

Universität Stuttgart

Institut für Nachrichtenvermittlung und Datenverarbeitung

Prof. Dr.-Ing. habil. Dr. h.c. P.J. Kühn

64. Bericht über verkehrstheoretische Arbeiten

**Methoden zur Planung von hierarchischen
Hochgeschwindigkeitsnetzen (MAN)**

von

Ulrich Gremmelmaier

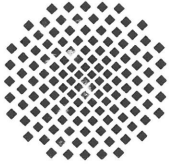
1996

D 93

© 1996 Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart

Druck: E. Kurz & Co., Druckerei + Reprografie GmbH., Stuttgart

ISBN 3-922403-74-3



University of Stuttgart

Institute of Communications Switching and Data Technics

Prof. Dr.-Ing. habil. Dr. h.c. P.J. Kühn

64th Report on Studies in Congestion Theory

**Methods for Planning of hierarchical
Metropolitan Area Networks**

by

Ulrich Gremmelmaier

1996

Methods for Planning of hierarchical Metropolitan Area Networks

Summary

Metropolitan Area Networks (MANs) have been developed to follow up Local Area Networks (LANs) as the next generation of high speed communication networks. MANs based on the IEEE 802.6 standard have already been introduced in the public sector as both a network for early deployment of broadband services and later as access network for Broadband-ISDN. It is expected that due to the growing demand of broadband services there is a need to apply MANs in wider geographical areas, i.e. up to several hundred kilometers, to supply a large number of customers with broadband facilities.

Building up a broadband network infrastructure requires large investments. Therefore, a network operator depends on a systematic planning and design process in order to make efficient use of resources like transmission lines and networking equipment. Consequently, there is a need for methods and tools which can be applied to MAN planning.

The main objective of this thesis is to develop and evaluate methods for the planning and cost-effective design of MANs based on DQDB (Distributed Queue Dual Bus). The methods are an essential part of a MAN planning tool which offers a variety of functions to design and analyze the network structure.

The introduction focuses on some of the major trends in data communications of public networks and gives an outline of the following chapters.

Chapter 2 introduces the basic concepts of Local and Metropolitan Area Networks such as switching techniques, media access control, routing and LAN-MAN-interworking. The main focus is on the architecture of MAN subnetworks according to the DQDB standard. In particular, the most important features of high speed data services relying on MANs, i.e. the Switched Multi-Megabit Data Service (SMDS) and the Connectionless Broadband Data Service (CBDS), are described. SMDS and CBDS have not only been designed for regional networks but will be offered in wide area networks. In the near future, they will be mainly realized by using MAN Switching Systems (MSS) as a transport network. Usually, a MSS comprises several DQDB subnetworks interconnected via routers. The last part of chapter 2 illustrates a typical MSS architecture.

Chapter 3 is mainly dedicated to modeling aspects of communication networks and optimization techniques suited for the topological design of MANs. Modeling aspects include

traffic models as well as graphs. Moreover, some theoretical background for the MAN planning and design process is provided in this chapter. Typically, topological design problems can be formulated as a combinatorial optimization problem. A significant number of these problems have been proved to be very difficult to solve. In most cases, heuristic algorithms are the only way to solve them. Therefore, the following heuristic algorithms are explained in detail: Local search algorithms, Simulated Annealing and Genetic Algorithms. In contrast to the local search algorithm, Simulated Annealing and Genetic Algorithms are statistical optimization methods.

The chapter closes with an overview of the planning process of communication networks concentrating on the basic topological design problems of packet switched wide area networks.

Planning Models are a prerequisite for effectively solving the planning problem. Chapter 4 presents a planning model which can be formulated as a non-linear integer optimization problem. The objective is to design a cost-effective network structure fulfilling *Quality of Service* requirements.

After discussing related approaches in the literature, several aspects being considered in the MAN planning model are investigated. They include simplifications of the underlying transmission network infrastructure and the capabilities of the network elements, i.e. stations and routers. Furthermore, a cost model is presented which allows for an efficient evaluation of the costs of a MAN topology. Also, it is shown how the traffic flow analysis and the queueing network analysis can be performed. Taking the above considerations into account an optimization model can be formulated. The model incorporates constraints imposed by topology, flow conservation, admissible load and path delay restrictions. The objective function comprises the sum of the total costs of all leased lines in a MSS. Finally, the complexity of the MAN design problem is discussed.

In order to solve the MAN design problem presented in chapter 4, heuristic algorithms have been developed. Chapter 5 provides a detailed description of three MAN design methods as well as a case study for their evaluation.

The first algorithm is a decomposition approach which consists of two parts: A *Clustering and Load sharing Algorithm* (CLA) provides for an initial assignment of MAN stations to subnetworks. The CLA uses the geographical distribution of the stations and the locality of their traffic flows. The second part comprises an algorithm to interconnect the subnetworks and to finally fix the bus topology in the subnetworks.

The second algorithm is based on Simulated Annealing which has been proved as an efficient method in many areas of combinatorial optimization. For this algorithm, the move set required for local transformations of the MAN topology is presented. Moreover, the overall structure of the algorithm is explained.

Genetic Algorithms form the basis of the third method. Based on experience with simpler planning problems, an efficient coding scheme and appropriate genetic operators have been designed.

In a case study, the results of the developed planning methods are compared with results derived from a distribution of randomly generated MAN topologies. Moreover, they are tested against a lower bound of the MAN design problem. The results of the case study show that the algorithms perform very well. Hence, it can be stated that topologies with low costs and sufficient Grade of Service can be derived. In some cases, the costs of the designed topologies are very close to the lower bound. In particular, combining the CLA and the algorithm based on Simulated Annealing takes advantage of the favourable characteristics of both. Regarding the run times, the methods can be well applied to problems of large size.

The last part of chapter 5 demonstrates how the methods can be used to extend an existing MSS and to improve its availability.

Chapter 6 concentrates on the MAN planning tool which is based on the methods described in chapter 5. This includes a presentation of the functionality and the modular software structure of the tool which allows for further extensions such as additional analysis methods. Moreover, some important implementation aspects are highlighted. A discussion of the general applicability of the planning tool closes the chapter.

Chapter 7 summarizes the main results and gives an outlook to possible extensions of the planning methods for other network types.

Inhaltsverzeichnis

Inhaltsverzeichnis	1
Abkürzungen	5
Formelzeichen	7
1 Einleitung	9
1.1 Planung zukünftiger Netzinfrastrukturen	9
1.2 Übersicht über die Arbeit	11
2 Lokale und Regionale Hochgeschwindigkeits-Kommunikationsnetze	13
2.1 Merkmale und Funktionsprinzipien	13
2.1.1 Dienstbegriff	13
2.1.2 Netzstrukturen	14
2.1.3 Vermittlungsverfahren	15
2.1.4 Schichtenmodell	16
2.1.5 Medienzugriffsverfahren	17
2.1.6 Adressierungskonzepte	19
2.1.7 Netzkopplung	21
2.1.8 Verkehrslenkung	23
2.2 Regionale Hochgeschwindigkeitsnetze auf der Basis von DQDB	24
2.2.1 Übersicht	24
2.2.2 DQDB	25
2.2.2.1 Topologie und Übermittlungstechnik	25
2.2.2.2 Stationsmodell	26

2.2.2.3	DQ-Protokoll	27
2.2.3	Hochgeschwindigkeits-Datendienste	29
2.2.3.1	Dienstbeschreibung	30
2.2.3.2	Zugangsnetz	31
2.2.4	MAN-Vermittlungssystem	32
2.2.4.1	Netzstruktur	32
2.2.4.2	Realisierung	34
2.2.4.3	Netzkopplung und Verkehrslenkung	35
2.2.5	Evolution zum B-ISDN	37
3	Grundlagen der Netz- und Konfigurationsplanung	39
3.1	Modellierung von Kommunikationsnetzen	39
3.1.1	Verkehrsmodelle	39
3.1.1.1	Stochastische Prozesse	40
3.1.1.2	Systemmodelle	41
3.1.2	Graphen	42
3.1.2.1	Begriffe und Definitionen	42
3.1.2.2	Algorithmen	44
3.2	Optimierungsverfahren	45
3.2.1	Kombinatorische Optimierung	45
3.2.2	Exakte und heuristische Verfahren	46
3.2.3	Simulated Annealing	49
3.2.4	Genetische Algorithmen	51
3.2.4.1	Übersicht	51
3.2.4.2	Codierung	52
3.2.4.3	Fitness, Selektion und Reproduktion	52
3.2.4.4	Genetische Operatoren	54
3.2.4.5	Arbeitsweise	56
3.3	Planung von Kommunikationsnetzen	56
3.3.1	Übersicht	56

3.3.2	Netzebenen	58
3.3.3	Zeithorizont	58
3.3.4	Verkehrstechnische Netzgestaltung	59
3.3.5	Topologischer Entwurf	60
3.3.5.1	Zentrale Strukturen	60
3.3.5.2	Dezentrale Strukturen	60
4	Ein Planungsmodell für MAN-Vermittlungssysteme	63
4.1	Übersicht	63
4.2	Modellierung	65
4.2.1	Netzelemente und Geographie	66
4.2.2	Kostenmodell	67
4.2.3	Verkehrslenkung und Flußanalyse	68
4.2.4	Warteschlangenmodell	71
4.3	Formulierung des Problems	75
4.4	Komplexität	78
5	Verfahren für den topologischen Entwurf von MAN-Vermittlungssystemen	81
5.1	Dekompositionsverfahren	81
5.1.1	Entwurf der Teilnetze	83
5.1.2	Verbindung der Teilnetze	86
5.1.3	Ablaufschema	89
5.2	Verfahren basierend auf Simulated Annealing	90
5.2.1	Struktur der Nachbarschaft	90
5.2.2	Ablaufschema	92
5.3	Verfahren basierend auf Genetischen Algorithmen	94
5.3.1	Codierung und Genetische Operatoren	95
5.3.2	Ablaufschema	97
5.4	Vergleich der Verfahren	98
5.4.1	Parameter der Netzelemente	98

5.4.2	Vergleichsstrategie	99
5.4.3	Vergleich mit zufällig ermittelten Topologien	101
5.4.4	Einfluß der Startbedingungen	105
5.4.5	Einfluß der maximal zulässigen Transferzeit	108
5.4.6	Abschätzung der Rechenzeiten	110
5.4.7	Diskussion	112
5.5	Erweiterungen	113
5.5.1	Ausbau bestehender MAN-Vermittlungssysteme	113
5.5.2	Erhöhung der Verfügbarkeit	116
6	Entwurf und Realisierung eines Planungswerkzeugs	119
6.1	Anforderungen	119
6.2	Funktionalität und Architektur	121
6.2.1	Hierarchiebildung	122
6.2.2	Netzdatenbasis	124
6.2.3	Planungsverfahren und Netzanalyse	125
6.2.4	Benutzerschnittstelle	126
6.3	Implementierungsaspekte	126
6.4	Entwurfsablauf	128
6.5	Bewertung und Einsatzmöglichkeiten	129
7	Zusammenfassung und Ausblick	131
	Literaturverzeichnis	134

Abkürzungen

ACF	Access Control Field
ACM	Association for Computing Machinery
AFI	Authority and Format Identifier
ANSI	American National Standards Institute
ATM	Asynchroner Transfer-Modus
B-ISDN	Broadband Integrated Services Digital Network
CAN	Customer Access Network
CBDS	Connectionless Broadband Data Service
CEQ	Customer Equipment
CGW	Customer Gateway
CLA	Clustering- und Lastteilungsalgorithmus
CPU	Central Processing Unit
CS	Circuit Switching
CSMA/CD	Carrier Sense Multiple Access with Collision Detection
DQDB	Distributed Queue Dual Bus
DSP	Domain Specific Part
DST	Destination Spanning Tree
EGW	Edge Gateway
ETSI	European Telecommunications Standards Institute
FDDI	Fibre Distributed Data Interface
FIFO	First In First Out
GLOBECOM	Global Telecommunications Conference
GN	Gathering Network
ICC	International Conference on Communications
ICCC	International Conference on Computer Communications
IDI	Initial Domain Identifier
IDP	Initial Domain Part
IEEE	Institute of Electrical and Electronics Engineers
IMI	Inter MSS Interface
INFOCOM	The Conference on Computer Communications
INI	Inter Network operator Interface
ISDN	Integrated Services Digital Network
ISO	International Organization for Standardization
ISS	International Switching Symposium
ITC	International Teletraffic Congress
ITU	International Telecommunication Union
IWU	Interworking Unit

KE	Kosteneinheit
LAN	Local Area Network
LLC	Logical Link Control
MAC	Media Access Control
MAN	Metropolitan Area Network
MSS	MAN Switching System
NMC	Network Management Center
NP	Non-deterministic Polynomial-time problem
OSI	Open Systems Interconnection
PDU	Protocol Data Unit
PMX	Partially Matched Crossover
PS	Packet Switching
QPSX	Queued Packet and Synchronous Circuit Exchange
SDH	Synchronous Digital Hierarchy
SAP	Service Access Point
SDU	Service Data Unit
SIP	SMDS Interface Protocol
SH	Segment Header
SMDS	Switched Multi-megabit Data Service
SNI	Subscriber Network Interface
ST	Spanning Tree
TSS	Telecommunication Standardization Sector
VC	Virtual Connection
VP	Virtual Path
VPN	Virtual Private Network
WAN	Wide Area Network

Formelzeichen

c	Variationskoeffizient
C_{ij}	Kapazität des Übertragungsabschnittes (i, j)
C_R	Router-Kapazität
C_S	Kapazität eines Teilnetzes
d	Entfernung
d_{ij}	Entfernung zwischen den Knoten i und j
D_i	Durchmesser von Cluster i
D_{max}	maximal zulässiger Durchmesser
$E[T_A]$	Erwartungswert der Zufallsvariablen T_A
$E[T_A^n]$	n -tes gewöhnliches Moment der Zufallsvariablen T_A
$f_A(t)$	Verteilungsdichtefunktion der Zufallsvariablen T_A
$F_A(t)$	Verteilungsfunktion der Zufallsvariablen T_A
$f_{i,q}^A$	Verkehrsfluß auf Bus A an Position i
$f_{i,q}^B$	Verkehrsfluß auf Bus B an Position i
f_{ij}	Verkehrsfluß über einen Übertragungsabschnitt (i, j)
f_{ij}^{mn}	Verkehrsfluß über eine Router-Router-Kopplung (i, j) , wobei i in Teilnetz m und j in Teilnetz n ist
G	allgemeiner Graph (Menge der Knoten und Kanten)
G_T	Topologiegraph
g	Zielfunktion
g_d	Gewichtungsfaktor (Dekompositionsverfahren)
γ_{ij}	Verkehrsfluß zwischen Station i und j
h	mittlere Bedienzeit (Zufallsvariable T_h)
$h^{(2)}$	2. Moment der Bedienzeit
K	Gesamtkosten
K_l	Kostenfunktion einer Übertragungsleitung vom Typ l
K_l^0	Fixkosten
κ_l	Normierungsfaktor der Kostenfunktion
λ	Ankunftsrate, Verkehrsfluß
λ_{int}^i	interner Verkehrsfluß in einem Cluster
λ_{ext}^i	externer Verkehrsfluß aus einem Cluster
M	Zahl der Teilnetze
N	Zahl der Stationen
P_A	Akzeptanzwahrscheinlichkeit (Simulated Annealing)
p_i	Belegungswahrscheinlichkeit eines Zeitschlitzes an Station i
\mathcal{R}	Menge der Pfade
ρ	Auslastung
ρ_{Cl}	Auslastungsgrenze (Clustering)

ρ_g	Schwellwert für das Einfügen einer direkten Router-Router-Kopplung
ρ_R	Auslastungsgrenze (Router)
ρ_S	Auslastungsgrenze (Teilnetz)
ρ_V	Auslastungsgrenze (Verbindung der Teilnetze)
S_i	Teilnetz i (Menge der Netzelemente in Teilnetz i)
t	Variable für die Zeit
t_b	mittlere Buszugriffszeit
t_p	mittlere Protokollbearbeitungszeit
t_l	Laufzeit (Infinite Server)
t_{mn}^R	mittlere Pfadtransferzeit zwischen Station m und Station n
t_{max}^R	maximal zulässige Pfadtransferzeit
$t_{max,ms}^R$	gemessene maximale Pfadtransferzeit
t_z	Dauer eines Zeitschlitzes
T_A	Zufallsvariable für ein Zeitintervall
τ	Faktor (Simulated Annealing)
Θ	Kontrollparameter (Simulated Annealing)
$\text{VAR}[T_A]$	Varianz der Zufallsvariablen T_A
x_{ij}^{mn}	Boolesche Variable, die eine Router-Router-Kopplung (i, j) bezeichnet, wobei i in Teilnetz m und j in Teilnetz n ist
y_{ij}^m	Boolesche Variable, die einen Übertragungsabschnitt (i, j) in Teilnetz m bezeichnet
$Z(M, k)$	Zahl der Möglichkeiten für eine Aufteilung (M, k)

Kapitel 1

Einleitung

1.1 Planung zukünftiger Netzinfrastrukturen

Die Kommunikationstechnik ist heute unbestritten einer der wichtigsten Wirtschaftsfaktoren moderner Industriegesellschaften und Prognosen deuten darauf hin, daß die Telekommunikationsindustrie über das Jahr 2000 hinaus diese Stellung noch ausbauen kann. Eine treibende Kraft ist dabei die steigende Nachfrage von Netzbenutzern nach Dienstleistungen, die im Bereich neuer Kommunikationsformen entstehen, wie z. B. Multimedia-Anwendungen, die Sprach-, Daten- und Bildkommunikation integrieren können. Die rasche Entwicklung auf den Gebieten der Mikroelektronik sowie der optischen Nachrichtentechnik bietet dafür die Voraussetzungen. Anwendungen wie Multimedia sind nur durch eine Weiterentwicklung heute verfügbarer Netztechniken und -infrastrukturen hin zur Breitbandtechnik realisierbar. Mit dem starken Anwachsen der Datenkommunikation stoßen darüber hinaus traditionelle Kommunikationsnetze an ihre Leistungsgrenzen [162], so daß auch dadurch der Aufbau einer leistungsfähigen Breitband-Netzinfrastruktur notwendig wird.

In den letzten Jahren haben sich dafür im Bereich öffentlicher Netze zwei Netztechniken herauskristallisiert, die aus unterschiedlichen Entwicklungslinien resultieren.

Regionale Hochgeschwindigkeitsnetze (*Metropolitan Area Networks*, MANs) haben sich aus dem Gebiet der Datenkommunikation entwickelt und zielen als wichtigste Anwendung auf die Kopplung lokaler Netze im regionalen Bereich ab, die auch auf den Weitverkehrsbereich ausgedehnt werden kann. Dies ist insbesondere erforderlich, da lokale Netze in ihrer geographischen Ausdehnung auf wenige Kilometer beschränkt sind und ein zunehmender Bedarf der Teilnehmer nach leistungsfähiger Kommunikation auch über den öffentlichen Bereich hinweg entsteht. Beispiele sind die Verbindung von Universitätsnetzen oder verschiedenen Firmenstandorten in einer Region. Netze auf der Basis von MANs wurden in einfacheren Konfigurationen bereits erprobt und in öffentlichen Netzen eingeführt.

Als Weiterentwicklung des Schmalband-ISDN (*Integrated Services Digital Network*) befindet sich mittlerweile das Breitband-ISDN (B-ISDN), dessen Funktionen durch die *International Telecommunication Union* (ITU) standardisiert werden, in der Erprobungsphase. Der vollständige Ausbau zu einer Breitband-Netzinfrastruktur auf der Basis des B-ISDN soll sich in zwei Phasen vollziehen [39, 51]: Zunächst werden Breitbanddienste auf der Basis regionaler Hochgeschwindigkeitsnetze angeboten, danach werden MANs schrittweise über das B-ISDN verbunden und allmählich die Rolle von Zugangsnetzen übernehmen. Auf lange Sicht wird erwartet, daß das B-ISDN dominieren wird.

Mit der Einführung dieser neuen Generation von Netzen wird auch die Flexibilität von Netzkonfigurationen deutlich gesteigert. Hier sind vor allem Funktionen zur Einrichtung sog. *Virtual Private Networks* (VPNs) zu nennen, mit denen einem privaten Kunden ein virtuelles Netz zur exklusiven Nutzung bereitgestellt wird. Zu diesem Zweck ist die Einrichtung von geschlossenen Benutzergruppen (*Closed User Groups*) vorgesehen. Damit wird beispielsweise auch der Aufbau eines unternehmensweiten Netzes (*Corporate Network*) vereinfacht, das geographisch verteilte Standorte eines Unternehmens unter Nutzung der Übermittlungsdienste des öffentlichen Netzes miteinander verknüpft und das für den Teilnehmer wie ein einziges Netz erscheint [38].

Breitband-Netzinfrastrukturen dieser Art sollten auf Grund der hohen Investitionen, die für den Netzbetreiber zu erwarten sind, systematisch entworfen und in Betrieb genommen werden, damit eine möglichst optimale Nutzung von Ressourcen erzielt werden kann. Durch die bevorstehende Deregulierung des Telekommunikationssektors in vielen Ländern erscheint eine sorgfältige Planung als Wettbewerbsfaktor für einen Netzbetreiber eher noch an Bedeutung zu gewinnen. Darüber hinaus ist bei einem hohen Grad an Flexibilität auch mit häufigen Änderungen der Netzkonfiguration zu rechnen. Diese Anforderungen können nur mit Hilfe rechnergestützter Planungsverfahren erfüllt werden.

Die vorliegende Arbeit befaßt sich mit der Planung und dem Entwurf von regionalen Hochgeschwindigkeitsnetzen. Dazu werden rechnergestützte Planungsverfahren benötigt, die sich zum Entwurf einer solchen Netzinfrastruktur eignen. Für dieses komplexe Entwurfsproblem gibt es bisher keinen allgemeinen Ansatz, der über einzelne Aspekte der Analyse und Dimensionierung regionaler Hochgeschwindigkeitsnetze hinausgeht. Angestrebt wird daher, ein umfassendes Verfahren zu entwickeln. Dabei werden vorrangig Gesichtspunkte der Konfigurationsplanung betrachtet mit dem Ziel, unter Einhaltung einer gegebenen Dienstgüte eine möglichst kostengünstige Netzkonfiguration zu erhalten. Auf Grund der Einführung immer leistungsstärkerer Rechnergenerationen können auch bei komplexen Problemen dieser Art umfangreiche Planungsverfahren entwickelt werden, die dennoch effizient sind.

Damit lassen sich die Ziele der Arbeit folgendermaßen umreißen:

- Darstellung der prinzipiellen Netzarchitekturen regionaler Hochgeschwindigkeitsnetze im Bereich öffentlicher Netze.

