koster, A., Kröller, A., Wessäly, R. und Zymolka, A. (2003). Kosten- und Qualitätsoptimierung in Kommunikationsnetzen. *Telekommunikation Aktuell*, **57**(7/8), 1–62.
- Chigan, C., Atkinson, G., Nagarajan, R. und Robertazzi, T. G. (2002a). Mathematical programming for joint restoration of IP over optical networks. In: *Proceedings of 36th Annual Conference on Information Sciences and Systems*, Princeton.

- Chigan, C., Atkinson, G., Nagarajan, R. und Robertazzi, T. G. (2002b). On joint restoration of packet-over-optical networks. In: *Proceedings of 18th Annual Conference of National Fiber Optic Engineers Conference (NFOEC02)*, Dallas.
- Chiu, A. L. und Strand, J. (2001). Joint IP/optical layer restoration after a router failure. <http://www.research.att.com/areas/opticalnetworking/IPoverWDMpublications.html>, Stand: 10.11.2004.
- Cisco Systems (2003). Leveraging Transport for Data Services with Virtual Concatenation (VCAT) and Link Capacity Adjustment Scheme (LCAS). [http://www.cisco.com/warp/public/cc/so/neso/meso/meipac/lcas\\_wp.pdf](http://www.cisco.com/warp/public/cc/so/neso/meso/meipac/lcas_wp.pdf), Stand: 19.1.2008.
- Conte, G., Listanti, M., Settembre, M. und Sabella, R. (2003). Protection and restoration strategies in WDM mesh networks. In: A. Bianco und F. Neri, Hrsg., *Next Generation Optical Network Design and Modelling*, S. 311–326. Kluwer, Boston.
- Dantzig, G., Fulkerson, R. und Johnson, S. (1954). Solution of a large-scale traveling-salesman problem. *Journal of the Operations Research Society of America*, **2**(4), 393–410.
- Dantzig, G. B. und Thapa, M. N. (1997). *Linear Programming*. Springer, Berlin.
- Demeester, P., Gryseels, M., Autenrieth, A., Brianza, C., Castagna, L., Signorelli, G., Clemente, R., Ravera, M., Jajszczyk, A., Janukowicz, D., Van Doorselaere, K. und Harada, Y. (1999). Resilience in multilayer networks. *IEEE Communications Magazine*, **37**(8), 70–76.
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, **1**, 269–271.
- Doucette, J. und Grover, W. D. (2001). Comparison of mesh protection and restoration schemes and the dependency on graph connectivity. In: *Proceedings 3rd International Workshop on the Design of Reliable Communication Networks (DRCN 2001)*, S. 121–128, Budapest.
- Doverspike, R. D., Phillips, S. und Westbrook, J. R. (2000). Transport network architectures in an IP world. [http://www.research.att.com/areas/transport\\_evolution](http://www.research.att.com/areas/transport_evolution), Stand: 10.11.2004.

- Duin, C. und Voß, S. (1999). The pilot method: A strategy for heuristic repetition with application to the Steiner problem in graphs. *Networks*, **34**, 181–191.
- EITO (2005). *European Information Technology Observatory 2005*. European Information Technology Observatory, Berlin.
- Ellinas, G., Bouillet, E., Ramamurthy, R., Labourdette, J.-F., Chauduri, S. und Bala, K. (2003). Routing and restoration architectures in mesh optical networks. *Optical Networks Magazine*, **4**(1), 91–106.
- Fourer, R., Gay, D. M. und Kernighan, B. W. (2003). *AMPL - A Modeling Language for Mathematical Programming*. Brooks/Cole, Duxbury, 2. Auflage.
- Fricke, M. und Höller, H. (2003). Ebenenübergreifende Planung von Telekommunikationsnetzen. In: B. Seiler, Hrsg., *taschenbuch der telekom praxis 2003*, S. 159–176. Schiele & Schön, Berlin.
- Fricke, M., Höller, H., Voß, S. und Neidlinger, S. (2005). Lösungsansätze für ein integriertes IP/MPLS- und Transportnetz. In: *ITG-Fachbericht 186 Photonische Netze, Vorträge der 6. ITG Fachtagung Photonische Netze in Leipzig*, S. 59–64. VDE Verlag, Berlin.
- Fumagalli, A. und Valcarenghi, L. (2000). IP restoration vs. WDM protection: Is there an optimal choice? *IEEE Network*, **14**, 34–41.
- GAMS Development Corporation (2006). GAMS. <http://www.gams.com>, Stand: 19.1.2008.
- Garey, M. R. und Johnson, D. S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman and Company, New York.
- Gatrell, J., Karas, J., Peschke, R. und Rodriguez, A. (2001). TCP/IP tutorial and technical overview. <http://www.ibm.com/redbooks>, Stand: 19.1.2008.
- Ghani, N. (2000). Integration strategies for IP over WDM. <http://citeseer.ist.psu.edu/ghani00integration.html>, Stand: 19.1.2008.
- Glover, F. und Kochenberger, G. A., Hrsg. (2003). *Handbook of Metaheuristics*. Kluwer.