- Entwicklung eines Modells zur Planung regionaler Hochgeschwindigkeitsnetze, das für hierarchische und nichthierarchische Netzstrukturen geeignet ist.
- Entwurf von Planungsverfahren, die sich sowohl zur Planung neuer Netzstrukturen als auch zur Erweiterung bestehender Netzstrukturen eignen.
- Vergleich der Leistungsfähigkeit der entwickelten Planungsverfahren.
- Umsetzung der gewonnenen Verfahren und Erkenntnisse in ein Netzplanungswerkzeug, das den Netzplaner beim Entwurf unterstützt.

1.2 Übersicht über die Arbeit

Kapitel 2 enthält eine Einführung in die Konzepte lokaler und regionaler Hochgeschwindigkeitsnetze. Dazu werden die für das Verständnis der nachfolgenden Kapitel benötigten Fachbegriffe erläutert. Der Schwerpunkt dieses Kapitels liegt auf der Darstellung der Architektur von MANs, die auf DQDB (*Distributed Queue Dual Bus*) basieren. Das schließt auch eine Beschreibung darauf aufbauender Dienstkonzepte mit ein.

Die Grundlage der Planungsverfahren bilden Verkehrsmodelle, graphentheoretische Modelle und Algorithmen sowie Verfahren der kombinatorischen Optimierung. Kapitel 3 umfaßt eine Darstellung dieser für die Netzplanung wichtigen Konzepte. Darüber hinaus werden Netzplanungsverfahren selbst durch eine Reihe von Merkmalen in das Gebiet des Netzmanagements eingeordnet und die Grundlagen topologischer Entwurfsverfahren beschrieben, wobei der Schwerpunkt auf dezentralen Strukturen liegt.

In Kapitel 4 wird nach einem Überblick über bisherige Ansätze ein neues Planungsmodell für MAN-Vermittlungssysteme aufgestellt. Insbesondere wird auf die einzelnen Modellierungsaspekte eingegangen, die die Voraussetzung für das Planungsmodell bilden. Das beinhaltet sinnvolle Vereinfachungen der in der Realität gegebenen Abhängigkeiten, ohne die das Planungsproblem nur in sehr kleiner Dimension gelöst werden könnte. Ziel ist es außerdem, zu prüfen, ob bereits bekannte Verfahren zur Lösung des aufgestellten Planungsproblems herangezogen werden können oder ob spezielle Lösungsverfahren entwickelt werden müssen. Dazu wird die Komplexität des Problems näher untersucht.

Kapitel 5 enthält eine Darstellung der Methoden zur Planung und zum Entwurf von MAN-Vermittlungssystemen. Zur Lösung des Problems werden drei neue Verfahren vorgestellt, die zum Teil auf statistischen Optimierungsverfahren beruhen. Die einzelnen Schritte der Planungsverfahren werden ausführlich erläutert. Da das Planungsproblem nicht exakt gelöst werden kann, bieten die drei Verfahren eine Möglichkeit, die Qualität der Lösungen zu vergleichen. Daher werden die Verfahren hinsichtlich ihrer Leistungsfähigkeit untersucht, was auch

den Vergleich mit unteren Schranken und zufällig ermittelten Lösungen einschließt. Außerdem werden Verfahren zur Erweiterung bestehender Netze vorgestellt, die sich aus einer Anwendung der oben genannten Verfahren ergeben. Eine Diskussion, wie die Verfügbarkeit von MAN-Vermittlungssystemen durch Planungsmaßnahmen erhöht werden kann, beschließt das Kapitel.

Kapitel 6 beschreibt den Entwurf und die Implementierung eines Softwarewerkzeugs zur Planung von MAN-Vermittlungssystemen, das im Kern auf den in Kapitel 5 vorgestellten Verfahren beruht. Das Kapitel umfaßt die Beschreibung der Funktionalität und der Softwarearchitektur sowie ausgewählter Implementierungsaspekte der entwickelten Software. Darüber hinaus wird der typische Entwurfsablauf beschrieben und das Planungswerkzeug hinsichtlich der Anwendbarkeit bewertet.

Kapitel 7 beschließt die Arbeit mit einer Zusammenfassung der wichtigsten Ergebnisse und gibt einen Ausblick auf zukünftige Entwicklungen.

Kapitel 2

Lokale und Regionale Hochgeschwindigkeits- Kommunikationsnetze

Regionale Hochgeschwindigkeitsnetze (*Metropolitan Area Network*, MANs) sind ursprünglich aus der Weiterentwicklung lokaler Netze (*Local Area Network*, LANs) entstanden. In diesem Kapitel sollen zunächst die grundsätzlichen Merkmale beider Netztypen beschrieben werden, soweit sie für das Verständnis der folgenden Kapitel benötigt werden. Der Schwerpunkt dieses Kapitels liegt dann auf der Darstellung der Architektur regionaler Hochgeschwindigkeitsnetze und der darauf aufbauenden Dienstkonzepte.

2.1 Merkmale und Funktionsprinzipien

2.1.1 Dienstbegriff

Dienste bezeichnen die Gesamtheit aller Kommunikationsformen wie Sprach-, Daten-, Fest- und Bewegtbildkommunikation [81]. Sie sind durch Dienstmerkmale gekennzeichnet, die die unterschiedlichen Eigenschaften eines Dienstes, z. B. die maximale Datenübertragungsrate oder Funktionen zur Punkt-zu-Mehrpunkt-Übermittlung, charakterisieren. Prinzipiell werden Übermittlungs- und Teledienste unterschieden. Übermittlungsdienste, wie etwa die Datenübertragung im ISDN (*Integrated Services Digital Network*), dienen der code- und anwendungsunabhängigen Nachrichtenübermittlung zwischen Benutzer-Netzschnittstellen. Teledienste dagegen ermöglichen die Kommunikation zwischen Endteilnehmern unter Einschluß

der Funktionen der Endgeräte. Beispiele hierfür sind das ISDN-Fernsprechen oder Videokonferenzen. Da der Telekommunikationsmarkt seit einigen Jahren einen raschen Aufschwung erfährt, ist im Bereich der Teledienste in Zukunft mit einer weitaus größeren Vielfalt, vor allem bei Multimedia- und mobilen Diensten, zu rechnen [6].

Im Zusammenhang mit Diensten ist der Begriff der Dienstgüte (*Quality of Service*), der die Gesamtheit der Qualitätsmerkmale eines Netzes aus der Sicht des Dienstbenutzers beschreibt, von wesentlicher Bedeutung. Aus der Dienstgüte können Anforderungen an die Leistungsfähigkeit eines Kommunikationsnetzes, die durch die Netzgüte definiert ist, abgeleitet werden. Aufgabe der Netzplanung und -dimensionierung ist es, ein Kommunikationsnetz so auszulegen, daß die Netzgüte stets gewährleistet ist. Typische Beispiele für Dienstgüteparameter sind die maximal zulässige Paketverlustwahrscheinlichkeit oder die Verfügbarkeit.

2.1.2 Netzstrukturen

Die Netzstruktur beschreibt den topologischen Aufbau des Nachrichtennetzes aus Netzknoten (Vermittlungsstellen), Einrichtungen und Verbindungs- bzw. Anschlußleitungen. Die Topologie bestimmt die Netzeigenschaften beispielsweise hinsichtlich der Verfügbarkeit oder der Ende-zu-Ende-Verzögerungszeit. Im Bereich lokaler Netze sind drei elementare Topologien bekannt, die in ihrer Ausdehnung auf wenige Kilometer beschränkt sind:

- Bus,
- Ring,
- baumförmige Topologie.

Die Festlegung der Topologie ist meist integraler Bestandteil der Norm, die die Eigenschaften des LAN definiert (s. Abschnitt 2.1.5). Durch die höheren Anforderungen bei MANs in Bezug auf die Verfügbarkeit wurden weitere, aus den einfachen Topologien abgeleitete Netzstrukturen entwickelt. Beispielhaft seien die Doppelringkonfiguration bei FDDI (*Fiber Distributed Data Interface*) [136] und Dualbussysteme (vgl. Abschnitt 2.2.2), die ringförmig verlegt werden, genannt. Bei diesen Topologien kann der Netzbetrieb auch nach dem Ausfall eines Übertragungsabschnittes aufrecht erhalten werden. Dies wird typischerweise durch automatische Rekonfigurationsmechanismen erreicht.

Darüber hinaus lassen sich durch die Verknüpfung elementarer Topologien sehr komplexe Netzstrukturen bilden, die sich in der Regel in hierarchische Netzebenen gliedern. Auf dem Gebiet der Datenkommunikation ist das sog. *Internet*, ein weltumspannendes Kommunikationsnetz mit mittlerweile über 2 Millionen angeschlossenen Endgeräten, ein eindrucksvolles Beispiel dafür [103].

2.1.3 Vermittlungsverfahren

Das Vermittlungsverfahren bestimmt, wie Nachrichtenblöcke durch das Kommunikationsnetz transportiert werden. Grundsätzlich werden zwei Vermittlungsverfahren unterschieden [99]:

- Bei der Durchschaltevermittlung (*Circuit Switching*, CS) wird einer Verbindung ein durchgeschalteter Kanal mit fester Übertragungsrate zur Verfügung gestellt, der exklusiv genutzt werden kann. Die Information wird dabei im Kanalmultiplex übertragen.
- Bei der Paketvermittlung (*Packet Switching*, PS) werden die zu übertragenden Nachrichten in Datenblöcke (Pakete) aufgeteilt und mit Adreßinformation versehen durch das Netz an die Zieladresse übermittelt. Die zugehörige Multiplextechnik bezeichnet man als Adreßmultiplex. Die Übermittlung der Pakete erfolgt abschnittsweise zwischen den Netzknoten. Ist ein Übertragungsabschnitt bereits belegt, werden die Pakete in den Netzknoten zwischengespeichert. Im Gegensatz zur Durchschaltevermittlung kann eine bessere Nutzung des Übertragungsmediums erzielt werden, da die Datenblöcke verschiedener Kommunikationsbeziehungen beliebig gemischt werden können.

Paketvermittlung kann auf zwei Arten realisiert werden. Bei der verbindungslosen Kommunikation werden einzelne Datenblöcke (Datagramme) mit der vollständigen Zieladresse von Netzknoten zu Netzknoten weitervermittelt, ohne daß dafür Betriebsmittel reserviert werden. Die Übermittlung der Datagramme erfolgt völlig unabhängig voneinander, so daß Datagramme mit der gleichen Zieladresse auf durchaus unterschiedlichen Wegen ihr Ziel erreichen können.

Bei der verbindungsorientierten Kommunikation wird analog zur Durchschaltevermittlung eine virtuelle Verbindung, die aus einer Verkettung logischer Kanäle besteht, zwischen den Endteilnehmern bereitgestellt. Eine virtuelle Verbindung wird während einer Verbindungsaufbauphase, die dem eigentlichen Datentransfer vorausgeht, hergestellt. Gleichzeitig werden der virtuellen Verbindung Betriebsmittel zugewiesen. Die Vermittlung der Pakete in den Netzknoten erfolgt dann während der Datentransferphase über die abschnittsweise Zuordnung logischer Kanalnummern. Im Gegensatz zum Datagrammbetrieb kann bei einer virtuellen Verbindung die Paketreihenfolge garantiert werden. Zusätzlich sind Sicherheitsmaßnahmen wie abschnittsweise Flußkontrolle und Fehlerbehandlung als Protokollfunktionen möglich.

LANs basieren auf der Paketvermittlungstechnik und bieten lediglich einen verbindungslosen Dienst an. MANs hingegen unterstützen hybride Vermittlungsverfahren, die die Integration von Durchschalte- und Paketvermittlung mittels synchroner Zeitmultiplextechnik ermöglichen. Dabei stellt ein sich zyklisch wiederholender Pulsrahmen konstanter Dauer gleichgroße Zeitschlitze bereit, die wahlweise mit paket- oder durchschaltevermitteltem Verkehr belegt werden können. Die Zyklusdauer wird in der Regel so gewählt, daß die CS-Kanäle digitalisierte Sprache übermitteln können.

2.1.4 Schichtenmodell

Die Verknüpfung von Computersystemen unterschiedlicher Hersteller und mit unterschiedlichen Betriebssystemen durch ein Kommunikationsnetz macht die Standardisierung von Schnittstellen und Funktionen, die zum Datenaustausch benötigt werden, unumgänglich. Die Internationale Organisation für Standardisierung (ISO) hat mit der Definition des Basisreferenzmodells für die Verbindung offener Systeme (*Open Systems Interconnection*, OSI) Randbedingungen für die Entwicklung von Diensten und Protokollen festgelegt [77, 116], die seit 1984 als internationaler Standard (IS) vorliegen.

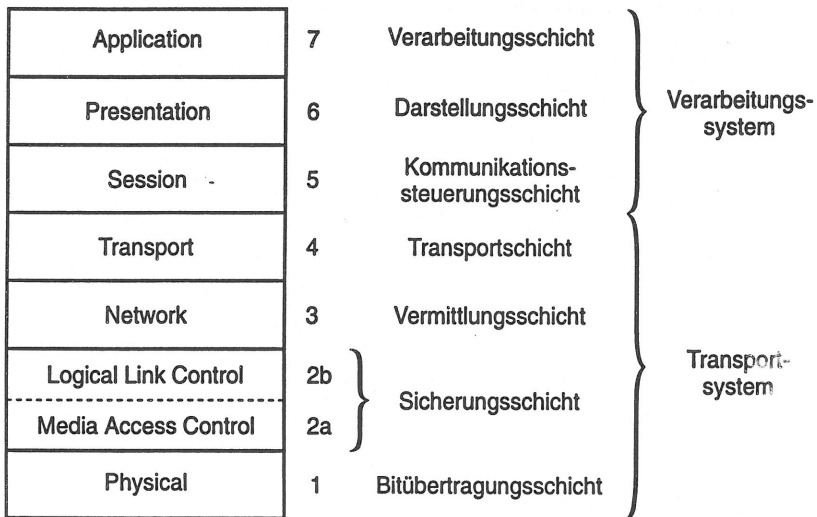


Bild 2.1: ISO-Basisreferenzmodell

Das ISO Basisreferenzmodell (Bild 2.1) beruht auf dem Prinzip der Schichtung, das die Komplexität eines vollständigen Kommunikationssystems in einfachere funktionale Einheiten aufteilt und von den unteren Schichten nach oben zu einer funktionalen Anreicherung führt. Jede Schicht ist für sich unabhängig und bietet der darüberliegenden eine Dienstschnittstelle an, über die spezielle Datenblöcke, sog. Dienstprimitive (*Service Data Unit*, SDU), zwischen den Schichten ausgetauscht werden. Diese Schnittstelle wird auch als Dienstzugangspunkt (*Service Access Point*, SAP) bezeichnet. Die Kommunikation zwischen zwei Instanzen der gleichen Schicht unterschiedlicher Systeme wird über ein Protokoll abgewickelt, mit dem Protokolldateneinheiten (*Protocol Data Unit*, PDU) ausgetauscht werden. PDUs bestehen aus Protokollsteuerinformation und den Nutzdaten, die in der jeweiligen Schicht aus den Dienstprimitiven extrahiert werden. Pakete werden somit beim Sender von Schicht zu Schicht,

beginnend mit der Verarbeitungsschicht, zunehmend mit Protokollsteuerinformation versehen, die beim Empfänger dann in umgekehrter Reihenfolge wieder entfernt wird.

Die Schichten des Basisreferenzmodells können in zwei Gruppen unterteilt werden, die unterschiedliche Aufgaben besitzen:

- Netzorientierte Schichten (Schichten 1 - 4, Transportsystem) mit Funktionen zur Einrichtung einer Verbindung und für den sicheren Datenaustausch zwischen Endsystemen.
- Anwendungsorientierte Schichten (Schichten 5 - 7, Verarbeitungssystem) zur Unterstützung von Anwendungsprozessen.

Die Bitübertragungsschicht definiert die mechanischen, elektrischen und optischen Eigenschaften des Mediums sowie die Bitdarstellung auf dem Medium. Die Sicherungsschicht stellt Funktionen für die Übermittlung von Datenblöcken, die Synchronisierung und die Fehlererkennung auf einem Übertragungsabschnitt bereit. Die Sicherungsschicht wird bei LANs und MANs in zwei Teilschichten, die Medienzugriffssteuerung (*Media Access Control*, MAC) (s. Abschnitt 2.1.5) und die eigentliche Sicherungsschicht (*Logical Link Control*, LLC), unterteilt, um eine Entkopplung zwischen der stark technologieabhängigen MAC-Teilschicht und den höheren Schichten zu erzielen. Durch die Vermittlungsschicht können Stationen über mehrere Teilnetze miteinander verbunden werden. Ihre Aufgaben sind in erster Linie die Verkehrslenkung, die Überlastabwehr und die Auswertung von Netzadressen. Die Transportschicht besitzt erstmalig Ende-zu-Ende-Signifikanz, d.h. diese Schicht stellt einen zuverlässigen Transportdienst zwischen Endsystemen bereit.

Das Verarbeitungssystem fügt dem Transportsystem Funktionen zur Dialogsteuerung (Sitzungsschicht), Datendarstellung (Darstellungsschicht) und eine ganze Reihe von Dienstelementen in der Verarbeitungsschicht, die spezifische Anwendungen wie etwa Datenbanken unterstützen, hinzu.

2.1.5 Medienzugriffsverfahren

Die Leistungsfähigkeit von LANs und MANs wird neben der Kapazität des Übertragungsmediums vor allem durch das Medienzugriffsprotokoll, das mit einem dezentralen Steuerungsmechanismus arbeitet, bestimmt. Dessen Aufgabe ist es, die vorhandene Übertragungskapazität auf alle angeschlossenen Stationen aufzuteilen. Eine Fülle von Medienzugriffsverfahren sind seit Beginn der siebziger Jahre entworfen und untersucht worden. Im folgenden sollen nur einige ausgesuchte Medienzugriffsverfahren skizziert werden, um ihre prinzipiellen Eigenschaften zu verdeutlichen. Für eine detaillierte Übersicht sei auf die Literatur verwiesen [7, 8].

Wichtige Vertreter, die vor allem für die private Datenkommunikation große Bedeutung erlangt haben, sind CSMA/CD (*Carrier Sense Multiple Access with Collision Detection*) [72], *Token Bus* [73], *Token Ring* [74] und das Hochgeschwindigkeits-LAN FDDI [4]. CSMA/CD ist als Produkt mit dem Namen *Ethernet* das derzeit wohl am weitesten verbreitete LAN. Die Datenübertragungsraten von LANs liegen typischerweise im Bereich von 1 bis 16 Mbit/s. Im Bereich öffentlicher Netzbetreiber hat sich das Medienzugriffsverfahren DQDB (*Distributed Queue Dual Bus*) durchgesetzt. Die Eigenschaften dieses Verfahrens werden an anderer Stelle beschrieben (s. Abschnitt 2.2).

Dezentrale Medienzugriffsverfahren werden in Wettbewerbs- und Reservierungsverfahren unterschieden. CSMA/CD, das ein Bussystem als Übertragungsmedium verwendet, basiert auf einem Wettbewerbsverfahren. Die sendebereiten Stationen hören den Übertragungskanal permanent ab und beginnen zu senden, sobald das Medium frei und damit ein eventuell laufender Übertragungsvorgang beendet ist. Durch die Signallaufzeit auf dem Medium kann es jedoch zu Kollisionen kommen. Tritt dieser Fall ein, beenden die Stationen den Sendevorgang und verschieben einen erneuten Versuch um ein zufällig ermitteltes Zeitintervall in die Zukunft.

Token Bus, *Token Ring* und FDDI benutzen ein *Token* (implizite Sendeberechtigung) zur Reservierung von Übertragungsbandbreite. Das *Token* wird von Station zu Station entlang eines logischen Rings weitergereicht, so daß jeweils nur eine Station sendeberechtigt ist. Diese Station behält das *Token* während eines, durch einen Zeitgeber begrenzten Zeitintervalls, das zur Paketübertragung genutzt wird. Erreicht der Paketkopf die sendende Station erneut, wird ein neues *Token* generiert und an die nächste Station weitergereicht (*Single Token-Betrieb*). Dieses Prinzip kann dadurch erweitert werden, daß bereits an das Ende der abgesendeten Pakete ein *Token* angehängt wird (*Multiple Token-Betrieb*). Das führt zu einer effizienteren Ausnutzung des Mediums, wenn die Ringverzögerung im Vergleich zur Paketbedienzeit deutlich größer ist. Mit allen Verfahren, die auf dem *Token-Prinzip* basieren, werden Kollisionen vermieden und die Übermittlungszeiten auch im Hochlastfall auf einen Maximalwert begrenzt.

Beim *Token Bus*, der auf einer Bustopologie realisiert wird, muß im Betrieb zusätzlich ein logischer Ring aufrecht erhalten werden, was einen aufwendigeren Steuerungsmechanismus im Vergleich zum *Token Ring* bedingt. Zusätzlich weist das *Token Bus*-Verfahren eine synchrone Verkehrsklasse mit festen Sendezeitverzögerungen auf, um Verkehr mit Echtzeitanforderungen zu unterstützen. Bei FDDI wird das Konzept so erweitert, daß auch bei höheren Übertragungsraten (100 Mbit/s) und größerer physikalischer Ringausdehnung (bis zu 100 Kilometern) in jedem Lastbereich der Maximaldurchsatz nicht zurückgeht. Hierzu wurde das zeitgesteuerte *Token*-Protokoll, wie es beim *Token Bus* verwendet wird, an die Anforderungen bei FDDI angepaßt (s. auch [152]).

2.1.6 Adressierungskonzepte

In jedem Kommunikationsnetz spielt die Frage, wie Netzressourcen bzw. Endteilnehmer lokalisiert werden können, eine wichtige Rolle. Diese Frage ist eng verknüpft mit der Namensgebung, dem Adressierungskonzept und der Verkehrslenkung [129]. Der Name bezeichnet eine bestimmte Ressource, die Adresse gibt Aufschluß über den Ort, an dem sich die Ressource befindet und die Verkehrslenkung entscheidet, wie man dorthin gelangt. Adressen werden bei OSI über einen Dienstzugangspunkt identifiziert. Ihre Struktur kann flach oder hierarchisch aufgeteilt sein.