- Glover, F., Klingman, D. und Philips, N. V. (1992). *Network Models in Optimization and their Applications in Practice*. Wiley.
- Greistorfer, P., Lokketangen, A., Voß, S. und Woodruff, D. L. (2005). Experiments concerning sequential versus simultaneous maximization of objective function and distance. *Journal of Heuristics*. Im Druck.
- Grover, W. D. (2004). *Mesh-Based Survivable Networks*. Prentice Hall, Upper Saddle River.
- Grover, W. D. und Doucette, J. (2001). Topological design of survivable mesh-based transport networks. *Annals of Operations Research*, **106**, 79–125.
- Hansen, P. und Mladenovic, N. (2003). Variable neighborhood search. In: F. Glover und G. A. Kochenberger, Hrsg., *Handbook of Metaheuristics*, S. 145–184. Kluwer.
- Harrison, E., Farrel, A. und Miller, B. (2001). Protection and restoration in MPLS networks. <http://www.dataconnection.com/news/whitepapers.htm>, Stand: 19.1.2008.
- Höller, H. und Voß, S. (2005). Software tools for a multilayer network design. *Journal of Telecommunications and Information Technology*, **3**, 57–63.
- Höller, H. und Voß, S. (2006). A heuristic approach for combined equipment-planning and routing in multi-layer SDH/WDM networks. *European Journal of Operational Research*, **171**(3), 787–796.
- Höller, H., Voß, S., Fricke, M. und Neidlinger, S. (2003). Schichtenübergreifende Netzplanung. In: *ITG-Fachbericht 175 Photonische Netze, Vorträge der 4. ITG Fachtagung Photonische Netze in Leipzig*, S. 21–28. VDE Verlag, Berlin.
- Höller, H., Voß, S., Fricke, M. und Neidlinger, S. (2004). Vergleich des Einflusses der Bedarfsstrukturen und Netztopologien von SDH/WDM Netzen auf das eingesetzte Schutzverfahren. In: *ITG-Fachbericht 182 Photonische Netze, Vorträge der 5. ITG Fachtagung Photonische Netze in Leipzig*, S. 115–120. VDE Verlag, Berlin.
- Höller, H., Melian, B. und Voß, S. (2006). Applying the pilot method to improve VNS and GRASP metaheuristics for the design of SDH/WDM networks. *European Journal of Operational Research*. Im Druck.