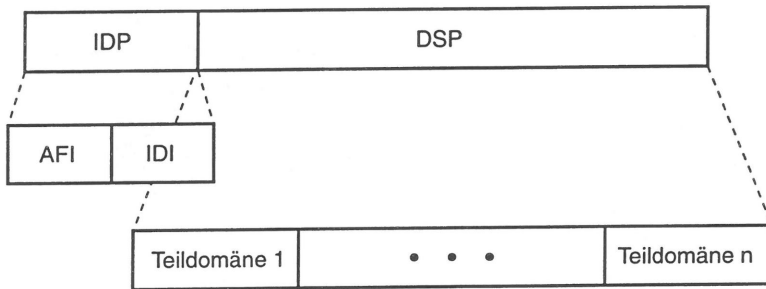


Bild 2.2: Struktur der OSI-Netzadresse

Auf der Vermittlungsschicht werden global gültige, hierarchisch aufgeteilte Adressen verwendet, da hier die Endsysteme eindeutig identifiziert werden können und eine einfache Abbildung auf teilnetzspezifische Adressen möglich ist. Teilnetze im Sinne von OSI umfassen LANs, MANs oder WANs mit unterschiedlicher Netztechnik. Die Adreßstruktur der OSI-Netzadresse zeigt Bild 2.2. Sie besteht aus einem *Initial Domain Part* (IDP) und einem *Domain Specific Part* (DSP). Der IDP wiederum umfaßt den *Authority and Format Identifier* (AFI) und den *Initial Domain Identifier* (IDI). Der AFI bezeichnet die registrierende Organisation (z.B. ISO), das Format des IDI und die abstrakte Syntax des DSP, d.h. ob dieser Wert binär oder dezimal codiert wird. Der IDI bezeichnet, welcher Adreßraum für den DSP gewählt wird und welche Organisation dafür zuständig ist.

Zur sinnvollen Aufteilung des Netzes kann der DSP in sog. Adreßdomänen (Verwaltungsbereiche) aufgeteilt und in Teildomänen hierarchisch verfeinert werden [78, 140]. Domänen auf einer Stufe dürfen sich nicht überlappen; Teildomänen müssen vollständig in einer übergeordneten Domäne enthalten sein. Durch diese Aufteilung des globalen Adreßraums kann bei Wahl von teilnetzunabhängigen Adressen, die beispielsweise nach einer Organisationsstruktur unterteilt werden, die Netzplanung und -administration sowie die Verkehrslenkung vereinfacht werden [45]. Dafür sind zwei Kriterien ausschlaggebend: Zum einen sollten

Netzadressen möglichst teilnetzunabhängig sein, um auch bei Topologieänderungen gültig zu bleiben. Zum anderen sollten Netzadressen so strukturiert werden, daß eine Gruppe von Adressen durch einen gemeinsamen Präfix identifiziert und Pakete, die für diese Adressen bestimmt sind, über den gleichen Übertragungsabschnitt zum nächsten Vermittlungsknoten gelenkt werden (hierarchische Verkehrslenkung). Die technischen und organisatorischen Randbedingungen sind dafür von ISO festgelegt [79].

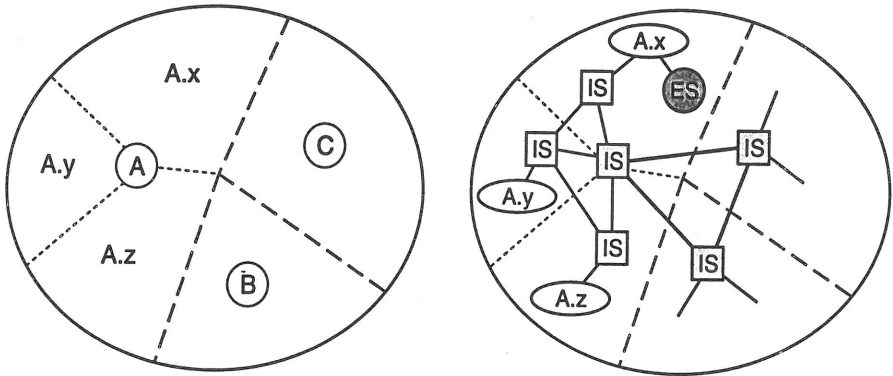


Bild 2.3: Beispiel für einen Adreßraum und die zugehörige Netzstruktur

Ein Beispiel, in Anlehnung an den ISO-Standard 10589 [80], zeigt Bild 2.3. Der Adreßraum sei in drei Verwaltungsbereiche mit jeweils einem Präfix („A“, „B“ oder „C“) aufgeteilt; der Verwaltungsbereich „A“ wiederum in drei weitere Teilbereiche. Die verwendeten Buchstaben stehen hier stellvertretend für Ziffern- oder Zeichenfolgen. Im rechten Teil des Bildes ist eine mögliche, korrespondierende Netzstruktur dargestellt. Ein Endsystem (ES), das an ein Teilnetz angeschlossen ist, wird durch den Adreßpräfix des Teilnetzes und eine Stationskennung identifiziert. Auf den höheren Ebenen werden Pakete über Router, auch als *Intermediate System* (IS) bezeichnet, durch die hierarchische Auswertung der Adressen vermittelt. Wird beispielsweise ein Paket von einem Endsystem, das an das Teilnetz „A.x“ angeschlossen ist, an ein beliebiges Endsystem in den Bereichen „B“ oder „C“ gesendet, so wertet der Router, der ebenfalls mit „A.x“ verbunden ist, nur den Anfang der Adresse aus und leitet das Paket an einen Router auf höherer Ebene (im Bild grau unterlegt) weiter, bis es zu einem Router im Bereich „B“ bzw. „C“ gelangt. Gleichzeitig sind also mit der Adreßstruktur auch die zulässigen Wege zur Weiterleitung von Paketen festgelegt.

Das Adressierungsschema nach IEEE [70]; das für LAN/MAN-Teilnetze vorgesehen ist, besitzt eine flache Adreßstruktur, die durch gerätespezifische Adressen (MAC-Adressen) realisiert wird. Diese Adressen werden von der Medienzugriffsteuerung ausgewertet. Bei einer flachen Adreßstruktur ist eine Bestimmung der Stationsposition aus der Adresse nicht

möglich. MAC-Adressen können wahlweise 2 oder 6 Byte umfassen; abhängig davon, ob ihre Gültigkeit lokal beschränkt ist oder nicht.

2.1.7 Netzkopplung

Für den Aufbau eines großen Netzverbundes ist die Aufteilung in Teilnetze nach administrativen, ökonomischen oder technischen Gesichtspunkten notwendig, um den Netzbetrieb sinnvoll zu gestalten. Die Kopplung von Teilnetzen wird durch spezielle Netzkoppeleinheiten (*Interworking Units*, IWU) erbracht. Innerhalb einer IWU bestimmt die Verkehrslenkung (*Routing*) aus einer gegebenen Adresse den Weg zu einem bestimmten System, d.h. in der IWU wird die Zieladresse einem bestimmten, teilnetzspezifischen Dienstzugangspunkt zugeordnet. Die eigentliche Protokollumsetzung (*Relaying*) wird getrennt davon betrachtet, da in manchen Fällen Protokolle bzw. Dienste der zu koppelnden Teilnetze verschieden sind. Ein Beispiel, das die Zusammenhänge zwischen Adressierung, Netzkopplung und Verkehrslenkung bei der Kopplung von lokalen und Weitverkehrsnetzen verdeutlicht, findet sich in [141].

Netzkoppeleinheiten lassen sich nach der Schicht des ISO Referenzmodells, in der die Kopplung erfolgt, klassifizieren:

- Schicht 1: *Repeater*.
- Schicht 2: *Bridge*.
- Schicht 3: *Router*.
- Schicht 4 oder höher: *Gateway*.

Repeater dienen zur Signalregeneration und können damit zur Erweiterung der physikalischen Topologie durch die Verbindung getrennter Netzsegmente eingesetzt werden. Sie werden vielfach in CSMA/CD-Netzen verwendet.

Auf Schicht 2 findet die Kopplung über *Bridges* statt. Sie ermöglichen die Aufteilung des Netzes zur Begrenzung der Last und der Anzahl der Stationen. Die eigentliche Netzkopplung findet in der Regel auf der MAC-Teilschicht statt; die Kopplung auf der LLC-Teilschicht ist prinzipiell ebenfalls möglich. Zwei Grundformen, die sich im wesentlichen durch die Art der Verkehrslenkung unterscheiden, haben in lokalen Netzen weite Verbreitung erfahren: *Spanning Tree Bridge* und *Source Routing Bridge*.

Bei *Spanning Tree Bridges* [70, 130] werden Pakete entlang eines logischen Baums (*Spanning Tree*), der automatisch durch ein verteiltes Protokoll zwischen allen *Bridges* aus der vorhandenen physikalischen Topologie gebildet wird, weitergeleitet. Durch die logische Topologie ist gewährleistet, daß Pakete nicht dupliziert werden können. Erreicht ein Paket die *Bridge*,

wird geprüft, ob sich die Zieladresse im selben LAN befindet und somit das Paket nicht weitergesendet werden soll. Ist das nicht der Fall, wird das Paket auf allen Ausgangsleitungen ausgesendet, sofern nicht die Zieladresse und eine zugeordnete Ausgangsleitung bereits im Adreßspeicher der *Bridge* vorhanden ist. Der Adreßspeicher wird durch das „Lernen“ von Adressen gefüllt, wobei die Ursprungsadresse eines Pakets, das eine *Bridge* passiert, zusammen mit der zugehörigen Eingangsleitung gespeichert wird. Durch die eindeutig festgelegte Topologie ist hier alternative Verkehrslenkung nicht möglich, so daß vorhandene Ressourcen nur schlecht ausgenutzt werden. In der Praxis ist die Administration solcher Netze durch die automatische Konfiguration des Gesamtsystems jedoch relativ einfach.

Source Routing Bridges verwenden dagegen ein Verkehrslenkungsverfahren, das den Weg vom Sender bis zum Empfänger vollständig bestimmt [33, 130]. Das Verfahren wurde ergänzend zu IEEE 802.5 (*Token Ring*) standardisiert. Die Auswahl der Pfade wird hier vollständig in den Stationen getroffen, die einen oder mehrere Pfade zu den Zielstationen speichern können. Ein Pfad wird durch eine Sequenz von *Token Ring*- und *Bridge*-Adressen beschrieben, die im Paketkopf enthalten ist. Um bessere oder neue Pfade zu erkunden, sendet eine Station spezielle Pakete aus, die vervielfältigt werden, und auf unterschiedlichen Wegen zum Ziel gelangen können. Aus den zurückgesendeten Antwortpaketen kann die Station dann den besten Pfad auswählen. Das wichtigste Merkmal dieses Verfahrens ist die Verlagerung der Verkehrslenkung aus den Netzkoppeleinheiten in die Endsysteme, die für das Auffinden, Unterhalten und Optimieren von geeigneten Wegen selbst verantwortlich sind.

Die mit der Kopplung auf Schicht 2 verbundenen Verkehrslenkungsverfahren sind jedoch hinsichtlich ihrer Anwendung in einem größeren Netzverbund, optimaler Auslastung von Betriebsmitteln und Zuverlässigkeit nur beschränkt verwendungsfähig. Nach ISO sollten die Verkehrslenkung eine Funktion der Vermittlungsschicht sein und *Router* zur Kopplung einer großen Zahl von Teilnetzen eingesetzt werden [79]. *Router* sind in der Lage, globale Netzadressen zur Verkehrslenkung zu verwenden. Sie können in der Regel dynamisch auf unterschiedliche Lastsituationen und Leitungsausfälle reagieren. Das wird durch ein Meldungsprotokoll ermöglicht, mit dem *Router* ihre Verkehrslenkungstabellen zyklisch untereinander oder mit einem Zentralrechner austauschen. Dadurch verschafft sich jeder *Router* ein aktuelles Bild über den Netzzustand. In Abschnitt 2.1.8 werden die wichtigsten Verfahren zur Verkehrslenkung beschrieben.

Die Netzkopplung auf der Transportschicht oder auf der Verarbeitungsschicht wird von einem *Gateway* durchgeführt. *Gateways* auf der Verarbeitungsschicht werden vor allem zur Kopplung herstellerspezifischer und standardisierter Protokollarchitekturen eingesetzt, zwischen denen sonst prinzipiell kein Datenaustausch möglich wäre. Dabei geht jedoch die Ende-zu-Ende-Signifikanz verloren, so daß ein *Gateway* die Endsysteme eines Teilnetzes gegenüber dem anderen Teilnetz verbirgt.

2.1.8 Verkehrslenkung

In Paketvermittlungsnetzen hat das verwendete Verkehrslenkungsverfahren signifikante Auswirkungen auf die Leistungsfähigkeit des Netzes. Ziel ist es, die günstigsten Wege¹ in einem Netz zu bestimmen, die zur Übermittlung der Pakete von einer Verkehrsquelle zur jeweiligen Verkehrssenke dienen. Dabei sollten andere Kommunikationsbeziehungen nicht beeinträchtigt und ein hoher Datendurchsatz bei möglichst geringen Transferzeiten erzielt werden.

Verkehrslenkungsverfahren können unter verschiedenen Gesichtspunkten wie

- zentrale oder dezentrale Steuerung,
- Auswahl der Wegeabschnitte (fest oder alternativ),
- Dynamik der Anpassung an einen veränderten Netzzustand (statische bzw. dynamische Verkehrslenkung),
- Orientierung an der Netzstruktur (hierarchische bzw. nichthierarchische Verkehrslenkung)

klassifiziert werden (vgl. auch [79, 143]).

Bei zentraler Steuerung melden die beteiligten Vermittlungsrechner (*Router*) ihren lokalen Netzzustand, z. B. welche Übertragungsabschnitte ausgefallen sind oder die aktuelle Last auf ihren Ausgangsleitungen, in regelmäßigen Zeitabständen an einen zentralen Steuerrechner. Aus diesen Statusinformationen kann der Steuerrechner die optimalen Wege bestimmen und diese an die Vermittlungsrechner in Form von Verkehrslenkungstabellen zurücksenden.

Bei dezentraler Steuerung (verteilte Verkehrslenkung) tauschen die Vermittlungsrechner die Informationen über den Netzzustand untereinander aus und berechnen daraus die günstigsten Wege.

Prinzipiell ist in beiden Fällen adaptive Verkehrslenkung, also die Anpassung an veränderte Netzzustände, möglich. Praktisch sind jedoch bei zentraler Steuerung die Verkehrslenkungsalgorithmen in ihrer Wirkung beschränkt, weil je nach Verzögerungszeit, die bis zum Eintreffen aller Meldungen entsteht, die Information beim Zentralrechner veraltet sein kann. Im Gegensatz dazu sind verteilte, adaptive Verkehrslenkungsverfahren in der Lage, schnell und robust auf Netzausfälle oder Lastschwankungen zu reagieren. Eine optimale Wegebestimmung kann aber nicht garantiert werden, da die Wegeberechnung auf Grund von Echtzeitanforderungen sehr schnell und damit approximativ ausgeführt wird. Zu beachten ist darüber hinaus, daß bei zentraler Steuerung der Ausfall des Zentralrechners den Netzbetrieb zum Erliegen bringt. Bei verteilter Steuerung ist der Ausfall eines Netzknotens meist weniger gravierend.

¹Im Zusammenhang mit der Verkehrslenkung wird der Begriff „Weg“ synonym mit „Pfad“ verwendet.

Der wohl bekannteste Algorithmus zur Wegbestimmung ist das Verfahren der kürzesten Wege bzw. Pfade (*Shortest Path Routing*) [13]. Dabei wird jedem Übertragungsabschnitt eine Metrik zugewiesen, d.h. eine Zahl, die eine bestimmte Größe repräsentiert. Die Metrik kann etwa von der Kapazität eines Übertragungsabschnitts, von der Last auf dem Übertragungsabschnitt oder von der Leitungslänge abhängig sein. Die Länge eines Pfades ergibt sich aus der Summe aller Metriken der Übertragungsabschnitte entlang des Pfades. Beim *Shortest Path Routing* wird nun versucht, Pakete jeweils von einer Verkehrsquelle zur -senke über einen Pfad minimaler Länge zu leiten, der zuvor aus der Netztopologie ermittelt wurde. Ein *Shortest-Path-Algorithmus* bildet beispielsweise die Basis für das Verkehrslenkungsverfahren im ARPANET, einem Paketvermittlungsnetz in den USA [13].

2.2 Regionale Hochgeschwindigkeitsnetze auf der Basis von DQDB

2.2.1 Übersicht

Aus der Erwartung neuer Applikationen mit hohen Bitratenanforderungen und dem Trend, diensteintegrierende Kommunikationssysteme bereitzustellen, wurde seit 1988 im Standardisierungsgremium IEEE 802.6 die Entwicklung von *Metropolitan Area Networks* (MANs) auf der Basis der LAN-Technik vorangetrieben [31, 93, 117]. Demnach weisen MANs folgende Funktionsmerkmale auf:

- Verwendung optischer Übertragungssysteme.
- Nutzung eines gemeinsamen, breitbandigen Übertragungsmediums mit Übertragungsraten typischerweise im Bereich von 100 Mbit/s.
- Netzausdehnung bis zu einigen hundert Kilometern.
- Hybride Vermittlungstechnik: Integration von Durchschaltvermittlung und Paketvermittlung.
- Realisierung des paketvermittelnden Dienstes durch Segmentierung von längeren Paketen in einheitliche Segmente fester Länge und ein verteiltes Medienzugriffsverfahren.
- Realisierung des durchschaltvermittelnden Dienstes über eine Pulsrahmenstruktur und zentrale Reservierung von Zeitschlitzten für eine Verbindung.
- Einrichtung und Betrieb in der Regel durch öffentlichen Netzbetreiber.
- Funktionen zum Abrechnungsmanagement und zur Netzsicherheit.

Wichtigste Einsatzgebiete sind die Kopplung lokaler Netze im Bereich der Bürokommunikation oder der rechnerintegrierten Fertigungsautomatisierung sowie die Ankopplung von

LANs an Großrechnersysteme. Ein weiteres Anwendungsgebiet ist die Übermittlung hochauflösender Bilder, wie sie beispielsweise in der medizinischen Diagnostik benötigt werden [5].

Für die Realisierung von MANs wurden verschiedene Verfahren vorgeschlagen. Neben *Distributed Queue Dual Bus* (DQDB), das schließlich als Norm durch das Standardisierungsgremium bei IEEE ausgewählt wurde, seien beispielhaft FDDI-II [136], ORWELL [46] und ATMR [76] genannt. Einige dieser Systeme sind jedoch über das Stadium eines Prototypen nicht hinausgekommen, da sie sich in der Standardisierung nicht durchsetzen konnten.

2.2.2 DQDB

2.2.2.1 Topologie und Übermittlungstechnik

1990 wurde die Norm IEEE 802.6 verabschiedet, die die Topologie, das verteilte Medienzugriffsprotokoll und die realisierten Dienste für ein MAN-Teilnetz auf der Basis von DQDB festlegt [75]. DQDB wurde ursprünglich unter dem Namen QPSX (*Queued Packet and Synchronous Circuit Exchange*) als Standardisierungsvorschlag eingereicht (vgl. [124]).

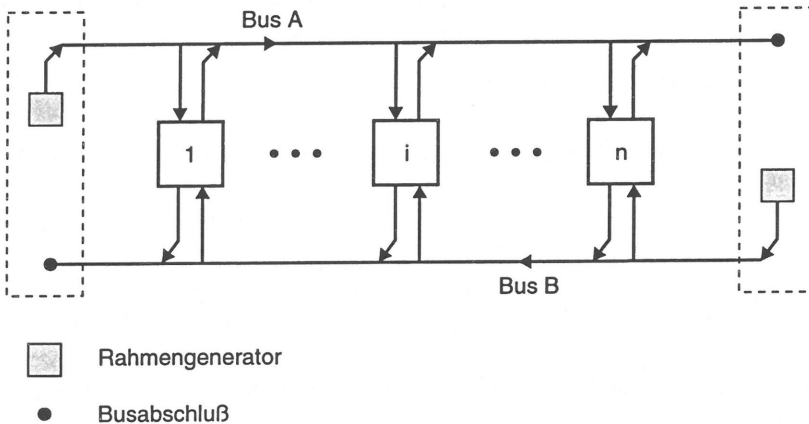


Bild 2.4: DQDB-Topologie

DQDB basiert auf einem Dualbussystem, das den Informationsfluß in jeweils gegenläufiger Richtung zwischen den angeschlossenen Stationen transportiert (Bild 2.4). Die beiden Kopfstationen erzeugen alle $125 \mu s$ einen Pulsrahmen mit einer Zeitschlitzstruktur, die die Integration von asynchronem (PS-) und isochronem (CS-) Verkehr ermöglicht. Die einzelnen Stationen können passiv oder aktiv an das Bussystem angekoppelt werden; in der Regel

wird ein solches System jedoch mit Hilfe von Punkt-zu-Punkt-Verbindungen aktiv ausgelegt. Die Anschlußeinheit wird mit Hilfe einer logischen ODER-Verknüpfung so realisiert, daß prinzipiell keine Daten vom Bus gelöscht werden können.

Zur Erhöhung der Verfügbarkeit kann ein DQDB-Teilnetz als sog. *Looped Bus* betrieben werden, wobei die Funktionen der beiden Kopfstationen (Station 1 bzw. n sowie jeweils Rahmengenerator und Busabschluß) in einer Station zusammengelegt werden, und der Bus zu einem Ring geschlossen wird. Der Informationstransport erfolgt nach wie vor entlang Bus A bzw. Bus B. Fällt ein Übertragungsabschnitt zwischen zwei Stationen aus, kann das Bussystem umkonfiguriert und der Netzbetrieb weiterhin aufrechterhalten werden. Dazu ist es allerdings notwendig, daß jede Station die Funktion von Kopfstationen übernehmen und die Rahmenstruktur erzeugen kann.

2.2.2.2 Stationsmodell

DQDB umfaßt neben der Bitübertragungsschicht und der Medienzugriffsteuerung (MAC-Teilschicht) des ISO-Basisreferenzmodells zusätzliche Funktionen für isochrone Dienste, die zusammen mit der MAC-Teilschicht als DQDB-Schicht bezeichnet werden. Das Stationsmodell, das den eigentlichen Rahmen der Standardisierung umfaßt, definiert dafür Funktionsblöcke, die den beiden Schichten zugeordnet werden können (Bild 2.5). Das Übertragungssystem ist nicht Gegenstand der DQDB-Standardisierung.

Die Bitübertragungsschicht (Schicht 1) stellt im wesentlichen Funktionen zum Schichtmanagement und zur Anpassung an unterschiedliche Übertragungssysteme bereit. Dazu zählen die Anpassung an die jeweiligen Datenübertragungsraten und die Abbildung der Zeitschlitzstruktur auf die Rahmenstruktur des Übertragungssystems. Als Übertragungssysteme sind standardisierte Systeme, wie sie in öffentlichen Netzen eingesetzt werden, mit Übertragungsraten zwischen 2,048 und 622,080 Mbit/s vorgesehen [93].

Die DQDB-Schicht (Schicht 2) enthält neben den gemeinsamen Funktionen zum Zugriff auf die Zeitschlitzstruktur Funktionsblöcke für das *Distributed-Queueing*-Protokoll (DQ-Protokoll), Funktionen zur Medienzugriffsteuerung von asynchronem Verkehr (mit *Queued-arbitrated* bezeichnet) und zur Vermittlung von isochronem Verkehr (*Pre-arbitrated*). Darauf aufbauend werden Anpassungsfunktionen, die für den verbindungslosen Dienst Funktionen zur Segmentierung und zum Zusammenfügen von Paketen bereitstellen, und geeignete Dienstschnittstellen definiert. Auch die DQDB-Schicht stellt mit den Funktionen des Schichtmanagements einen direkten Zugang für die Netzüberwachung bereit. Es sei angemerkt, daß die Standardisierung von verbindungsorientierten, asynchronen Diensten noch nicht abgeschlossen ist. Im nächsten Abschnitt wird nun genauer auf die Funktionen zur Medienzugriffsteuerung von asynchronem Verkehr, das sog. *Distributed-Queue*-Protokoll, eingegangen.

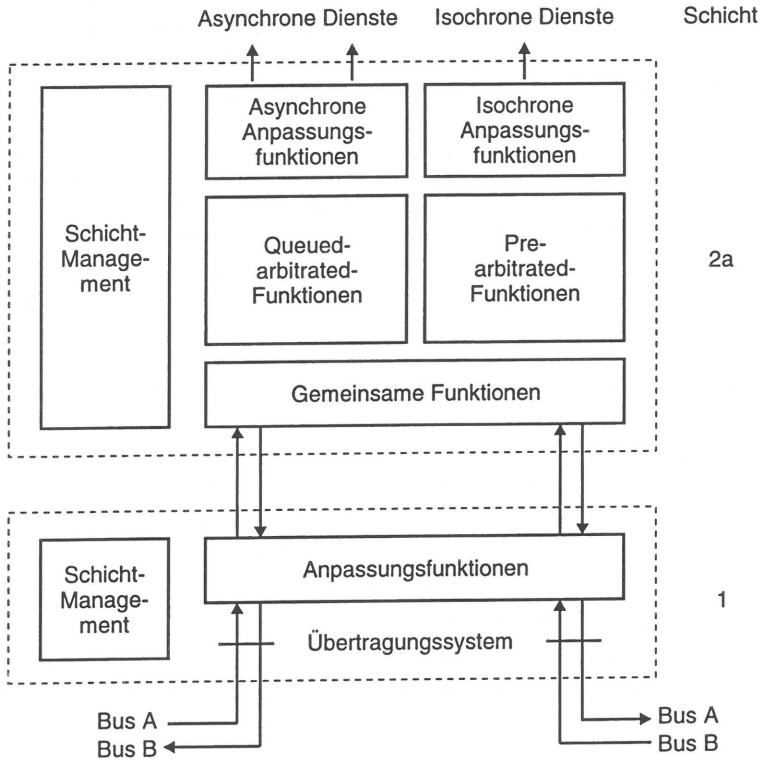


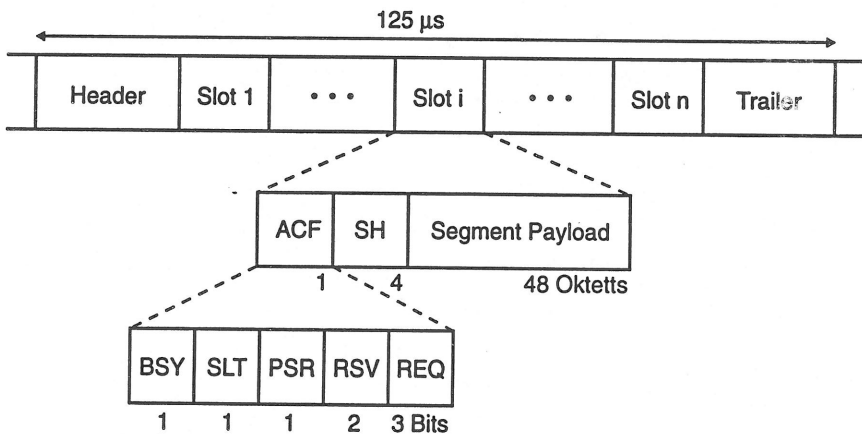
Bild 2.5: DQDB-Stationmodell

2.2.2.3 DQ-Protokoll

Der mit dem DQ-Protokoll realisierte verbindungslose Dienst unterstützt drei Prioritätsklassen. In jeder Klasse erfolgt die Übermittlung von Segmenten in den asynchron belegten Zeitschlitzen. Die Nutzinformation wird dabei in das Feld Segment Payload eingefügt (Bild 2.6). Als Protokollsteuerinformation stehen innerhalb des *Access Control Fields* (ACF) das *Busy-Bit* (BSY) und drei *Request-Bits* (REQ) zur Verfügung. Das Feld *Slot Type* (SLT) dient zur Unterscheidung, ob CS- oder PS-Verkehr transportiert wird. Der *Segment Header* (SH) enthält im wesentlichen Felder zur Identifizierung einer Verbindung und zur Kennzeichnung von aufeinanderfolgenden und zusammengehörigen Segmenten. Im Falle von isochronem Verkehr wird das Feld *Segment Payload* durch eine Struktur ersetzt, in der jeder CS-Verbindung ein Oktett pro Rahmen zugewiesen wird.

Das Prinzip des DQ-Protokolls beruht auf einer Warteschlange, die auf alle Stationen verteilt ist und für jeden Bus durch Reservierungsanforderungen aufgebaut wird. Diese werden über den jeweils entgegengesetzten Bus an alle folgenden Stationen gesendet. Idealerweise entsteht so eine Warteschlange nach dem FIFO-Prinzip (*First In, First Out*), die den Zugriff auf freie Zeitschlitz konfliktfrei regelt.

Für die Protokollabwicklung stehen in jeder Station pro Bus und Prioritätsklasse zwei Zähler zur Verfügung: *Request-Zähler* und *Countdown-Zähler*. Die Protokollbeschreibung beschränkt sich im folgenden auf eine Prioritätsklasse und einen Bus als Senderichtung.



Abkürzungen:

ACF	Access Control Field	PSR	Previous Segment Received
SH	Segment Header	RSV	Reserved
BSY	Busy	REQ	Request
SLT	Slot Type		

Bild 2.6: DQDB-Rahmenstruktur

Im Anfangszustand werden beide Zähler zurückgesetzt. Für jede Anforderung auf dem entgegengesetzt gerichteten Bus (*Request-Bit* hat den Wert „1“) wird der Request-Zähler um Eins erhöht, so daß auch stromabwärts liegende Stationen Zugriff auf den Bus erhalten; für jeden freien Zeitschlitz (*Busy-Bit* hat den Wert „0“), der an einer Station vorbeikommt, wird der *Request-Zähler* um Eins verringert. Besteht nun ein Sendewunsch, wird das zu sendende Segment in die lokale Warteschlange eingereiht und das der Priorität entsprechende *Request-Bit* im nächsten möglichen Zeitschlitz auf dem entgegengesetzt gerichteten Bus gesetzt. Außerdem wird der aktuelle Inhalt des *Request-Zählers* in den *Countdown-Zähler*

kopiert und der *Request*-Zähler auf den Wert Null gesetzt. Für jeden freien Zeitschlitz, der in dieser Phase die Station passiert, wird der *Countdown*-Zähler um Eins verringert. Ist der Zählerstand des *Countdown*-Zählers bei Null angelangt, erhält die Station auf den nächsten freien Zeitschlitz Zugriff und kann senden. Der Zählerstand des *Request*-Zählers ändert sich währenddessen durch die bereits oben beschriebenen Mechanismen.

Das DQ-Protokoll kann im Hochlast- bzw. Überlastfall in Abhängigkeit von der Busposition zur Benachteiligung bestimmter Stationen führen, was vor allem auf die Laufzeiten der Reservierungen (*Requests*) und der freien Zeitschlitz auf dem Medium zurückzuführen ist. Durch die Einführung eines Bandbreiten-Ausgleichsmechanismus, der jeder Station nur einen bestimmten Anteil an der maximal erreichbaren Übertragungsrates zur Verfügung stellt, kann jedoch die Übertragungskapazität gleichmäßiger aufgeteilt werden. Darüber hinaus gibt es Vorschläge, die Übertragungskapazität durch die Mehrfachnutzung von Zeitschlitz zu erhöhen. Das kann mit Hilfe spezieller Stationen erfolgen, die bereits gelesene Zeitschlitz wieder freischalten. Diese Stationen werden *Erase Nodes* genannt. Für eine ausführliche Diskussion dieser Vorschläge sei auf die Literatur verwiesen [56, 66, 120, 137].

2.2.3 Hochgeschwindigkeits-Datendienste

Durch die schnelle Verbreitung lokaler Netze wächst der Bedarf an Datenkommunikationsdiensten im Weitverkehrsbereich rasch an. Insbesondere besteht bei Anwendern der zunehmende Wunsch nach einer flächendeckenden Versorgung mit Hochgeschwindigkeits-Datendiensten, um die, meist als „Kommunikationsinseln“ auftretenden LANs adäquat zu verbinden. Für diesen Zweck wurden daher spezielle Dienste für die Hochgeschwindigkeits-Datenkommunikation spezifiziert: SMDS (*Switched Multi-Megabit Data Service*) [12, 32, 61] in Nordamerika von Bellcore und CBDS (*Connectionless Broadband Data Service*) [43] in Europa durch das Standardisierungsgremium ETSI (*European Telecommunications Standards Institute*).

Ziel ist es, die Kopplung von lokalen Netzen über große Distanzen hinweg mit hohem Durchsatz und geringen Verzögerungszeiten zu ermöglichen. Festgelegt ist dabei lediglich die Dienstschnittstelle zwischen der Teilnehmereinrichtung und dem Netz, das den Dienst bereitstellt. Das dafür verwendete Zugangsprotokoll basiert auf DQDB. Das Netz selbst wird nicht näher spezifiziert; es wird zunächst von einer Realisierung mittels gekoppelter MAN-Teilnetze ausgegangen. Ein solcher Netzverbund wird auch als MAN-Vermittlungssystem (*MAN Switching System*, MSS) bezeichnet.

Im folgenden sollen zunächst der Dienst und das Zugangsprotokoll sowie die Ankopplung von LANs beschrieben werden. Da CBDS die wesentlichen Merkmale von SMDS übernommen hat, beschränkt sich die folgende Beschreibung auf SMDS.

2.2.3.1 Dienstbeschreibung

SMDS war zunächst als verbindungsloser Datendienst für regionale Gebiete ausgelegt. Mittlerweile werden dieser und vergleichbare Dienste (CBDS) auch überregional angeboten. SMDS soll eine Dienstgüte bieten, die mit derjenigen in LANs vergleichbar ist, und eine einfache Anbindung von LANs ermöglichen. Das bedeutet, daß in einem LAN idealerweise fast keine Änderungen vorzunehmen sind und der Anschluß durch eine spezielle Netzkoppel-einheit möglich ist.

Die wesentlichen Merkmale von SMDS sind gekennzeichnet durch die Ausrichtung auf die Datenkommunikation:

- Verbindungslose Paketübermittlung bis zu einer Paketlänge von 9188 Oktetts. Damit sind alle in LANs verwendeten Paketformate abgedeckt.
- Netzadressierung nach dem ISDN-Numerierungsschema E.164 [82], das mit dem Numerierungsschema in Telephonnetzen vergleichbar ist.
- Adressierung einer Gruppe von Teilnehmern mit einer einzigen Adresse (Gruppen-adressierung), analog zu einer *Multicast*-Funktion in LANs.
- Funktionen zur Prüfung, ob die Adresse innerhalb einer Teilnehmergruppe korrekt angegeben wird (*Address Screening*). Dadurch wird die Bereitstellung virtuell privater Netze (*Virtual Private Network*, VPN), die ausschließlich für eine Teilnehmergruppe eingerichtet werden, unterstützt.
- Funktionen zur Prüfung, ob die Quelladresse an der Schnittstelle zwischen Teilnehmer und Netz korrekt angegeben wird (*Address Validation*).
- Definition von Netzzugangsklassen, die einen begrenzten Schutz vor einer Überlastung des MSS bieten. Das wird durch die Überwachung und Beschränkung der mittleren Übertragungsrate über einen definierten Zeitraum erreicht. Kurzfristig kann mit der maximalen Übertragungsrate gesendet werden; wird jedoch ein vorher festgelegter, langfristiger Mittelwert überschritten, so werden zu sendende Pakete abgelehnt. Netzzugangsklassen sind beispielsweise für 4, 10 oder 16 Mbit/s definiert, so daß sie mit den Übertragungsraten der meistverwendeten LANs (*Ethernet*, *Token Bus* und *Token Ring*) kompatibel sind.

Ferner werden Dienstgüteparameter hinsichtlich der maximal zulässigen Ende-zu-Ende-Verzögerungszeit zwischen zwei beliebigen Dienstschnittstellen, der Paketverlustwahrscheinlichkeiten und der Verfügbarkeit des Dienstes spezifiziert. Darüber hinaus ist auch die Mindestzahl anschließbarer Teilnehmer festgelegt, ebenso wie die Struktur von Gruppenadressen. Diese Beschränkungen sind bei der Planung eines MSS zu beachten.

2.2.3.2 Zugangsnetz

Ein Dienst wie SMDS wird für den Teilnehmer lediglich an der Dienstschnittstelle (*Subscriber Network Interface, SNI*) sichtbar. Daher sollte die Dienstschnittstelle einerseits einfach und damit kostengünstig, andererseits aber auch flexibel gestaltet werden, um Erweiterungen zu ermöglichen. Bei SMDS sind deshalb unterschiedliche Konfigurationen zum Anschluß der Teilnehmer über ein Zugangsnetz vorgesehen, die sowohl Punkt-zu-Punkt-Verbindungen als auch busförmige Zugangsnetze umfassen.

Die Kommunikation der Teilnehmer mit dem Netz erfolgt über das *SMDS Interface Protocol (SIP)*. Es basiert auf den DQDB-Protokollfunktionen für den verbindungslosen Dienst und wird in drei Schichten unterteilt, die im wesentlichen analog zu DQDB aufgebaut sind. Die oberste Schicht unterstützt die Adressierung nach E.164 und besitzt eine Fehlererkennung für Pakete, die in den darunterliegenden Schichten verlorengegangen sind. In dieser Schicht werden auch die obengenannten Funktionen zur Prüfung von Adressen realisiert. Die mittlere Schicht stellt Funktionen für die Segmentierung bzw. das Zusammenfügen von Paketen der darüberliegenden Schicht sowie den Zugriff auf die Zeitschlitzstruktur des Übertragungsmediums bereit. Die unterste Schicht umfaßt die Funktionen der Bitübertragungsschicht.

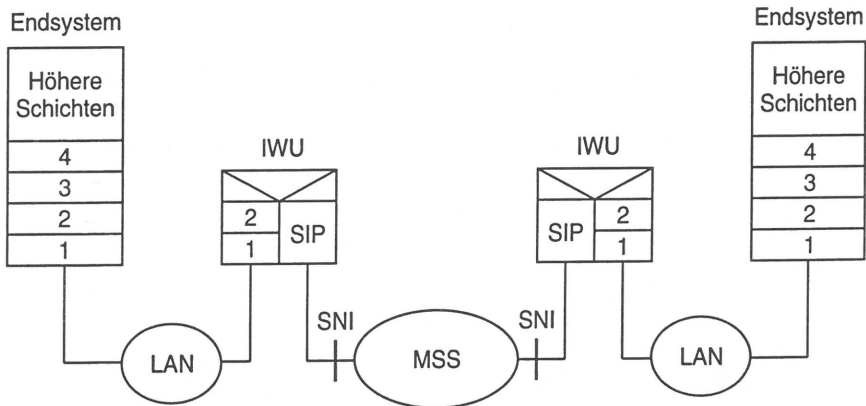


Bild 2.7: LAN-LAN-Kopplung über SMDS

Die Ankopplung von LANs kann auf Schicht 2 oder Schicht 3 des ISO Basisreferenzmodells erfolgen [28]. Die prinzipielle Architektur zur LAN-LAN-Kopplung über SMDS zeigt Bild 2.7. Die Netzkoppeleinheit (IWU) sollte so ausgewählt werden, daß zum einen der zu erwartende Ausbau von SMDS als weltweiter Datendienst als auch die flexible Konfiguration von virtuell privaten Netzen für unterschiedliche Benutzergruppen berücksichtigt wird.

Bei der Kopplung über Schicht 2 bietet sich die Anwendung von *Remote Bridging* an [71]. Dabei werden LANs über eine sog. *Remote Bridge* (RB) mit dem MSS, das die Konnektivität zwischen den RBs sicherstellt, verbunden. Aus der so entstandenen Topologie wird durch die RBs mit Hilfe eines verteilten Algorithmus ein logischer Baum gebildet – die Ausdehnung des *Spanning Tree*-Verfahrens über ein öffentliches Netz hinweg. Einschränkend ist festzustellen, daß mit dieser Konfiguration nur innerhalb eines virtuell privaten Netzes kommuniziert werden kann. Ferner müssen, um die Adreßumsetzung für LAN-Adressen zu gewährleisten, im MSS geeignete Mechanismen, z. B. Adreß-*Server*, vorgesehen werden. Dort werden MAC- und E.164-Adressen einander zugeordnet. Das kann gegebenenfalls sehr aufwendig werden, wie ein ähnliches Problem zeigt, das bei der Kopplung von LANs an das ISDN auftritt [111]. Bei dem beschriebenen System wird versucht, den Aufwand für die Adreßauflösung dadurch zu minimieren, daß stufenweise, mit abnehmender Lokalität, nach der gewünschten Zieladresse gesucht wird. Zunächst wird lokal in der betroffenen RB gesucht; als letzte Instanz ist daran ein Adreß-*Server* beteiligt, wenn die Zieladresse vorher nicht ermittelt werden konnte. Damit verbunden ist jedoch ein aufwendiger Steuerungsmechanismus, der zusätzlichen Kommunikationsaufwand zwischen den RBs und die präzise Einstellung von Zeitgebern einschließt.

Die oben genannten Beschränkungen können durch *Router* vermieden werden. Durch die Verwendung von Netzadressen kann, sofern eine physikalische Verbindung besteht, jeder beliebige Teilnehmer, auch in Netzen unterschiedlicher Netzbetreiber, erreicht werden. *Router*, die das globale Adressierungsschema und die Verkehrslenkung nach OSI verwenden, sind außerdem in der Lage, hierarchische Verkehrslenkungsverfahren zu unterstützen. Dies ist dann sinnvoll, wenn eine große Zahl von Teilnehmern zu vernetzen ist, und eine administrative Aufteilung in verschiedene Domänen notwendig wird. Unter Verwendung dynamischer Verkehrslenkungsverfahren ist eine effektive Ankopplung umfangreicher lokaler Netze an ein SMDS-Netz möglich, wie in [95] gezeigt wird.

2.2.4 MAN-Vermittlungssystem

Die Anforderungen von Diensten wie SMDS, aber auch von zukünftigen Diensten aus dem Bereich der Mobilkommunikation [104, 110], bestimmen die Architektur eines MAN-Vermittlungssystems. Dazu zählen die Topologie der MAN-Teilnetze und der Zugangsnetze, die Strukturierung in Hierarchieebenen und die Funktionen der verschiedenen Netzkoppeleinheiten.

2.2.4.1 Netzstruktur

Grundsätzlich können mehrere MAN-Teilnetze, die jeweils die Zugangsnetze verbinden, zu einem MAN-Vermittlungssystem (MSS) zusammengefaßt werden. Das ist vor allem dann

erforderlich, wenn Stationen in ausgedehnten geographischen Gebieten erreicht werden sollen oder die Kapazität eines einzelnen MAN-Teilnetzes nicht ausreicht. Mehrere MSS können wiederum zu einem Verwaltungsbereich eines Netzbetreibers (*Network Operator Domain*) zusammengefaßt und über spezielle Netzkoppeleinheiten und Transitnetze verbunden werden. Da in Zukunft durch die Deregulierung mehrere Netzbetreiber miteinander konkurrieren werden, sind unterschiedliche Schnittstellen innerhalb der Verwaltungsbereiche der Netzbetreiber und zwischen diesen erforderlich. Diese Schnittstellen wurden ebenso wie die Teilnehmer-Netzschnittstelle von ETSI standardisiert [42, 44]. Genormt werden die Schnittstellen an Bezugspunkten, wobei die physikalischen Eigenschaften und die über die Schnittstelle abgewickelten Protokolle festgelegt werden. Für die Netzplanung ergibt sich die Konsequenz, daß diese Schnittstellen mit unterschiedlichen Merkmalen als zusätzliche Nebenbedingungen berücksichtigt werden müssen.

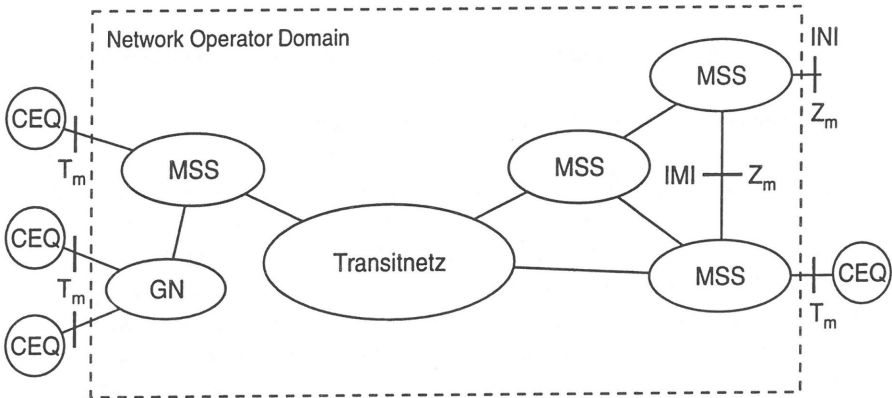


Bild 2.8: MAN-Architekturmodell nach ETSI

Demnach können folgende Schnittstellen und Konfigurationen unterschieden werden (vgl. Bild 2.8):

- Die Teilnehmereinrichtung (*Customer Equipment, CEQ*) kann über eine Punkt-zu-Punkt-Verbindung oder ein Zugangsnetz (*Gathering Network, GN*) an ein MSS angebunden werden (Bezugspunkt T_m). An dieser Stelle endet auch der Hoheitsbereich des Netzbetreibers.
- Innerhalb des MSS gibt es keine Vorgaben durch die Standardisierung, d.h. auch keine Festlegung der Netztechnik. Vorschläge zur Realisierung eines MSS, die im folgenden Abschnitt behandelt werden, basieren durchgängig auf der DQDB-Technik (vgl. Abschnitt 2.2.4.2).

- Zwischen MSS sind abhängig von der administrativen Aufteilung innerhalb des Verwaltungsbereichs eines Netzbetreibers und zwischen den Verwaltungsbereichen unterschiedlicher Netzbetreiber Schnittstellen definiert (Bezugspunkt Z_m : *Inter MSS Interface*, IMI oder *Inter Network operator Interface*, INI) [44]. Der Vermaschungsgrad zwischen diesen Netzbereichen kann beliebig gewählt werden. Mit dem IMI bzw. INI sind auch die Bezugspunkte definiert, an denen mehrere MSS überregional verbunden werden können.