- Jakab, T., Arató, D. und Kállay, E. (2000). Availability analysis of optical network architectures. In: *Proceedings 2nd International Workshop on the Design of Reliable Communication Networks (DRCN 2000)*, München.
- Jereb, L., Jakab, T. und Unghváry, F. (2002). Availability analysis of multi-layer optical networks. *Optical Networks Magazine*, **3**(2), 84–94.
- Kennington, J., Lewis, K., Olinick, E., Ortynski, A. und Spiride, G. (2003). Robust solutions for the DWDM routing and provisioning problem: models and algorithms. *Optical Networks Magazine*, **4**(2), 74–84.
- Kerivin, H. und Mahjoub, A. (2005). Design of survivable networks: A survey. *Networks*, **46**(1), 1–21.
- Köhn, M. und Gauger, C. (2003). Dimensioning of SDH/WDM multilayer networks. In: *ITG-Fachbericht 175 Photonische Netze, Vorträge der 4. ITG Fachtagung Photonische Netze in Leipzig*, S. 29–33. VDE Verlag, Berlin.
- Kiefer, R. (1996). *Digitale Übertragung in SDH- und PDH-Netzen*. Expert-Verlag, Renningen.
- Krauss, O. (2002). *DWDM und Optische Netze*. Publicis, Erlangen.
- Kuhn, D. R. (1997). Sources of failure in the public switched telephone network. *IEEE Computer Magazine*, **30**(4), 31–36.
- Lai, W. und McDysan, D. (2002). Network hierarchy and multilayer survivability. RFC3386, <http://www.ietf.org>.
- Lee, B. und Kim, W. (2002). *Integrated Broadband Networks*. Artech House, Boston.
- Lee, E. K. und Lewis, D. P. (2006). Integer programming for telecommunications. In: M. G. C. Resende und P. M. Pardalos, Hrsg., *Handbook of Optimization in Telecommunications*, S. 67–102. Springer, New York.
- LINDO Systems Inc. (2006). LINGO. <http://www.lindo.com/products/lingo/lingom.html>, Stand: 19.1.2008.
- Lochmann, D. (1995). *Digitale Nachrichtentechnik*. Verlag Technik GmbH, Berlin.
- Maier, G., Pattavina, A., De Patre, S. und Martinelli, M. (2002). Optical network survivability: Protection techniques in the WDM layer. *Photonic Network Communications*, **4**, 251–269.

- Marti, R. (2003). Multi-start methods. In: F. Glover und G. A. Kochenberger, Hrsg., *Handbook of Metaheuristics*, S. 355–368. Kluwer, Boston.
- Matsumoto, M. und Nishimura, T. (1998). Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation*, **8**(1), 3–30.
- Melian, B. (2003). *Metaheuristic Optimization for WDM Network Planning*. Ph.D. thesis, University of La Laguna, Tenerife, Spain.
- Melian, B., Laguna, M. und Moreno-Perez, J. A. (2004). Capacity expansion of fiber optic networks with WDM systems: Problem formulation and comparative analysis. *Computers & Operations Research*, **31**(3), 461–472.
- Melian, B., Laguna, M. und Moreno-Perez, J. A. (2005a). Minimizing the cost of placing and sizing wavelength division multiplexing and optical cross-connect equipment in a telecommunications network. *Networks*, **45**(4), 199–209.
- Melian, B., Höller, H. und Voß, S. (2005b). VNS and GRASP metaheuristics for the design of SDH/WDM networks. In: *Proceedings of the 18th Mini Euro Conference on VNS, 23-25 November 2005, Puerto de la Cruz, Spain*.
- Melian, B., Höller, H. und Voß, S. (2006). Designing WDM networks by a variable neighborhood search. *Journal of Telecommunications and Information Technology*, **4**, 15–20.
- Michalewicz, Z. und Fogel, D. B. (2004). *How to Solve It: Modern Heuristics*. Springer, Berlin, 2. Auflage.
- Minoux, M. (2006). Multicommodity network flow models and algorithms in telecommunications. In: M. G. C. Resende und P. M. Pardalos, Hrsg., *Handbook of Optimization in Telecommunications*, S. 163–184. Springer, New York.
- Mobiot, F., Sanso, B. und Girard, A. (2003). A method for the integrated design of reliable GMPLS networks. In: *Proceedings of the Fourth International Workshop on Design of Reliable Communication Networks (DCRN 2003)*, S. 53–60.
- Orlowsky, S. und Wessäly, R. (2004). An integer programming model for multi-layer network design. <http://www.zib.de/PaperWeb/abstracts/ZR-04-49>, Stand 19.1.2008.