Für die Verbindung von MSS sind sowohl einfache Übertragungsstrecken als auch Transitnetze vorgesehen. Ein Transitnetz kann zum Beispiel durch das zukünftige Breitband-ISDN realisiert werden (vgl. Abschnitt 2.2.5).

2.2.4.2 Realisierung

Die Netzelemente, die zur Realisierung eines MSS in einem öffentlichen Netz eingesetzt werden, umfassen neben den Verbindungs- und Anschlußleitungen im wesentlichen drei Arten von Netzkoppeleinheiten [58, 62, 93, 124]: *Customer Gateways* (CGW), die LANs oder Nebenstellenanlagen mit einem Zugangnetz (*Customer Access Network*, CAN) verbinden – im Sinne von ETSI also ein *Gathering Network* –, *Edge Gateways* (EGW), die den Übergang vom privaten in den öffentlichen Bereich markieren und *Subnetwork Router* (RTR) zur Verbindung der MAN-Teilnetze untereinander. Eine typische Anordnung, bei der die EGWs über eine *Looped-Bus*-Konfiguration verbunden sind, zeigt Bild 2.9. Die Struktur des eigentlichen Transportnetzes, das aus den MAN-Teilnetzen gebildet wird, kann durch ein oder mehrere Teilnetze, die eine Funktion als *Backbone*-Netz wahrnehmen, hierarchisch ausgebaut werden. Diese Teilnetze enthalten lediglich *Router* und besitzen in der Regel eine höhere Übertragungskapazität.

CGWs bzw. EGWs besitzen neben der Medienzugriffsfunktion auch höhere Schichten wie die Vermittlungsschicht, die im wesentlichen der Verkehrslenkung dient, sowie Funktionen des Netzmanagements. Das Netzmanagement umfaßt das Abrechnungsmanagement, Funktionen zur Adreßüberprüfung innerhalb von geschlossenen Benutzergruppen und zur Überwachung von Netzzugangsklassen sowie Funktionen zur Erfassung von Leistungsgrößen. Darüber hinaus sind auch Funktionen des Fehlermanagements vorgesehen. Mit den Netzmanagementfunktionen können Dienste wie SMDS oder CBDS realisiert werden. Die Netzelemente sind mit einer Netzüberwachungsstation (*Network Management Center*, NMC) verbunden, die alle relevanten Meßdaten zentral sammelt. Das NMC verwaltet auch die statischen Verkehrslenkungstabellen, die bei Bedarf in die Netzkoppeleinheiten geladen werden.

Wird ein MAN im privaten Sektor betrieben, werden jeweils ein CGW und ein EGW zu einer Funktionseinheit (*Customer/Network Interface Unit*) zusammengefaßt, da die Trennung in

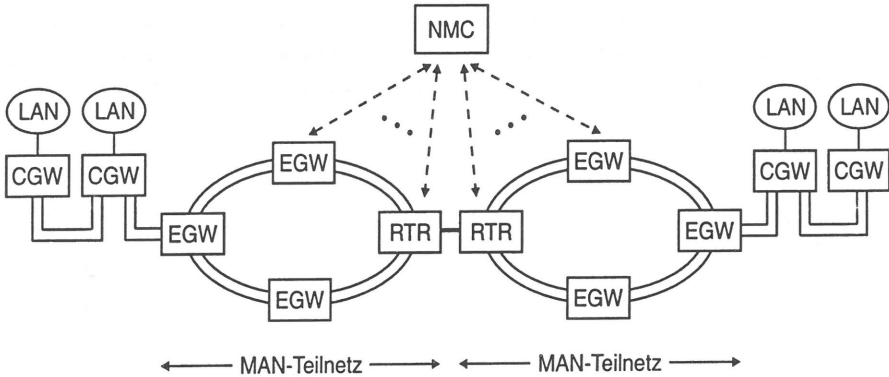


Bild 2.9: Beispiel für ein MSS

Zugangsnetz und eigentliches Transportnetz beim Betrieb unter der Kontrolle einer Organisation entfallen kann. MAN-Vermittlungssysteme, die auf dieser Architektur basieren, wurden in verschiedenen Feldversuchen erfolgreich getestet und werden nun, zunächst in einfachen Konfigurationen, als Breitband-Netzinfrastruktur eingesetzt (vgl. [2, 48, 123, 132, 133]).

2.2.4.3 Netzkopplung und Verkehrlenkung

Bei der Kopplung der MAN-Teilnetze innerhalb eines MSS steht die möglichst ökonomische Nutzung der Netzressourcen im Vordergrund, wobei gleichzeitig die geforderte Dienstgüte eingehalten werden sollte. Um eine ausreichende Verfügbarkeit zu gewährleisten, sollte die Topologie, die durch die Vermaschung der MAN-Teilnetze entsteht, mindestens zwei voneinander unabhängige Pfade zwischen jeweils zwei Teilnetzen aufweisen. Auf diese Weise können durch das Verkehrlenkungsverfahren alternative Pfade ermittelt werden, wenn eine Übertragungsstrecke oder ein Netzknoten ausfällt.

Da MANs aus der LAN-Technik weiterentwickelt wurden, gibt es eine Reihe von Vorschlägen, um die bekannten Verfahren der LAN-Technik zu erweitern und an die Anforderungen eines MAN anzupassen [18, 25, 26, 126, 150, 153]. Die vorgeschlagenen Verfahren verwenden meist eine Kombination von *Spanning Tree*-Verfahren und *Source Routing* oder *Shortest Path Routing*. In allen Fällen wird zumindest hierarchische Adressierung in zwei Stufen (Stations- und Teilnetzadressen) verwendet, und die Verkehrlenkung tabellengesteuert abgewickelt.

Probleme treten sowohl beim *Spanning Tree*-Verfahren als auch bei *Source Routing* hinsichtlich einer Anwendung auf große Netzverbundsysteme auf. Das *Source Routing* bleibt zwar auf die MAN-Stationen beschränkt und wird nicht auf die angeschlossenen LANs ausgedehnt.

aber der durch das Suchen der optimalen Pfade entstehende Verkehr kann bei steigender Stationszahl zu einer relativ hohen Netzlast führen. Das *Spanning Tree*-Verfahren kann neben der bereits erwähnten schlechten Ressourcennutzung in Abhängigkeit von der Anzahl der Teilnetze auch zu deutlich steigenden Pfadlängen führen.

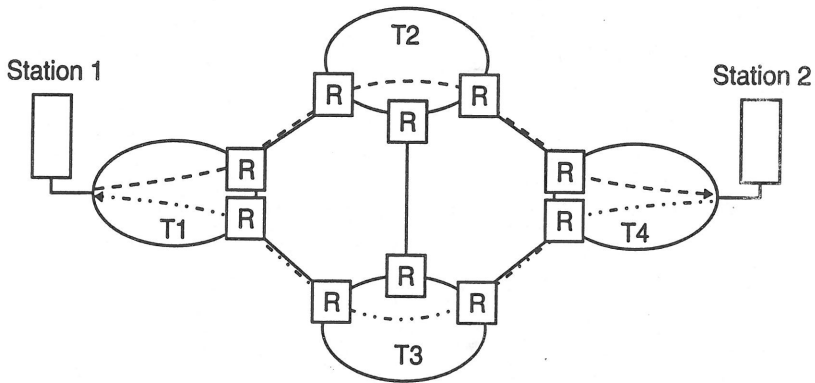


Bild 2.10: Verkehrslenkung im MSS

Eine Verbesserung kann mit Hilfe von *Shortest Path Routing* [126] erzielt werden. Die logische Topologie besteht dabei aus einer Menge von sog. *Shortest Path Trees* (SPTs), die für jedes potentielle Zielnetz statisch oder dynamisch konfiguriert werden können. Der Weg von einem beliebigen Netz zum jeweiligen Zielnetz ist dadurch eindeutig bestimmt. Jede Netzkoppeleinheit besitzt Tabellen, die das nächste Teilnetz auf dem Weg zum Zielnetz angeben, und Informationen über die Zuordnung zu den jeweiligen *Shortest Path Trees* enthalten. Auf diese Weise wird, verglichen mit den oben genannten Verfahren, eine bessere Netzauslastung und eine höhere Flexibilität erreicht. Ein ähnliches Verfahren, das auf zwei Hierarchieebenen ausgedehnt werden kann, ist in [145] beschrieben. Die Verkehrslenkungsschema basiert auch in diesem Fall auf *Shortest Path Trees*.

Ein Beispiel verdeutlicht die prinzipielle Funktionsweise des *Shortest Path Routing* am Beispiel eines MAN-Vermittlungssystems. Bild 2.10 zeigt vier MAN-Teilnetze, die über Router verbunden sind. In jedem Teilnetz befindet sich für einen Übertragungsabschnitt, der zwei Teilnetze verbindet, eine Router-Anschlußeinheit. Die bereits erwähnten SPTs können mit Hilfe der Verkehrslenkungstabellen so konfiguriert werden, daß die Stationen 1 und 2 bis auf die Teilnetze, an die sie angeschlossen sind, über getrennte Pfade in Hin- bzw. Rückrichtung kommunizieren können. Zusätzlich können die Router in den Teilnetzen T1 bzw. T4 so platziert werden, daß die Kommunikationsbeziehung auch in den Teilnetzen zwei unterschiedliche Wege benutzt, d. h. über jeweils verschiedene Busrichtungen abgewickelt wird. Auf

diese Weise ist eine gleichmäßigere Auslastung der Netzressourcen möglich. Beim *Spanning Tree*-Verfahren könnte beispielsweise nur ein Pfad zur Verfügung gestellt werden.

2.2.5 Evolution zum B-ISDN

Die Realisierung von Breitbanddiensten durch die MAN-Technik kann im öffentlichen Bereich als erster Schritt in einer Netzevolution zum Breitband-ISDN (B-ISDN) betrachtet werden [51, 39]. Bei einem vollständigen Ausbau des B-ISDN sollen MANs die Rolle von Zugangsnetzen übernehmen.

Das B-ISDN ist eine Weiterentwicklung des Schmalband-ISDN und basiert auf dem Prinzip der Diensteintegration. Als Übermittlungstechnik wird jedoch der Asynchrone Transfer-Modus (ATM) verwendet [65, 102]. Die Kommunikation erfolgt mittels verbindungsorientierter Paketvermittlung über virtuelle Verbindungen (*Virtual Connection*, VC) bzw. virtuelle Pfade (*Virtual Path*, VP), die mehrere virtuelle Verbindungen umfassen können. Die Information wird in kleinen Paketen fester Länge, sog. Zellen, transportiert. Eine Zelle, die 53 Oktetts umfaßt, besteht aus einem Zellkopf und einem Informationsfeld, wobei der Zellkopf Steuerinformation enthält, und das Informationsfeld zum Transport der eigentlichen Nutzinformation dient. Das Format eines DQDB-Zeitschlitzes ist im übrigen an das ATM-Zellformat angepaßt, wodurch die Kopplung beider Netztypen vereinfacht wird. Darüber hinaus verwendet das B-ISDN ebenfalls das E.164-Adressierungsschema.

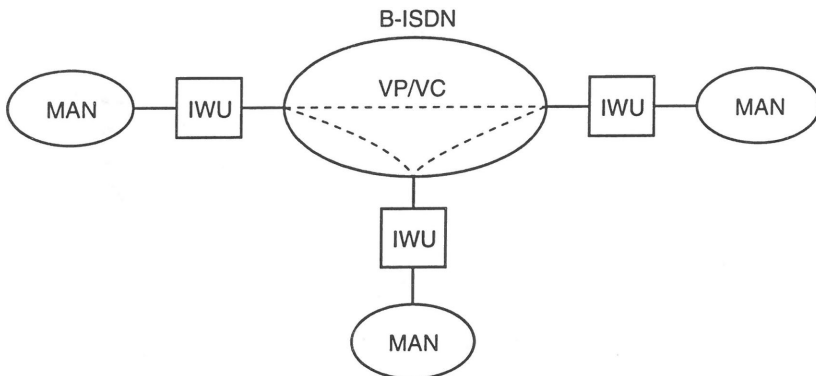


Bild 2.11: MAN-Kopplung über das B-ISDN

Prinzipiell ist die Kopplung von MANs über virtuelle Verbindungen bzw. virtuelle Pfade des B-ISDN möglich (Bild 2.11). Solange sich das B-ISDN im Aufbau befindet, werden lediglich

festgeschaltete VP- oder VC-Verbindungen zwischen den angeschlossenen MANs verfügbar sein. Dadurch kann zwar eine gewisse Anzahl von MANs gekoppelt werden. Bei sehr großen Systemen ist jedoch die Vernetzung über *Connectionless Server* innerhalb des B-ISDN zu bevorzugen [148], da der Vermaschungsgrad durch VP- oder VC-Verbindungen sehr groß werden kann. Dafür sind Netzkoppeleinheiten erforderlich, denen die Anpassungsaufgaben der unterschiedlichen Protokolle zufallen. Wie bereits erwähnt, werden diese Anpassungsaufgaben durch das einheitliche Segment- bzw. Zellformat wesentlich vereinfacht. Eine ausführliche Diskussion der Kopplungsprobleme ist in [142] beschrieben.

Kapitel 3

Grundlagen der Netz- und Konfigurationsplanung

Dieses Kapitel, das sich mit den Grundlagen der Planung und des Entwurfs von Kommunikationsnetzen befaßt, gliedert sich in drei Teile. Neben den grundsätzlichen Modellierungsaspekten werden geeignete Verfahren zur Optimierung von Kommunikationsnetzen dargestellt. Der dritte Teil enthält dann eine Übersicht über wichtige Gesichtspunkte der Netzplanung, insbesondere über die Probleme, die beim topologischen Entwurf von Kommunikationsnetzen auftreten.

3.1 Modellierung von Kommunikationsnetzen

Die Lösung von Planungsproblemen setzt eine sinnvolle Modellbildung voraus, für die sich insbesondere Verkehrsmodelle und Graphen eignen. Die folgenden Abschnitte sollen die wesentlichen Begriffe und Ergebnisse bereitstellen.

3.1.1 Verkehrsmodelle

Verkehrsmodelle ermöglichen es, die Leistungsfähigkeit von Kommunikationssystemen quantitativ zu untersuchen. Ausgehend vom realen System werden die wesentlichen Charakteristika und die beeinflussenden Leistungsgrößen in einem Modell nachgebildet. Das Modell beschreibt sowohl die Beziehungen von Komponenten untereinander als auch das zeitliche Ablaufgeschehen. Eine Leistungsuntersuchung kann mit Hilfe eines Modells simulativ oder analytisch durchgeführt werden, wobei die Simulationstechnik meist eine sehr detaillierte Modellierung und Auswertung des statistischen Verhaltens ermöglicht. Im Hinblick auf die Konfigurationsplanung und, damit verbunden, die Optimierung von Kommunikationsnetzen

können Verkehrssimulationen als Mittel zur Bestimmung von zeitabhängigen Kenngrößen allerdings nur bedingt eingesetzt werden. Die Anwendung von Simulationsmethoden bei Optimierungsverfahren ist auf Grund des hohen Rechenaufwands selbst auf Hochleistungsrechnern für praktisch relevante Problemgrößen nicht möglich. Daher stützen sich die verkehrstheoretischen Methoden für die Optimierung auf analytische bzw. approximative Verfahren.

3.1.1.1 Stochastische Prozesse

Zeitbehaftete Komponenten eines Kommunikationssystems können mit stochastischen Prozessen beschrieben werden. Stochastische Prozesse lassen sich durch die zeitlichen Abstände von aufeinanderfolgenden Ereignissen (Zufallsvariable T_A) mit der Verteilungsfunktion

$$F_A(t) = P\{T_A \leq t\} \quad (3.1)$$

charakterisieren.

Daraus ergibt sich die Verteilungsdichtefunktion gemäß

$$f_A(t) = \frac{dF_A(t)}{dt} \quad (3.2)$$

Im Fall von diskreten Zufallsvariablen genügt die Angabe der Wahrscheinlichkeitsverteilung:

$$f_k = P\{T_A = k\}, \quad k = 0, 1, \dots, \quad (3.3)$$

wobei k auf eine Basiszeiteinheit normiert ist.

Verteilungsfunktionen oder Verteilungen können mit Hilfe von Messungen bestimmt werden. Oft sind nur der Erwartungswert $E[T_A]$ und der Variationskoeffizient

$$c = \frac{\sqrt{\text{VAR}[T_A]}}{E[T_A]} = \sqrt{\frac{E[T_A^2]}{(E[T_A])^2} - 1} \quad (3.4)$$

bekannt, woraus durch Approximation zugehörige Verteilungen gefunden werden können. Die gewöhnlichen Momente $E[T_A^n]$ sind wie folgt definiert:

$$E[T_A^n] = \int_{0-}^{\infty} t^n f_A(t) dt, \quad (3.5)$$

für kontinuierliche Zufallsvariablen bzw.

$$E[T_A^n] = \sum_{i=0}^{\infty} i^n f_i, \quad (3.6)$$

für diskrete Zufallsvariablen.

Im Rahmen dieser Arbeit sind vor allem Prozesse von Interesse, die eine gute Abschätzung des realen Systemverhaltens ohne hohen Rechenaufwand ermöglichen. Günstige Eigenschaften in dieser Hinsicht weist der Poisson-Prozeß auf. Der Poisson-Prozeß, der zur Klasse der Erneuerungsprozesse zählt, d. h. die Ankunftsabstände sind statistisch unabhängig voneinander und gehorchen einer gemeinsamen Verteilung, wird meist als Approximation bei der Überlagerung mehrerer Verkehrsquellen verwendet. Er besitzt darüber hinaus die Eigenschaft der Gedächtnisfreiheit (Markoff-Eigenschaft) und gehört damit auch zur Klasse der Markoff-Prozesse. Diese Prozesse sind in ihrer zukünftigen Entwicklung nur vom gegenwärtigen Zustand abhängig. Die Phasendauern zwischen den Ereignissen sind negativ-exponentiell verteilt mit der Ankunftsrate λ . Damit lauten die charakteristischen Größen:

$$F(t) = 1 - e^{-\lambda t} \quad (3.7)$$

$$f(t) = \lambda e^{-\lambda t} \quad (3.8)$$

$$E[T_A] = \frac{1}{\lambda} \quad (3.9)$$

$$c = 1 \quad (3.10)$$

Bei Systemen, die wie DQDB mit einer festen Taktrate $1/\Delta t$ arbeiten, können Ereignisse, wie der Zugriff auf einen Zeitschlitz, nur zu bestimmten Zeitpunkten auftreten. Es sind also stochastische Prozesse von Interesse, die mit einer diskreten Zufallsvariablen bzw. einer Verteilung $P\{T_A = k \Delta t\}$ beschrieben werden. Für die Modellierung von Bedienzeiten in einem getakteten System ist ein Bernoulli-Prozeß geeignet, dessen Ankunftsabstände einer um eine Zeiteinheit Δt verschobenen geometrischen Verteilung mit

$$f_k = (1 - q) q^{k-1} \quad (3.11)$$

$$E[T_A] = \frac{1}{1 - q} \quad (3.12)$$

$$c = \sqrt{q} \quad (3.13)$$

genügen. Die Parameter der Verteilung sind gemäß $0 \leq q < 1$ und $k = 1, 2, \dots$ beschränkt. Der Bernoulli-Prozeß weist wie der Poisson-Prozeß die Eigenschaft der Gedächtnisfreiheit auf.

3.1.1.2 Systemmodelle

Ein Systemmodell kann mit Hilfe von

- Verkehrsquellen,
- Bedieneinheiten
- und Warteschlangen

beschrieben werden [100]. Verkehrsquellen erzeugen Anforderungen, d. h. Rufe oder Pakete, die nach einem definierten Ankunftsprozeß von außen auf das System treffen. Bedieneinheiten, die meist Prozessoren repräsentieren, bearbeiten diese Anforderungen nach einem bestimmten Bedienprozeß. Bei Wartesystemen dienen Warteschlangen dazu, entstehende Konflikte beim Zugriff auf Betriebsmittel aufzulösen. Eine Warteschlange ist durch ihre Abfertigungsstrategie und die Anzahl der Warteplätze gekennzeichnet. Wichtige Systemgrößen sind z. B. der maximale Durchsatz, Verlustwahrscheinlichkeiten bei endlicher Puffergröße oder die Durchlaufzeiten.

Mit der erweiterten Notation nach Kendall [100] werden der Typ des Ankunftsprozesses, der Typ des Bedienprozesses, die Anzahl der Bedieneinheiten und gegebenenfalls Strukturmerkmale bezeichnet. Als Beispiel sei ein $M/G/1-s$ -System betrachtet, wobei „M“ für einen Markoff-Ankunftsprozeß, „G“ (*General*) für einen Bedienprozeß mit allgemeiner Verteilung, die „1“ für eine Bedieneinheit und „s“ für die Anzahl der Warteplätze stehen.

Werden mehrere Bedienstationen, die aus einer Warteschlange für die Anforderungen und aus einer oder mehreren Bedieneinheiten bestehen können, durch ideale Kanäle miteinander verbunden, so entsteht ein Warteschlangennetz. Ein offenes Warteschlangennetz kann von außen Anforderungen erhalten bzw. bediente Anforderungen nach außen abgeben. Ein geschlossenes Warteschlangennetz enthält dagegen eine konstante Anzahl von Anforderungen. Verzögerungsknoten (*Infinite Server*) sind spezielle Bedienstationen, die jede Anforderung sofort bedienen. Sie dienen zur Modellierung von Laufzeiten innerhalb eines Kommunikationssystems, wenn die Bedienzeiten der einzelnen Bedieneinheiten konstant gewählt werden.

3.1.2 Graphen

Graphen können u. a. zur Modellierung technischer Transportsysteme eingesetzt werden, wie etwa Netze zur Wasser- oder Energieversorgung, oder eben Kommunikationsnetze. Im folgenden werden die für die Arbeit wichtigen Begriffe und Definitionen zusammengestellt sowie einige, für die Netzplanung wichtige Algorithmen behandelt. Für eine eingehende Behandlung der Thematik sei auf die Literatur verwiesen [20, 88].

3.1.2.1 Begriffe und Definitionen

Ein Graph $G = (V, E)$ umfaßt eine Menge V von Knoten und eine Menge E von Kanten, wobei eine Kante $e \in E$ eine zweielementige Teilmenge von V ist, die durch den Anfangs- und Endknoten definiert wird [88]. Für technische Probleme genügt es, eine endliche Zahl von Knoten zu betrachten. Für eine Kante (a, b) heißen die Knoten a bzw. b inzident mit dieser Kante; a und b sind adjazent. Der Grad eines Knotens a , oft ein Maß für die Komplexität

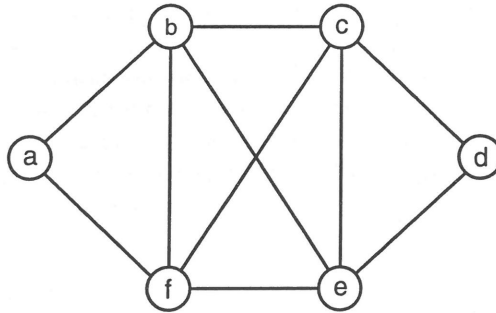


Bild 3.1: Beispiel für einen Graphen

eines Netzes, ist die Anzahl der mit a inzidenten Kanten. Graphen können gerichtete oder ungerichtete Kanten besitzen.

An einem Beispiel (Bild 3.1) sollen einige weitere Definitionen erläutert werden, die insbesondere bei Kommunikationsnetzen eine Rolle spielen. Ein Kantenzug ist eine beliebige Folge von Kanten in einem Graphen G , in Bild 3.1 z. B. durch die Knoten a, b, c, e, b, c, d definiert. Sind die Kanten paarweise verschieden, wird der Kantenzug als Weg bezeichnet (Beispiel: a, b, e, f, b, c, d); sind darüber hinaus auch die Knoten paarweise verschieden, wird der Kantenzug als einfacher Weg oder Pfad bezeichnet (Beispiel: a, b, c, d). Ein einfacher Kreis entsteht aus einem Pfad, wenn Anfangs- und Endknoten identisch sind.

Kanten können mit Gewichten belegt werden, die Eigenschaften wie die geographische Länge, den Informationsfluß über eine Leitung oder die Kapazität einer Leitung bezeichnen. Ein solcher Graph wird auch als Netz und das Kantengewicht als Länge der Kante bezeichnet. Der Verkehrsfluß in einem Netz wird typischerweise durch einen gerichteten Graphen dargestellt.

Für Zuverlässigkeitsbetrachtungen ist der Begriff der Konnektivität („Zusammenhang“) von Bedeutung. Ein Graph ist zusammenhängend, wenn es zwischen jedem Knotenpaar mindestens einen Pfad gibt, d. h. wenn mindestens ein Knoten oder eine Kante entfernt werden muß, um den Graphen in mehrere Subgraphen zu teilen. Ein Graph G wird als k -fach zusammenhängend bezeichnet, wenn je zwei Knoten durch mindestens k knotendisjunkte Pfade verbunden sind.

Ein Baum ist ein zusammenhängender Graph, der kreisfrei ist. Im Zusammenhang mit Netzplanungsverfahren wird oftmals ein erzeugender Baum (*Spanning Tree*) $H = (V', E')$ aus einem gegebenen Graphen $G = (V, E)$ bestimmt. H ist ein Subgraph von G , dessen Knotenmenge V' identisch mit V ist. In Bild 3.1 bilden beispielsweise die Kanten (a, b) , (a, f) , (b, c) , (b, e) , (e, d) einen erzeugenden Baum.

3.1.2.2 Algorithmen

Der Entwurf einer Netztopologie setzt die Modellierung des Kommunikationsnetzes in Form eines Graphen voraus. Bei verteilten Kommunikationssystemen, wie etwa einem Paketvermittlungsnetz, werden typischerweise die Übertragungsabschnitte als Kanten und die Vermittlungsrechner als Knoten eines Graphen dargestellt. Wie bereits erwähnt, können den Knoten und Kanten, auch mehrfach, Gewichte zugeordnet werden, die in der Regel positiv-reelle Werte annehmen.

In vielen Fällen ist die Ermittlung von Kenngrößen oder Strukturen eines Graphen von Interesse. Eine wichtige Kenngröße ist etwa der Durchmesser (*diameter*) eines Graphen, der als maximaler Abstand, der zwischen einem beliebigen Knotenpaar auftreten kann, definiert ist. Offensichtlich werden für die Netzplanung Topologien benötigt, die unter allen möglichen Realisierungen den minimalen Durchmesser aufweisen, was sich vor allem auf die Netzgüte günstig auswirkt. In der Literatur sind dafür eine Reihe fundamentaler, graphentheoretischer Algorithmen bekannt [88], die in vielen Fällen als Grundbausteine eines Netzplanungsverfahrens verwendet werden. Graphentheoretische Fragestellungen sind auch eng verknüpft mit Problemen der kombinatorischen Optimierung (vgl. Abschnitt 3.2.1). Zwei Problemstellungen, die sowohl für die Verkehrslenkung als auch für den topologischen Entwurf von Bedeutung sind, sollen nachfolgend skizziert werden: Das Problem des erzeugenden Baums mit minimalem Gewicht (*Minimum Spanning Tree*) und das Problem der kürzesten Pfade in einem Netz (*Shortest Path*). Zur Lösung dieser Probleme sind in beiden Fällen mehrere Algorithmen bekannt [92].

Das Problem des *Minimum Spanning Tree* kann mit dem Algorithmus nach Prim gelöst werden, bei dem sukzessive die Elemente des minimal erzeugenden Baums aus einem gegebenen Graphen G mit gegebenen Kantengewichten bestimmt werden. G sei zusammenhängend. Zunächst wird ein beliebiger Knoten von G als erstes Element des Baums gewählt. Danach wird bei jedem Schritt die Kante an den Baum angefügt, die das geringste Gewicht unter allen Kanten besitzt, die noch nicht in den Baum eingefügt sind, und von denen jeweils ein Knoten zum Baum gehört. Dies geschieht so lange, bis der Baum vollständig ist. Mit Hilfe dieses Algorithmus kann beispielsweise die Topologie von LANs, die mit *Bridges* gekoppelt sind, optimiert werden (vgl. Abschnitt 2.1.7). Die Netzkoppeleinheiten werden dabei als Kanten, die mit vorgegebenen Gewichten belegt werden, und die LANs als Knoten modelliert.

Das *Shortest Path*-Problem wurde bereits in Abschnitt 2.1.8 im Zusammenhang mit der Verkehrslenkung diskutiert. Der Algorithmus nach Dijkstra ist z. B. ein Lösungsverfahren für den kürzesten Pfad zwischen zwei beliebigen Knoten. Dieser Algorithmus beruht auf der Beobachtung, daß kürzeste Pfade wiederum kürzeste Pfade enthalten. Ein kürzester Pfad, der die Knoten i, j, k der Reihe nach enthält, besteht aus der Folge der kürzesten Pfade (i, j)

und (j, k) . Aus dieser Beobachtung kann ein rekursiver Algorithmus abgeleitet werden, der mit jeder Iteration den momentan kürzesten Teilpfad erweitert.

Ausgehend vom Ursprungsknoten werden die benachbarten Knoten hinsichtlich ihrer Länge verglichen und durch „Marken“, einem Gewicht, das den bis dahin kürzesten Pfad repräsentiert, gekennzeichnet. Der Knoten mit dem kleinsten Gewicht wird gewählt und bildet mit dem Ursprungsknoten den ersten Teilabschnitt des kürzesten Pfades. Von diesem Knoten aus werden nun schrittweise die Längen der Pfade zu den jeweiligen Nachbarknoten bestimmt und diesen wiederum Gewichte zugewiesen, falls nicht bereits ein Nachbarknoten mit kleinerem Gewicht existiert. Erneut wird der Knoten mit der kleinsten Marke ausgewählt, falls er sich nicht schon im bisher vorhandenen kürzesten Teilpfad befindet. Diese Schritte werden so lange wiederholt, bis der Zielknoten erreicht ist.

3.2 Optimierungsverfahren

Bei der Konfigurationsplanung bzw. dem topologischen Entwurf von Kommunikationsnetzen stehen Probleme der diskreten (kombinatorischen) Optimierung im Vordergrund [92, 112]. So werden die Netzgüte und die Netzkosten durch die Kapazität der Übertragungsabschnitte oder der Vermittlungsknoten, die häufig nur in diskreter Menge zur Verfügung steht, entscheidend beeinflusst. Allerdings können nicht alle Probleme der Konfigurationsplanung auf kombinatorische Optimierungsprobleme abgebildet werden, wie in Abschnitt 3.3.5 noch ausgeführt wird.

3.2.1 Kombinatorische Optimierung

Der folgende Abschnitt gibt einen Überblick über die Eigenschaften kombinatorischer Optimierungsprobleme. Ein kombinatorisches Optimierungsproblem besitzt die folgenden Charakteristika [92]:

- Die Variablen nehmen relativ wenige, diskrete Werte an. Häufig handelt es sich um positive, ganze Zahlen, z. B. 0 und 1.
- Der Lösungsraum kann durch den Ausschluß offensichtlich nichtoptimaler Lösungen eingeschränkt werden. Damit kann durch eine Approximation in vielen Fällen eine gute empirische Lösung gefunden werden.
- Algorithmen zur Lösung kombinatorischer Optimierungsprobleme sind oftmals stark von der Problemstruktur abhängig.

Die Formulierung kombinatorischer Optimierungsmodelle kann mit unterschiedlichen mathematischen Hilfsmitteln erfolgen [135]. Beispiele sind graphentheoretische Methoden oder

die Verwendung Boolescher Variablen. Ein einheitliches Grundmodell gibt es jedoch bisher nicht. Aarts und Korst [1] formulieren eine Ausprägung eines kombinatorischen Optimierungsproblems in allgemeiner Weise als Paar (S, g) , wobei der Lösungsraum (Zustandsraum) S die endliche Menge aller möglichen Lösungen und die Zielfunktion g eine Abbildung mit

$$g: S \rightarrow \mathbb{R} \tag{3.14}$$

ist. Soll die Zielfunktion minimiert werden, gilt für das Optimum i_{opt} :

$$g(i_{opt}) \leq g(i), \forall i \in S \tag{3.15}$$

Werden die Zielfunktion und die Nebenbedingungen in Abhängigkeit von Variablen dargestellt, die nur ganzzahlige Werte annehmen können, spricht man von einem ganzzahligen Optimierungsproblem. Bei Variablen, die nur die Werte 0 und 1 annehmen, wird auch von binärer oder Boolescher Optimierung gesprochen (vgl. [68]).

Ein Beispiel für ein kombinatorisches Optimierungsproblem ist das Problem des *Graph Partitioning*, bei dem ein Graph G mit n Knoten, $n = 2, 4, 6, \dots$, in zwei Teilgraphen gleicher Größe (in Bezug auf die Knotenzahl) aufgeteilt werden soll, so daß die Zahl der Kanten, die Endpunkte in beiden Teilgraphen besitzen, minimal wird. Zu beachten ist, daß G nicht zusammenhängend sein muß. Der Lösungsraum S ist gegeben durch

$$S = \{\text{Alle Partitionen der Knoten in zwei Teilmengen gleicher Größe}\} \tag{3.16}$$

mit der Kardinalität (Anzahl der Lösungen)

$$|S| = \frac{n!}{2 \cdot (n/2)! \cdot (n/2)!} \tag{3.17}$$

Die Zielfunktion g berechnet die Zahl der Kanten, die die beiden Teilgraphen verbindet.

Bezüglich ihrer Komplexität können kombinatorische Optimierungsprobleme – vereinfacht betrachtet – in zwei Klassen eingeteilt werden [158]. Zur ersten Klasse gehören Probleme, die in polynomialer Zeit gelöst werden können. Das bedeutet, daß ein Algorithmus zur exakten Lösung des Optimierungsproblems existiert, dessen Rechenschritte n proportional zu einer polynomialen Funktion wie etwa n^2 oder n^3 wachsen. Zur zweiten Klasse gehören Probleme, für die kein Algorithmus bekannt ist, der weniger als eine exponentiell steigende Anzahl von Rechenschritten benötigt, um das Optimum zu finden. Diese Klasse wird in der Komplexitätstheorie auch als NP-vollständig (*Non-deterministic Polynomial-time problem*) bezeichnet. Das bereits erwähnte *Graph Partitioning*-Problem ist NP-vollständig [84]. Eine ausführliche Diskussion der Komplexitätstheorie findet sich in [158].

3.2.2 Exakte und heuristische Verfahren

Wie bereits ausgeführt, können nicht alle kombinatorischen Optimierungsprobleme exakt in polynomialer Zeit gelöst werden. Daher sind Verfahren, die gute Approximationen für

NP-vollständige Probleme liefern, von signifikanter Bedeutung. Prinzipiell können Optimierungsverfahren in exakte und heuristische Verfahren unterschieden werden. Heuristische Verfahren können das Erreichen des globalen Optimums nicht garantieren. Sie sind aber in der Lage, auf Grund ihres Konstruktionsprinzips, Lösungen zu finden, die dem globalen Optimum sehr nahe kommen.

Exakte Verfahren beruhen auf graphentheoretischen Algorithmen (vgl. Abschnitt 3.1.2.2) oder auf der Enumeration des Zustandsraums wie beispielsweise bei *Branch and Bound* [68]. Bei solchen Verfahren muß im schlechtesten Fall der Zustandsraum vollständig abgesucht werden. Bei *Branch and Bound* wird der Zustandsraum schrittweise in zwei oder mehr disjunkte Teilräume aufgeteilt, so daß der vollständige Zustandsraum mit Hilfe eines Entscheidungsbaums dargestellt werden kann. Für jeden dieser Teilräume wird abgeschätzt, ob er das Optimum enthält oder nicht. Dazu werden jeweils Schranken des Zielfunktionswertes berechnet sowie Verzweigungsstrategien eingesetzt, um den nächsten, zu untersuchenden Teilraum zu ermitteln. Durch geeignete Verzweigungsstrategien kann in bestimmten Fällen die exakte Lösung mit vertretbarem Aufwand gefunden werden. Dennoch steigt die Rechenzeit eines *Branch and Bound*-Verfahrens bei Problemen mit vielen Variablen mit Erhöhung der Dimension in der Regel exponentiell [135].

Heuristische Verfahren umfassen einerseits deterministische Strategien wie lokale Suchverfahren, Eröffnungsverfahren, um eine Anfangslösung zu finden, oder sog. Relaxationsverfahren [17, 92, 109, 108, 119] und andererseits statistische Optimierungsverfahren. Bei Relaxationsverfahren werden bestimmte Nebenbedingungen aufgehoben und durch Gewichtung in die Zielfunktion einbezogen, um dann mit bekannten Optimierungsverfahren eine Lösung für das vereinfachte Problem zu ermitteln. Dabei muß jedoch geprüft werden, ob tatsächlich eine Lösung des ursprünglichen Problems vorliegt. Ein Beispiel ist die sog. Lagrange-Relaxation, die vor allem bei linearen ganzzahligen Optimierungsproblemen als Approximationsverfahren eingesetzt wird [52].

Zu den statistischen Optimierungsverfahren zählen neben Monte-Carlo-Verfahren, die den Lösungsraum in rein zufälliger Folge absuchen, *Simulated Annealing* und Genetische Algorithmen (vgl. Abschnitte 3.2.3 und 3.2.4).

Heuristische Verfahren können ferner nach folgenden Kriterien eingeteilt werden:

- **Prinzipielle Strategien zur Lösungsfindung:** Suchverfahren benötigen eine Anfangslösung und versuchen dann durch iterative Verbesserungen eine lokal oder global optimale Lösung zu finden. Eröffnungsverfahren ermitteln im Gegensatz zu iterativen Verfahren, die mehrere Lösungen erzeugen, direkt eine Lösung.
- **Strategien während der Optimierungsphase:** Hier können Verfahren mit und ohne „Gedächtnis“ zum Einsatz kommen. Die Optimierung wird dabei durch einen oder mehrere Parameter gesteuert, um Zyklen bei der Suche im Lösungsraum zu vermeiden.

So wird beispielweise bei einem rein zufallsgesteuerten Verfahren der Lösungsraum generell ohne „Gedächtnis“ abgesucht.

- Anpassung an das Problem: Der Grad der Anpassung an ein spezifisches Optimierungsproblem kann die allgemeine Anwendbarkeit des Verfahrens beschränken.

Bei komplexen kombinatorischen Optimierungsproblemen sind in vielen Fällen lokale Suchverfahren zur Approximation geeignet. Sie stehen außerdem in enger Beziehung zu *Simulated Annealing* (vgl. Abschnitt 3.2.3). Für einen solchen Algorithmus wird ein elementarer Mechanismus benötigt, um aus einer Lösung eine andere abzuleiten, die in unmittelbarer Umgebung liegt. Diese Umgebung wird auch als Nachbarschaft bezeichnet. Im mathematischen Sinne ist eine Nachbarschaft eine (sehr kleine) Teilmenge des Lösungsraums, die alle Lösungen umfaßt, die durch den elementaren Mechanismus erreicht werden können.

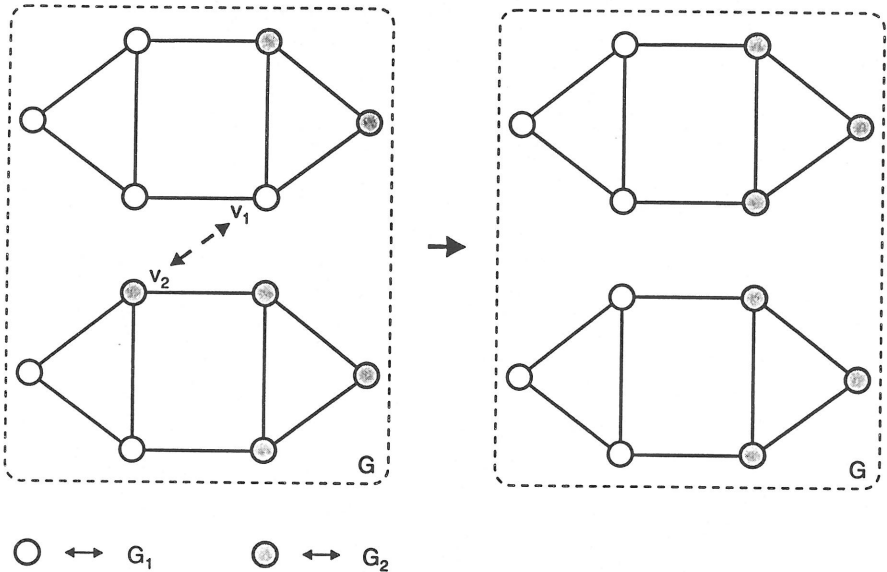


Bild 3.2: Lokal optimale Lösung des *Graph Partitioning*-Problems

Um das Beispiel aus Abschnitt 3.2.1 wieder aufzugreifen, sei ein solcher Mechanismus für das *Graph Partitioning*-Problem beschrieben [84]. Eine Nachbarschaft entsteht durch den Austausch eines Knotens v_1 im Teilgraph G_1 mit einem anderen Knoten v_2 im Teilgraph G_2 , so daß wieder eine gültige Aufteilung (Partition) entsteht. Da dies für jeden Knoten eines Teilgraphen möglich ist, enthält eine Nachbarschaft genau $n^2/4$ Partitionen. Das lokale Suchverfahren kann nun so organisiert werden, daß in mehreren Iterationen jeweils alle benachbarten Partitionen untersucht werden. Dabei wird jeweils die Partition mit dem ge-

ringsten Zielfunktionswert ausgewählt, falls sie existiert. Andernfalls ist das lokale Optimum gefunden. Alternativ kann während jedes Iterationsschrittes innerhalb der Nachbarschaft nach der nächstmöglichen Partition mit einem geringeren Zielfunktionswert als die bis dahin beste Partition gesucht werden. Diese wird dann als momentan beste Lösung gewählt. Falls keine Verbesserung mehr erzielt werden kann, ist das lokale Optimum gefunden.

Der Nachteil lokaler Suchverfahren liegt genau in ihrer Definition. Es ist prinzipiell nicht möglich, Aussagen über die Güte des erreichten, lokalen Optimums zu machen. Entscheidend für die Güte ist insbesondere die ermittelte Anfangslösung. Bild 3.2 zeigt rechts eine lokal optimale Partition mit dem Zielfunktionswert 4, d. h. 4 Kanten enden sowohl in G_1 als auch in G_2 , wohingegen der Wert 0 optimal wäre. Offensichtlich gibt es eine beträchtliche Anzahl von Startlösungen, die zu dieser Aufteilung führen, wenn der Reihe nach je ein beliebiger weißer bzw. grau unterlegter Knoten in der Startlösung gegeneinander ausgetauscht werden (wie beispielsweise beim Graphen G in der linken Hälfte von Bild 3.2). Einen Ausweg bieten Verfahren an, die in bestimmtem Maße auch Verschlechterungen und damit Wege aus den lokalen Optima zulassen. Beispiele dafür sind *Simulated Annealing* und Genetische Algorithmen, auf die in den folgenden Abschnitten vertieft eingegangen wird.

3.2.3 Simulated Annealing

Simulated Annealing ist aus einem Verfahren der Festkörperphysik entstanden, das insbesondere zum Ausglühen von Metallen eingesetzt wird, um gewünschte Materialeigenschaften durch eine möglichst homogene Kristallstruktur zu erhalten. Der Prozeß des *Annealing* wird in zwei Schritten durchgeführt: Zunächst wird der Festkörper sehr lange auf eine bestimmte Temperatur erhitzt. Danach wird er langsam abgekühlt, so daß die Atome den Zustand niedrigster Energie annehmen können und sich in einer, idealerweise perfekten Kristallstruktur anordnen. Die Abkühlung erfolgt in Temperaturschritten, bei denen der Prozeß jeweils eine längere Zeit stabil gehalten wird.

Dieser Prozeß wurde von Metropolis et. al. [1] mit einem einfachen Algorithmus simuliert. Daraus entwickelten Kirkpatrick et. al. [96] einen Algorithmus zur Lösung kombinatorischer Optimierungsprobleme, der auf Probleme wie die optimale Platzierung von Schaltungsteilen beim Entwurf von VLSI-Schaltungen und das Rundreiseproblem (*Traveling Salesman Problem*) angewendet wurde. Eine detaillierte Untersuchung zu *Simulated Annealing* findet sich in [1].

Prinzipiell ist *Simulated Annealing* zur Lösung beliebiger Optimierungsprobleme geeignet. Nachfolgend soll jedoch nur die Anwendung für die kombinatorische Optimierung betrachtet werden. Der Kern eines Algorithmus, der auf *Simulated Annealing* basiert, ist ein iteratives Suchverfahren, das durch die Erhöhung des Zielfunktionswertes (bei Minimierung) mit

einer bestimmten Wahrscheinlichkeit lokale Minima vermeiden kann. Zu Beginn wird eine Anfangslösung, die die Nebenbedingungen erfüllt, erzeugt und dann iterativ verändert. Jede neue Lösung wird zufällig aus einer Nachbarschaft ausgewählt, auf die Gültigkeit der Nebenbedingungen überprüft und mit Hilfe der Zielfunktion bewertet. Falls der Wert der Zielfunktion kleiner ist als zuvor, wird die Lösung in jedem Fall akzeptiert; andernfalls wird die Lösung angenommen, falls die Wahrscheinlichkeit P_A ,

$$P_A = \exp\left(\frac{g(i) - g(i-1)}{\Theta}\right), \quad (3.18)$$

größer ist als eine gleichförmig verteilte Zufallszahl aus dem Intervall $(0, 1]$. $g(i-1)$ bzw. $g(i)$ ist der Zielfunktionswert der vorherigen bzw. neuen Lösung; Θ ist ein Kontrollparameter, der die Akzeptanzwahrscheinlichkeit steuert, und nach einer bestimmten Anzahl von Schritten entsprechend einer Berechnungsvorschrift vermindert wird. Θ entspricht der Temperatur, die den physikalischen *Annealing*-Prozeß charakterisiert. Mit dem Absenken von Θ nimmt im Verlauf des Optimierungsprozesses auch die Wahrscheinlichkeit ab, daß Verschlechterungen akzeptiert werden.