- Phillips, S., Reingold, N. und Doverspike, R. (2002). Network studies in IP/optical layer restoration. [http://www.research.att.com/areas/transport\\_evolution](http://www.research.att.com/areas/transport_evolution), Stand: 10.11.2004.
- Pitsoulis, L. S. und Resende, M. G. C. (2001). Greedy randomized adaptive search procedures. <http://www.research.att.com/~mgcr/papers.html>, Stand: 19.1.2008.
- Poestel, J. (1994). Internet protocol. RFC760, <http://www.ietf.org>.
- Polya, G. (1957). *How to solve it*. Doubleday, Garden City, 2nd Auflage.
- Pongpaibool, P. (2004). *Survivability of GMPLS-based IP-over-optical networks*. Ph.D. thesis, Carnegie Mellon University, Department of Electrical and Computer Engineering.
- Ramamurthy, S. und Mukherjee, B. (1999a). Survivable WDM mesh networks, part I - protection. In: *Proceedings of the INFOCOM'99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies.*, Band 2, S. 744–751. IEEE.
- Ramamurthy, S. und Mukherjee, B. (1999b). Survivable WDM mesh networks, part II - restoration. In: *Proceedings of ICC'99, Montreal, Canada*, Band 3, S. 2023–2030.
- Ramaswami, R. und Sivarajan, K. N. (2002). *Optical Networks*. Morgan Kaufmann, San Francisco, 2. Auflage.
- Reß, H. und Viebeck, G. (2000). *Datenstrukturen und Algorithmen*. Carl Hanser Verlag, München.
- Resende, M. G. C. (2001). A bibliography of GRASP. <http://www.research.att.com/~mgcr/doc/graspbib.pdf>, Stand: 19.1.2008.
- Resende, M. G. C. und Ribeiro, C. C. (2003). Greedy randomized adaptive search procedures. In: F. Glover und G. A. Kochenberger, Hrsg., *Handbook of Metaheuristics*, S. 219–249. Kluwer, Boston.
- Rosen, E., Viswanathan, A. und Callon, R. (2001). Multiprotocol label switching architecture. RFC3031, <http://www.ietf.org>.
- Sexton, M. und Reid, A. (1997). *Broadband Networking: ATM, SDH and SONET*. Artech House, Boston.

- Soni, S., Gupta, R. und Pirkul, H. (1999). Survivable network design: The state of the art. *Information Systems Frontiers* 1, S. 303–325.
- Stern, T. E. und Bala, K. (1999). *Multiwavelength optical networks: a layered approach*. Addison-Wesley, Reading.
- Suurballe, J. W. (1974). Disjoint paths in a network. *Networks*, 4, 125–145.
- Suurballe, J. W. und Tarjan, R. E. (1984). A quick method for finding shortest pairs of disjoint paths. *Networks*, 14, 325–336.
- Theelen, B. D., Voeten, J. P. M., van der Putten, P. H. A., Dorren, H. J. S. und Stevens, M. P. J. (2002). Concurrent support of higher-layer protocols over WDM. *Photonic Network Communications Journal*, 4(1), 47–62.
- Voß, S., Fink, A. und Duin, C. (2005). Looking ahead with the Pilot Method. *Annals of Operations Research*, 136, 285–302.
- Wagner, R. (2003). Mersenne twister random number generator. <http://www.personal.engin.umich.edu/~wagnerr/MersenneTwister.html>, Stand: 19.1.2008.
- Wessály, R. (2000). *Dimensioning Survivable Capacitated Networks*. Doktorarbeit, Technische Universität Berlin, Berlin, Germany.
- Wessály, R. und Pióro, M. (2006). SNDlib. <http://sndlib.zib.de>, Stand: 19.1.2008.
- Wilde, A. (1999). *SDH in der Praxis*. VDE-Verlag, Berlin.
- Wu, T. H. (1992). *Fiber Network Service Survivability*. Artech House, Boston.
- Xu, L., Perros, H. G. und Rouskas, G. N. (2000). Transporting IP packets over light: A survey. Technical Report TR-2000-03, North Carolina State University.
- Yuan, D. (2001). An annotated bibliography in communication network design and routing. In: *Optimization Models and Methods for Communication Network Design and Routing*. Linköpings Universitet, Institute of Technology, Linköping.
- Zang, H. (2003). *WDM Mesh Networks: Management and Survivability*. Kluwer, Boston.

- Zhang, H. und Durrezi, A. (2002). Differentiated multi-layer survivability in IP/WDM networks. In: *8th IEEE-IFIP Network Operations and Management Symposium*, S. 681–694.
- Zhu, K., Zang, H. und Mukherjee, B. (2003). Exploiting the benefits of virtual concatenation in optical transport networks. OFC2003, March 2003, Anaheim.