Bei der Entwicklung eines speziellen Algorithmus auf der Basis von *Simulated Annealing* sind allgemeine und problemspezifische Parameter zu unterscheiden [84]. Dabei ist anzumerken, daß die Güte eines solchen Verfahrens entscheidend von der Qualität des Mechanismus abhängt, der eine Nachbarschaft erzeugt und der problemspezifisch zu entwickeln ist. In dieser Hinsicht stellt *Simulated Annealing* eine Verallgemeinerung eines lokalen Suchverfahrens dar. Zu den allgemeinen Parametern, die gewählt werden müssen, zählen

- die Anfangswerte des Kontrollparameters Θ und der Akzeptanzwahrscheinlichkeit P_A ,
- die Berechnungsvorschrift zur Verminderung von Θ ,
- die Zahl der Iterationen (Zustandsübergänge) für jeden Wert von Θ
- und das Abbruchkriterium.

Diese Parameter werden als *Cooling Schedule* bezeichnet.

Für *Simulated Annealing* wurden theoretische Untersuchungen über das Konvergenzverhalten durchgeführt, die auf einer Modellierung als statistischem Gleichgewichtsprozeß beruhen (vgl. [1]). Sie zeigen, daß das Verfahren asymptotisch in das globale Optimum konvergiert. Dazu ist jedoch eine unendliche Zahl von Zustandsübergängen erforderlich. Folgerungen, wie die Parameter des *Cooling Schedule* zu wählen sind, lassen sich daraus nur begrenzt ableiten [84]. Deshalb können sowohl die Parameter als auch die Mechanismen zum Generieren von Nachbarschaften lediglich experimentell überprüft werden. Bekannte graphentheoretische Probleme wurden in dieser Hinsicht bereits untersucht [84, 85].

Es können jedoch einige Aussagen über das *Cooling Schedule* gemacht werden, die beim Entwurf eines speziellen Optimierungsalgorithmus auf der Basis von *Simulated Annealing*

nützlich sind. Der Anfangswert von Θ wird typischerweise sehr groß gewählt, z. B. für eine Akzeptanzwahrscheinlichkeit $P_A = 0.99$, damit zu Beginn fast alle Zustandsübergänge möglich sind. Weiterhin lautet eine häufig verwendete Vorschrift, wie Θ zu berechnen ist:

$$\Theta_i = \tau \cdot \Theta_{i-1}, \quad (3.19)$$

wobei $\tau = 0.9 \dots 0.99$ ist, und i den i ten Iterationsschritt bezeichnet. Die Anzahl der Zustandsübergänge für jeden Wert von Θ sollte gleich der Kardinalität der Nachbarschaft gewählt werden, so daß eine hohe Wahrscheinlichkeit besteht, den größten Teil einer Nachbarschaft zu erreichen. Schließlich wird der Prozeß terminiert, falls sich die Zielfunktion nicht mehr wesentlich verbessert.

3.2.4 Genetische Algorithmen

3.2.4.1 Übersicht

Genetische Algorithmen haben ihren Ursprung in den Mechanismen der Evolution und der Genetik [53, 60, 115, 147]. Die prinzipielle Idee ist es, die genetischen Basismechanismen nachzubilden und auf Optimierungs- oder Suchprobleme anzuwenden. In der Natur kann das Prinzip der Evolution stark vereinfacht mit dem Schlagwort „Survival of the fittest“ beschrieben werden. Das bedeutet, daß ein Ausleseprozeß von Individuen stattfindet, bei dem sich diejenigen durchsetzen, die am besten auf ihre sich ändernde Umgebung angepaßt sind. Da die Eigenschaften von Individuen durch ihre Gene bestimmt sind, manifestiert sich der Anpassungsprozeß auch in der Änderung des Erbgutes. Im Prinzip wirken die natürliche Auslese und die genetischen Mechanismen zusammen und bilden die treibenden Kräfte der Evolution. Die Auslese der Individuen bewirkt also gleichzeitig die Auslese des Erbgutes.

Die Gesamtheit der Gene wird auch als Gen-Pool bezeichnet, wobei ein Chromosom eine gewisse Anzahl Gene umfaßt. Als Basismechanismus für den Austausch von Genen wird das sog. *Crossover* angesehen, bei dem bestimmte Abschnitte zweier Chromosomen ausgetauscht werden. Erst durch das *Crossover* und die Selektion kann sich der Gen-Pool fortentwickeln und Individuen hervorbringen, die besser an ihre Umgebung angepaßt sind.

Mehrere Autoren veröffentlichten Vorschläge, wie die Mechanismen der Genetik in statistische Optimierungsverfahren umzusetzen seien [53]. Mittlerweile hat sich für diese Verfahren der Begriff Genetische Algorithmen und eine einheitliche Nomenklatur durchgesetzt. Genetische Algorithmen versuchen, durch die Manipulation einer Population mit Hilfe von sog. genetischen Operatoren, ein Optimierungsproblem zu lösen. Dazu wird das Problem mit Hilfe von Zeichenketten in eine dem Algorithmus zugängliche Form abgebildet. Der prinzipielle Ablauf umfaßt die Erzeugung einer Startpopulation und deren Fitnessbewertung, um daraus iterativ mit Hilfe der Selektion und den genetischen Operatoren neue Generationen

zu gewinnen. Dieser Vorgang wird durch ein Abbruchkriterium, wie z. B. eine vorgegebene Zahl von Generationen, terminiert.

Genetische Algorithmen sind eng verwandt mit den sog. Evolutionsstrategien [134]. Ein wichtiger Unterschied besteht darin, daß Genetische Algorithmen primär *Crossover* verwenden, während Evolutionsstrategien sich auf die Mutation als Suchmechanismus abstützen und in der Regel auch nicht die Repräsentation durch Zeichenketten unterstützen [147]. Außerdem erfolgt die Selektion bei Evolutionsstrategien deterministisch, während mit Genetischen Algorithmen die Individuen über einen Zufallsmechanismus ausgewählt werden.

Wie die Verfahren des *Simulated Annealing* besitzen Genetische Algorithmen einen problem-spezifischen und einen problemunabhängigen Teil. In den folgenden Abschnitten werden Aspekte der problemspezifischen Repräsentation (Codierung) und die wichtigsten, problem-unabhängigen Komponenten diskutiert.

3.2.4.2 Codierung

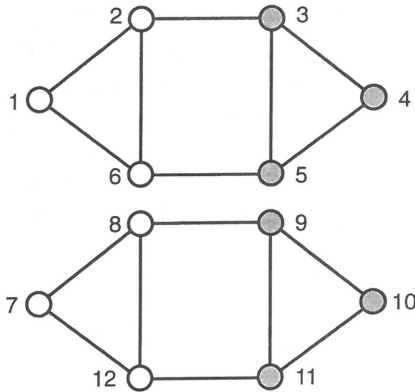
Um ein Problem für die Bearbeitung durch Genetische Algorithmen zugänglich zu machen, muß es in Form von Zeichenketten (*Strings*), die den Chromosomen entsprechen, codiert werden. Die Elemente der Zeichenkette nehmen üblicherweise nur binäre Werte an; es sind aber auch andere ganzzahlige Werte möglich. Ein Gen wird dabei durch den Wirkungsort (Position der Gens) und den Wert eines Elementes einer Zeichenkette definiert.

Zur Verdeutlichung soll wiederum das *Graph Partitioning*-Problem dienen (vgl. Abschnitt 3.2.1). Bild 3.3 zeigt eine Zeichenkette für ein Problem mit zwölf Knoten. Die Zugehörigkeit eines Knotens zu einem Teilgraphen wird durch den Wertebereich $\{0, 1\}$ (binäre Codierung) gekennzeichnet. Mit Hilfe des Wirkungsortes wird markiert, welcher Knoten sich in dem jeweiligen Teilgraphen befindet. Im Beispiel befindet sich Knoten 5 im zweiten Teilgraphen und Knoten 6 im ersten Teilgraphen.

3.2.4.3 Fitness, Selektion und Reproduktion

Ein Individuum ist durch seine Fitness gekennzeichnet. Die Zielfunktion, die die einzelnen Zeichenketten bewertet, wird zu diesem Zweck auf eine Fitnessfunktion abgebildet. Der Wertebereich der Fitnessfunktion umfaßt in jedem Fall positiv reelle Werte. Für das *Graph Partitioning*-Problem ist eine mögliche Fitnessfunktion der Kehrwert der Anzahl von Kanten zwischen den beiden Partitionen.

Zur Auswahl der Individuen folgen die Genetischen Algorithmen dem Prinzip „Survival of the fittest“, d. h. eine Zeichenkette, die einen höheren Fitnesswert erzielt, hat eine größere



Wirkungsort	1	2	3	4	5	6	7	8	9	10	11	12
Wert	0	0	1	1	1	0	0	0	1	1	1	0

Bild 3.3: Beispiel für eine Zeichenkette mit binärer Codierung (*Graph Partitioning*)

Chance, in der nächsten Generation vertreten zu sein. Im einfachsten Fall wird ein Individuum Nachkommen proportional zu seiner Fitness erhalten. Ist die Fitness eines Individuums gleich der mittleren Fitness der Population, so wird im Mittel ein Nachkomme pro Individuum erzeugt.

Zur Realisierung der Selektion kann das sog. Rouletteverfahren verwendet werden, bei dem jedes Individuum einen Sektor proportional zu seiner Fitness erhält. Es wird dann so lange ein Sektor zufällig ausgewählt und für das korrespondierende Individuum jeweils ein Nachkomme für die neue Population erzeugt, bis die neue Population aufgefüllt ist. Damit ist die Zahl der Nachkommen pro Individuum im Mittel proportional zur Sektorgröße. Die Populationsgröße wird in der Regel konstant gehalten.

Durch den beschriebenen Mechanismus haben auch Individuen mit einer geringen Fitness eine gewisse „Überlebenschance“. Diese Individuen können in den folgenden Generationen dazu beitragen, noch bessere Lösungen zu finden. Deshalb sind auch Genetische Algorithmen ganz allgemein in der Lage, aus lokalen Optima wieder herauszufinden. Ein Problem stellt in manchen Fällen die frühzeitige Konvergenz des Verfahrens in ein lokales Optimum dar. In diesem Fall besteht die Population hauptsächlich aus ähnlichen oder gleichen Individuen. Die Verbesserung der Individuen ist damit durch die stark eingeschränkte Vielfalt der Population fast nicht mehr möglich.

Daher werden zwei Ansätze verfolgt, um einerseits ein frühzeitiges Dominieren bestimmter Individuen zu beschränken und andererseits, wenn eine Population aus ähnlichen Individuen besteht, den besten eine signifikant unterschiedliche Fitness zuzuweisen. Der erste Ansatz versucht, durch eine Änderung der bereits beschriebenen Selektionsmethode, die auf einer Gleichverteilung beruht, die Mannigfaltigkeit einer Population zu erhalten. So verspricht die sog. *Elitist Selection* in dieser Hinsicht eine verbesserte Leistungsfähigkeit [53]. Bei diesem Verfahren bleibt bei der Erzeugung einer neuen Population immer das Chromosom mit dem höchsten Fitnesswert erhalten.

Eine andere Möglichkeit bietet die Skalierung der Fitnessfunktion, bei der Individuen mit höherer Fitness nach einer bestimmten Funktion stärker gewichtet werden. Außerdem kann durch die Skalierung ein negativer Wertebereich der Zielfunktion in einen positiven Wertebereich für die Fitness transformiert werden. Häufig verwendete Skalierungsfunktionen sind lineare Funktionen oder Funktionen, die die skalierte Fitness mit Hilfe einer Potenzfunktion berechnen.

3.2.4.4 Genetische Operatoren

Die Individuen einer Population werden durch sog. genetische Operatoren verändert. Basisoperatoren sind *Crossover* und *Mutation*, die innerhalb des von Holland entwickelten Verfahrens (*Simple Genetic Algorithm*) zuerst verwendet wurden. Nach der zufälligen Auswahl zweier Chromosomen aus einer Population kann ein *Crossover*-Operator auf sie angewendet werden. Hierfür wurden verschiedene Verfahren zum Austausch der Gene entwickelt. Das einfachste Verfahren ist die zufällige Bestimmung einer Position im Chromosom und die Vertauschung der nachfolgenden Gene an ihren jeweiligen Positionen (*One-Point-Crossover*, s.

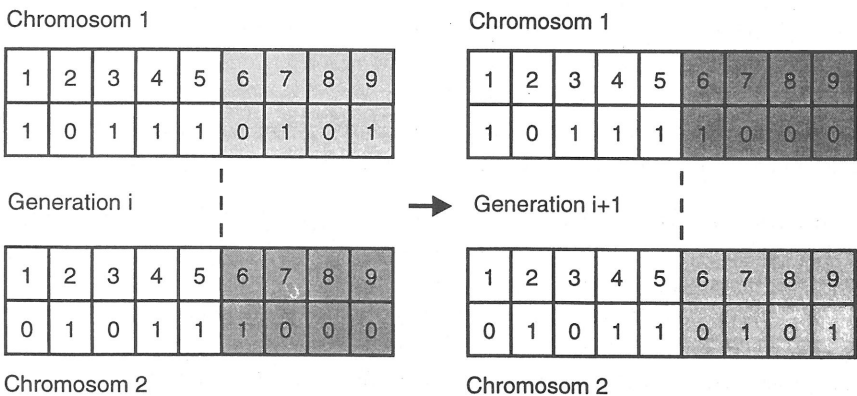


Bild 3.4: *One-Point-Crossover*-Operator

Bild 3.4). Auf diese Weise entsteht ein neues Chromosomenpaar.

Daran anschließend werden zufällig ausgewählte Gene durch den Mutations-Operator verändert. Bei binärer Codierung wird das durch das zufällige Ändern des Wertes einer Zeichenkette (von 0 auf 1 und umgekehrt) erreicht. Die Mutationen beeinflussen sich nicht gegenseitig. Die Steuerung dieser Mechanismen erfolgt über die *Crossover*- und Mutationswahrscheinlichkeiten. Die Mutationswahrscheinlichkeit liegt meist um Größenordnungen unter derjenigen für das *Crossover*.

In der Literatur sind eine Reihe weiterer Operatoren wie Inversion oder PMX (*Partially Matched Crossover*) bekannt [60]. PMX und Inversion verwenden einen erweiterten *Crossover*-Mechanismus. Diese Operatoren werden vor allem dann benötigt, wenn eine Codierung verwendet wird, bei der die Reihenfolge der Elemente einer Zeichenkette erhalten bleiben soll. An einem Beispiel sei der Mechanismus des PMX-Operators, der auch bei der topologischen Optimierung mittels Genetischer Algorithmen verwendet werden kann, genauer erläutert.

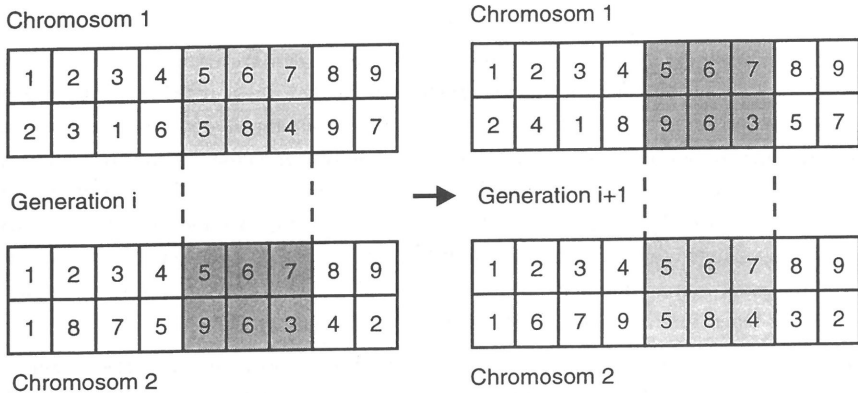


Bild 3.5: PMX-Operator

Dazu sei eine Folge natürlicher Zahlen $1, 2, \dots, n$ betrachtet, deren Reihenfolge bezüglich einer Zielfunktion optimiert werden soll. Das bereits erwähnte Rundreiseproblem, bei dem der kürzeste Weg durch n Städte gesucht wird, kann auf diese Weise repräsentiert werden. Wie beim *One-Point-Crossover*-Operator werden zunächst zwei Individuen ausgewählt. Für beide Individuen werden zufällig zwei Positionen bestimmt, zwischen denen die Gene ausgetauscht werden (s. Bild 3.5). Dabei können Zeichenketten entstehen, in denen ein oder mehrere Gene zweifach vorhanden sind. Um wieder zwei gültige Individuen zu erhalten, bei denen die Zahlenfolge eindeutig ist, müssen auch Genwerte außerhalb der zuvor bestimmten Positionen vertauscht werden.

Im Beispiel werden in Generation $i + 1$ die Werte 9, 6, 3 gegen 5, 8, 4 in Chromosom 1 bzw. 5, 8, 4 gegen 9, 6, 3 in Chromosom 2 vertauscht, da diese Genwerte nach dem Austausch der Gene innerhalb der Zeichenkette jeweils doppelt vorhanden sind.

3.2.4.5 Arbeitsweise

Die Arbeitsweise der Genetischen Algorithmen kann mit Hilfe der *Building Block*-Hypothese beschrieben werden [60], die zwar in der Forschung noch nicht gänzlich akzeptiert ist [53], aber in vielen Anwendungsfällen eine brauchbare Grundlage zur Entwicklung des problem-spezifischen Teils Genetischer Algorithmen liefert. Dazu wird der Begriff des sog. Schemas eingeführt. Ein Schema beschreibt eine Teilmenge von Zeichenketten, die sich hinsichtlich fester Gen-Positionen ähnlich sind. So repräsentiert das Schema $\{101**\}$ eine Menge mit vier Zeichenketten: $\{10100, 10101, 10110, 10111\}$. Die Länge eines Schemas wird durch die äußeren Gen-Positionen, die durch das Schema fixiert sind, bestimmt. Im Beispiel beträgt die Länge zwei Positionen. Außerdem kann für ein Schema die mittlere Fitness ermittelt werden. Diese berechnet sich als Mittelwert der Fitness aller Zeichenketten, die das Schema enthalten.

Die Arbeitsweise der Genetischen Algorithmen ist durch den Wettbewerb zwischen unterschiedlichen Schemata gekennzeichnet. Schemata mit hoher Fitness und einer geringen Länge werden *Building Blocks* genannt. Die *Building-Block*-Hypothese besagt, daß Individuen mit hoher Fitness aus einzelnen *Building Blocks* zusammengesetzt werden. Mutation und *Crossover* unterstützen den Prozeß der Selektion, da insbesondere durch den *Crossover*-Operator der Austausch von *Building Blocks* gefördert wird. Wichtig ist dabei, daß die Schemata eine geringe Länge besitzen, da sie sonst mit hoher Wahrscheinlichkeit durch den *Crossover*-Operator geteilt werden und dann nur eine geringe „Überlebenschance“ besitzen. Durch den Mutations-Operator können zusätzlich völlig neue Schemata entstehen.

Mit jeder Generation können mit Hilfe der Genetischen Algorithmen im Vergleich zur Zahl der Individuen eine sehr große Zahl von Schemata untersucht werden. Die Anzahl untersuchter Schemata liegt bei einer Populationsgröße von n Individuen ungefähr in der Größenordnung von n^3 [60]. Aus diesem Grund wird die Arbeitsweise der Genetischen Algorithmen auch als „implizit parallel“ bezeichnet.

3.3 Planung von Kommunikationsnetzen

3.3.1 Übersicht

Die Planung von Kommunikationsnetzen umfaßt eine Reihe von Aktivitäten, die sich durch die Wahl der Netzebene und die Technologie, für die geplant wird, den Zeithorizont und die

Art der Netzgestaltung unterscheiden [10, 159]. Ziel der Netzplanung ist es, Netzressourcen, wie beispielsweise Übertragungsleitungen, Vermittlungseinrichtungen oder auch Dienstelemente in geeigneter Menge, rechtzeitig und zu möglichst geringen Kosten in einem Kommunikationsnetz bereitzustellen [64]. Aus dieser Sicht stellen viele Aufgaben der Netzplanung Optimierungsprobleme dar.

Netzplanung ist aber auch als kontinuierlicher Prozeß mit schrittweisen Verbesserungen aufzufassen, insbesondere wenn die mittlerweile schnell voranschreitende technologische Entwicklung so unterschiedlicher Gebiete wie die Mobil- oder die Breitbandkommunikation betrachtet wird. Die Planung sollte immer dann überdacht und gegebenenfalls geändert werden, wenn neue Netztechniken signifikante Verbesserungen erwarten lassen. Hinzu kommen die Änderungen der regulatorischen Rahmenbedingungen, wie etwa die Aufhebung von Netzmonopolen, die sich typischerweise in mehreren Etappen vollzieht [139]. Ein wichtiger Parameter ist auch die Nutzungsdauer von Netzressourcen.

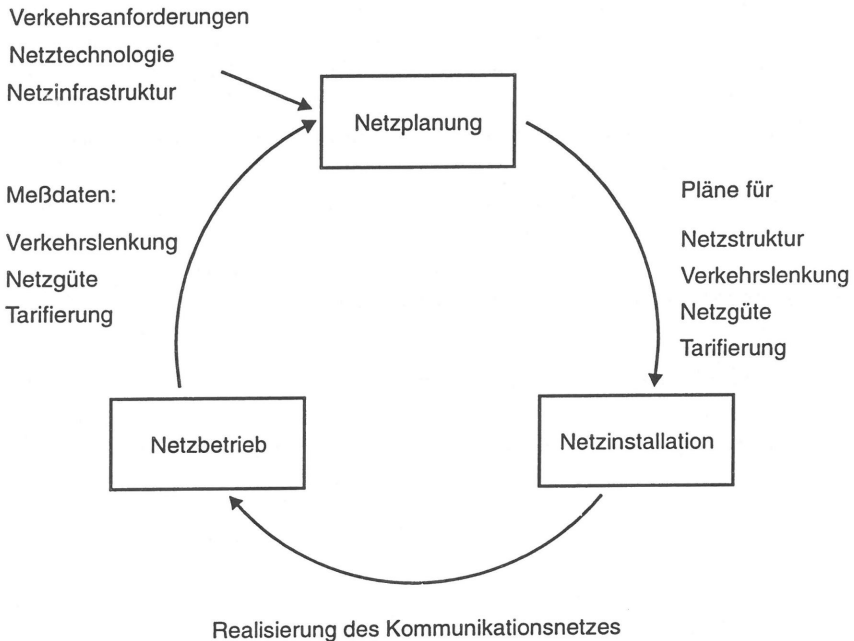


Bild 3.6: Lebenszyklus eines Kommunikationsnetzes

Der Planungsprozeß selbst kann in den sog. Lebenszyklus eines Kommunikationsnetzes eingeordnet werden, der auch die wesentlichen Phasen des Netzmanagements umfaßt (s. Bild 3.6):

Wird ein Netz neu geplant, müssen zunächst die Verkehrsanforderungen der Teilnehmer, die geographischen Standortdaten und andere technische und administrative Nebenbedingungen ermittelt werden. Ferner sollten geeignete Netztechnologien ausgewählt und die Entwurfsparameter festgelegt werden. Mit Hilfe der entwickelten Pläne wird das Netz realisiert. Nach der Installation des Netzes bilden dann im Betrieb die Daten über das Verkehrsaufkommen oder Komponentenausfälle, die in periodischen Abständen gesammelt werden, die Grundlage, um das Netz zu überwachen und hinsichtlich seiner Netzgüte zu analysieren. Die Netzanalyse ist wiederum die Basis für Konfigurationsänderungen oder Netzerweiterungen, die bei Maßnahmen zur Verbesserungen der Netzgüte, steigender Teilnehmerzahl oder der Einführung neuer Dienste notwendig werden.

3.3.2 Netzebenen

Zur Vereinfachung des Planungsprozesses wird ein Kommunikationsnetz in eine physikalische und eine logische Ebene aufgeteilt [159]. Die physikalische Ebene umfaßt die Übertragungsstrecken und Einrichtungen wie Multiplexer oder *Cross Connects*; die logische Ebene ein Verbindungsnetz, das die logischen Verbindungen zwischen Vermittlungsstellen oder Netzkoppeleinheiten beschreibt. Typischerweise werden diese Ebenen getrennt geplant. Weitverkehrsstrecken werden auf Grund von langfristigen Bedarfsprognosen mit Hilfe von Entwurfsregeln, die geographische Randbedingungen, vorhandene Infrastruktur und Abschätzungen des Verkehrsaufkommens zwischen Knotenpunkten (Ballungsräume, größere Städte) miteinbeziehen, eingerichtet. Sie bilden dann die Basis zur Realisierung des eigentlichen Kommunikationsnetzes. Ein Beispiel ist der Ausbau des Glasfasernetzes in Deutschland und die Realisierung von Breitbandnetzen unter Nutzung von Teilstrecken.

Durch die zunehmende Verknüpfung von Übertragungs- und Vermittlungstechnik wird die strikte Trennung der beiden Ebenen in Zukunft mehr und mehr aufgehoben, so daß die Planung durch die Berücksichtigung wechselseitiger Abhängigkeiten eher aufwendiger wird. Diese Tendenz wird voraussichtlich bei zukünftigen photonischen Netzen mit rein optischer Übertragungs- und Vermittlungstechnik noch deutlicher werden [146].

3.3.3 Zeithorizont

Die Planungsaktivitäten können nach ihrem zeitlichen Horizont unterschieden werden, den Rahmen für Investitionsentscheidungen bildet und Auswirkungen auf das Planungsmodell hat [159]:

- kurzfristig (mehrere Monate - 1 Jahr)
- mittelfristig (mehrere Jahre)

- langfristig (mehr als 10 Jahre)

Je nach Planungshorizont wird dann versucht, basierend auf bestehenden Daten, eine Prognose über die Entwicklung des Bedarfs nach Telekommunikationsdiensten aufzustellen. Langfristige Vorhersagen sind jedoch oft mit erheblicher Unsicherheit behaftet, insbesondere, wenn Informationen benötigt werden, in welchem Zeitraum neue Dienste bereitgestellt werden, oder wann ein konkreter Bedarf entsteht. Diesem Umstand kann bis zu einem gewissen Grad durch Szenarien über die technische und ökonomische Bedarfsentwicklung entgegengewirkt werden [16]. Der Bedarf ist insbesondere auch von der demographischen und soziologischen Entwicklung eines Wirtschaftsraumes abhängig.

3.3.4 Verkehrstechnische Netzgestaltung

Für die Netzdimensionierung ist die verkehrstechnische Netzgestaltung (*Traffic Engineering*) von großer Bedeutung. Dabei können folgende Aktivitäten unterschieden werden [47, 83, 156, 157]:

- Messung und Analyse des Nachrichtenverkehrs
- Überlastabwehr und Lastbalancierung
- Kapazitätserweiterung
- Grundlegende Netzstrukturierung

Treten kurzfristige Lasterhöhungen in einem Kommunikationsnetz auf, kann durch Mechanismen der Überlastabwehr und der Lastbalancierung die erforderliche Leistungsfähigkeit erhalten bleiben und Staueffekte vermieden werden. Eine Lasterhöhung kann jedoch, ausgelöst durch erhöhtes Verkehrsaufkommen der Teilnehmer, über einen längeren Zeitraum bestehen. Abhängig von den vorhandenen Netzressourcen muß dann die verfügbare Netzkapazität, beispielsweise durch das Schalten weiterer Kanäle oder durch leistungsfähigere Vermittlungsknoten, erhöht werden. Wenn diese Maßnahmen jedoch nicht ausreichen, wird eine grundlegende Neugestaltung des Netzes unumgänglich. Die beiden letztgenannten Aktivitäten werden als Kapazitäts- bzw. Konfigurationsplanung oder topologischer Entwurf bezeichnet, da sie dem Entwurf einer, hinsichtlich gegebener Kriterien, optimalen Netzstruktur und -konfiguration dienen (*Network Design*). Derartige Planungsprobleme können mit Optimierungsverfahren gelöst werden. Wesentlich ist dabei die Strukturierung des Netzes unter wirtschaftlichen Gesichtspunkten, d. h. die auf Grund der Strukturierung entstehenden Netzkosten sind zu minimieren. Die Konfigurationsplanung erstreckt sich auch auf Aspekte der Netzverfügbarkeit und Fragen der Adressierung oder Sicherheit.

Die Netzdimensionierung richtet sich meist nach der Spitzenbelastung, die in der sog. Hauptverkehrsstunde auftritt. Dabei spielen sowohl der Wochentag und saisonale Einflüsse als auch die Tarifierung eine Rolle [159].

3.3.5 Topologischer Entwurf

Dieser Abschnitt befaßt sich mit allgemeinen Problemstellungen der Konfigurationsplanung, die bei Planung und Entwurf zentral oder dezentral gesteuerter Kommunikationsnetze auftreten. Die folgenden Überlegungen bilden den Ausgangspunkt für das Planungsmodell, das in Kapitel 5 vorgestellt wird.

3.3.5.1 Zentrale Strukturen

Zentrale Strukturen spielen heute nur noch in Teilnehmerzugangnetzen, die den Anschluß zum öffentlichen Netz bereitstellen, eine Rolle. Vor allem in Rechnernetzen sind zentrale Strukturen, wie sie noch in Netzen mit der *Systems Network Architecture* [151] dominierend sind, durch verteilte Systeme weitgehend verdrängt worden. Teilnehmerzugangnetze sind meist baumförmig strukturiert, wobei die Endsysteme (*Terminals*) über das Zugangsnetz an einen zentralen Vermittlungsknoten angeschlossen sind. Die Kommunikation findet nur zwischen dem Vermittlungsknoten und den jeweiligen Endsystemen statt; die Verkehrslenkung ist damit durch die Topologie festgelegt.

Der Entwurf zentraler Strukturen, etwa als Zugangnetze für MSS, soll im Rahmen dieser Arbeit nicht weiter betrachtet werden. Für eine ausführliche Darstellung der Entwurfsverfahren sei auf die Literatur verwiesen [92].

3.3.5.2 Dezentrale Strukturen

In diesem Abschnitt sollen die wichtigsten Probleme beim Entwurf von paketvermittelnden Kommunikationsnetzen dargestellt werden. Dabei liegt der Problembeschreibung meist eine ideale Betrachtungsweise zu Grunde, die die Anforderungen bei realen Planungsaufgaben weitgehend vernachlässigt (vgl. auch [92, 97]).

Beim Entwurf eines Paketvermittlungsnetzes sind im wesentlichen drei Entscheidungen zu treffen [54, 59, 92]:

1. Die Wahl der Übertragungsabschnitte, die die Netzknoten miteinander verbinden (Entwurf der Netzstruktur).
2. Die Wahl der Pfade, über die der Verkehrsfluß von einer Verkehrsquelle zur -senke gelenkt wird (Problem der statischen Verkehrslenkung).
3. Die Wahl der Kapazität, die zur Übermittlung des Verkehrsflusses benötigt wird (Problem der Kapazitätszuweisung).

Parameter	Kapazitätszuweisung	Verkehrslenkung	Topologischer Entwurf
Gegeben	f_{ij}, x_{ij}	$\gamma_{ij}, C_{ij}, x_{ij}$	γ_{ij}
Zielfunktion	$\min g = g(C_{ij}, d_{ij})$	$\min g = g(C_{ij}, d_{ij})$	$\min g = g(C_{ij}, d_{ij})$
Variablen	C_{ij}	f_{ij}	C_{ij}, f_{ij}, x_{ij}
Nebenbedingungen	$f_{ij} \leq C_{ij}$, Einhaltung der geforderten Netzgüte	$f_{ij} \leq C_{ij}$, Einhaltung der geforderten Netzgüte	$f_{ij} \leq C_{ij}$, Einhaltung der geforderten Netzgüte

Tabelle 3.1: Optimierungsprobleme in Paketvermittlungsnetzen

Bei der Bestimmung dieser Parameter sind eine Reihe von Nebenbedingungen zu beachten, die sich beispielsweise aus Dienstgüteparametern, Beschränkungen der zulässigen Pfade bei der Verkehrslenkung oder durch die Kapazität der Übertragungsabschnitte und der Vermittlungsrechner ergeben. Die gleichzeitige Festlegung dieser drei Netzparameter wird als topologischer Entwurf bezeichnet.

Zur Verdeutlichung werden die drei Probleme anhand eines graphentheoretischen Modells erläutert. Es wird angenommen, daß ein Graph G mit n Knoten gegeben sei, dessen Knoten die Vermittlungsstellen und dessen Kanten die Übertragungsabschnitte im Netz repräsentieren. Die oben genannten Größen (Übertragungsabschnitte, Verkehrsfluß, Kapazität) können mit Hilfe sog. Entscheidungsvariablen folgendermaßen beschrieben werden: Die Variable x_{ij} gibt an, ob eine Kante (i, j) vorhanden ist. Mit C_{ij} wird die Kapazität einer Kante (i, j) und mit f_{ij} der Fluß über eine Kante (i, j) bezeichnet, wobei $i, j = 1 \dots n$. Der Fluß über eine Kante ist stets gleich der Summe der Einzelflüsse, die diese Kante benutzen, und darf die Kapazität dieser Kante nicht überschreiten.

Typischerweise soll das Netz kostenoptimal für eine gegebene Zielfunktion g dimensioniert werden. Es kann aber auch eine Optimierung der Topologie hinsichtlich bestimmter Dienstgüteparameter, wie z. B. der mittleren Transferzeit, erwünscht sein. Tabelle 3.1 zeigt die Formulierung der drei Probleme und die jeweiligen Freiheitsgrade im Vergleich. Beim Problem der Kapazitätszuweisung sind die Verkehrsflüsse für jeden Übertragungsabschnitt (i, j) gegeben, in den beiden anderen Fällen sind diese erst noch aus den gegebenen Verkehrsanforderungen γ_{ij} , die den Verkehrsfluß zwischen zwei beliebigen Netzknoten i und j , beschreiben, zu ermitteln. Ist die Zielfunktion eine Funktion der Netzkosten, hängt diese im allgemeinen von der Größe der Kapazität eines Übertragungskanals und von der Entfernung d_{ij} ab. Neben der Beschränkung des Verkehrsflusses wird meist auch die Beschränkung der Transferzeit gefordert. Bei Anforderungen an die Verfügbarkeit sind verschiedene Netzparameter wie die Konnektivität des Netzes oder die Ausfallwahrscheinlichkeit zu betrachten.

Das Problem der Kapazitätszuweisung ist für eine lineare Kostenfunktion und die Annahme,

daß die Kapazität eine kontinuierliche Variable darstellt, exakt lösbar [13]. Wenn angenommen wird, daß die Kostenfunktion nichtlinear und konkav ist, kann eine exakte Lösung nur noch für spezielle Formen der Kostenfunktion ermittelt werden [92]. In der Realität kann jedoch oftmals nur aus einer diskreten Menge von Übertragungsabschnitten mit fester Kapazität ausgewählt werden, z. B. Mietleitungen oder ISDN-Kanäle. Bei umfangreichen Problemen kann der Aufwand, die Lösung zu berechnen, dann sehr groß sein; damit können Näherungsverfahren notwendig werden, die lediglich lokal optimale Lösungen erzielen [92].

Das Problem der Kapazitätszuweisung tritt in der Regel nicht als isoliertes Problem auf, sondern als Teilproblem allgemeiner Optimierungsaufgaben. Üblicherweise wird die statische Verkehrslenkung auf der Basis vorgegebener Kapazitäten der Übertragungsabschnitte optimiert [131]. Daher ist im allgemeinen die simultane Lösung beider Probleme anzustreben. Das ist ein schwieriges kombinatorisches Optimierungsproblem und ebenfalls nur heuristisch zu lösen [57, 59].

Dies sei an einem stark vereinfachten Beispiel verdeutlicht. Betrachtet werde ein Übertragungsabschnitt zwischen zwei Netzknoten mit k Kanälen, deren Kapazität jeweils C_i , $i = 1 \dots k$, sei. Die Kapazitäten C_i sollen so gewählt werden, daß $\sum_{i=1}^k C_i$ minimal wird. Dabei soll der Verkehrsfluß auf die k Kanäle aufgeteilt werden, so daß die mittlere Transferzeit einen bestimmten Wert nicht überschreitet. Die optimale Lösung umfaßt nur einen Kanal, der genug Kapazität besitzt, um die Nebenbedingung für die mittlere Transferzeit zu erfüllen. Wird der Verkehrsfluß auf mehrere Kanäle verteilt, steigt die mittlere Transferzeit. Die Auswahl des Kanals spielt keine Rolle, so daß dieses Problem k lokale Minima besitzt. Hinzu kommt, daß diese Lösung eventuellen Anforderungen an die Verfügbarkeit nicht genügt, und damit die Zahl der lokalen Minima weiter erhöht wird.

Somit ist auch das Problem des topologischen Entwurfs eines Paketvermittlungsnetzes auf Grund der vielen Freiheitsgrade nur bei geringer Dimension exakt lösbar. Praktisch anwendbare Verfahren basieren durchgängig auf heuristischen Algorithmen [92].

Kapitel 4

Ein Planungsmodell für MAN-Vermittlungssysteme

Das folgende Kapitel behandelt die wesentlichen Modellierungsaspekte, die zur algorithmischen Lösung des MSS-Entwurfsproblems benötigt werden. Eine formale Beschreibung dieses Planungsproblems soll ferner Aufschluß darüber geben, ob eine Lösung mit Hilfe von Standardverfahren möglich ist, oder ob ein spezifisches Lösungsverfahren entwickelt werden muß.

4.1 Übersicht

Die prinzipiellen Planungsmodelle für Paketvermittlungsnetze, wie sie in Abschnitt 3.3.5 beschrieben wurden, sind für reale Planungsaufgaben, insbesondere bei lokalen und regionalen Hochgeschwindigkeitsnetzen, nicht ausreichend. Daher wurden in den letzten Jahren neue Modelle und Verfahren entwickelt, die sich hauptsächlich auf die Planung und den Entwurf lokaler Netze bzw. den Entwurf von *Backbone*-Netzen im Weitverkehrsbereich zur Datenübermittlung zwischen verteilten Rechnersystemen konzentrieren. Für MANs sind Planungsmodelle erst in Ansätzen entwickelt bzw. nur für Teilprobleme bekannt.

Beim topologischen Entwurf lokaler Netze wird primär das Ziel verfolgt, einen Verbund von LANs, die über *Bridges* oder *Repeater* gekoppelt sind, so zu strukturieren, daß die Last möglichst gleichmäßig auf die Netzsegmente verteilt ist, und außerdem die mittlere Transferzeit zwischen den angeschlossenen Stationen minimiert wird. Dieses Ziel kann mit verschiedenen Methoden erreicht werden:

- In [63] wird ein Verfahren zur optimalen Platzierung von *Bridges* bei gegebener Netzstruktur beschrieben. Bei diesem Verfahren werden die durch die Platzierung der

Bridges möglichen Netzsegmente als Kanten eines Graphen modelliert, denen wiederum Gewichte, die der mittleren Pakettransferzeit eines Netzsegmentes entsprechen, zugeordnet werden. Die Knoten dieses Graphen stellen die *Bridges* dar. Die Entscheidung, an welchen Positionen tatsächlich *Bridges* eingesetzt werden sollen, kann durch die Lösung eines Standardproblems der ganzzahligen Optimierung (*Set Partitioning*) getroffen werden. In das Optimierungsmodell können auch die Kosten von *Bridges* miteinbezogen werden.

- Ist die physikalische Netzstruktur noch nicht festgelegt, kann diese als *Minimum Spanning Tree* durch ein Optimierungsverfahren ermittelt werden. Dazu gibt es verschiedene Ansätze, die sich auf Grund der Komplexität des Problems auf Verfahren des *Simulated Annealing* abstützen, und damit entweder eine Minimierung der mittleren Systemtransferzeit [41] oder eine Minimierung der Gesamtkosten [49, 50] anstreben. Kosten entstehen sowohl durch Netzkomponenten als auch durch Mietleitungen, die verteilte LAN-Standorte über den Weitverkehrsbereich hinweg verbinden. Ein vergleichbarer Ansatz, allerdings mit einem deterministischen Näherungsverfahren, findet sich in [160].
- Ist lediglich eine Lastaufteilung auf der Basis der Verkehrsmatrix gesucht, kann das in [94] skizzierte Verfahren verwendet werden. Dabei wird versucht, den internen Verkehr von Stationen auf dem gleichen Netzsegment zu maximieren, und gleichzeitig die Last unter den Netzsegmenten gleichmäßig aufzuteilen. Die tatsächliche Netzstruktur kann mit diesem Verfahren aber nicht festgelegt werden.
- Ein Entwurfsverfahren, das auf Genetischen Algorithmen beruht, beschreibt [40]. Dieser Ansatz löst gleichzeitig das Problem der Lastteilung und erzeugt eine *Minimum-Spanning-Tree-Topologie*.

Meist werden LANs als Knoten und ihre Verbindungen über Netzkoppeleinheiten als Kanten eines Graphen modelliert und daraus die Verkehrsflüsse im Netz, die Gesamtkosten und die Dienstgüteparameter abgeleitet. Sie bilden die Grundlage, um die Zielfunktion und die Nebenbedingungen eines Optimierungsmodells zu formulieren. Lösungen für die genannten Optimierungsprobleme lassen sich in der Regel nur mit heuristischen Verfahren erzielen.

Mit dem Entstehen neuer Netzinfrastrukturen, die LANs über den regionalen und den Weitverkehrsbereich hinweg verbinden sollen, treten auch Planungsprobleme bei MANs und bei Paketvermittlungsnetzen zur Unterstützung von SMDS in den Vordergrund. In [23, 24] werden verschiedene Planungsaspekte bei MANs mit DQDB-Technik diskutiert. Zwar sind auch Planungsverfahren in Ansätzen beschrieben, über die Analyse vorgegebener Konfigurationen gehen sie jedoch nicht hinaus. Für eine spezielle Konfiguration gekoppelter MAN-Teilnetze wird in [128] ein lineares Optimierungsmodell präsentiert, was zur Durchsatzoptimierung bei günstiger Wahl der Protokollkonfiguration geeignet ist. Eine Lösung wird jedoch nur für einfache Spezialfälle angegeben. Modelle und Algorithmen zur Dimensionierung und

Erweiterung von SMDS-Paketvermittlungsnetzen finden sich in [106, 107]; spezifische MAN-Strukturen werden nicht berücksichtigt. Ein Modell zur Planung von Hochgeschwindigkeitsnetzen mit ringförmiger Topologie wie beispielsweise FDDI beschreibt Leblanc [101]. Insbesondere dient das Modell zur Abschätzung des Verkehrs und, darauf aufbauend, zur Netzerweiterung einzelner, ringförmiger Teilnetze. Ein weiteres Verfahren zur Planung ringförmiger Netze ist in [30] enthalten.

Für paketvermittelnde Weitverkehrsnetze wurden die Planungsverfahren, wie sie in Abschnitt 3.3.5 beschrieben sind, weiterentwickelt, so daß auch Probleme mit umfangreichen Nebenbedingungen, die bei praktisch relevanten Planungsaufgaben erfüllt sein müssen, zumindest heuristisch gelöst werden können [3, 57, 138]. Ferner wurden Verfahren entwickelt, die eine Planung von Datennetzen über mehrere Planungsperioden ermöglicht [35, 36], oder die die optimale Platzierung von Netzkoppeleinheiten bei der Zusammenführung zweier existierender Netze unterstützt [14, 105]. Eine Anwendung von Netzplanungsverfahren auf den Entwurf virtuell privater Netze beschreibt [127].

Zusammenfassend läßt sich feststellen, daß geeignete Modelle und Verfahren zur Planung und zum Entwurf von MAN-Vermittlungssystemen bisher nicht entwickelt wurden. Ein solches Planungsmodell sollte insbesondere

- Aspekte berücksichtigen, die aus Anforderungen im Netzbetrieb resultieren,
- die Planung von hierarchischen MAN-Vermittlungssystemen unterstützen und
- für MAN-Vermittlungssysteme mit einer nennenswerten Zahl von Stationen Lösungen liefern.

Die genannten Aspekte werden im weiteren Verlauf dieses Kapitels ausführlich behandelt.

4.2 Modellierung

MAN-Vermittlungssysteme sind durch Eigenschaften charakterisiert, wie sie teilweise auch in Weitverkehrsnetzen und in lokalen Netzen bestimmend sind. Ein wichtiger Faktor ist die geographische Ausdehnung, die zum einen ein Parameter der Netzkosten ist, wenn MAN-Vermittlungssysteme im öffentlichen Netz eingesetzt werden, oder Verbindungsleitungen teilweise über öffentliches Gelände führen (Mietleitungen), und die zum anderen Einfluß auf die Leistungsfähigkeit hat. MAN-Vermittlungssysteme sollten, bei ähnlichen Anforderungen wie in Weitverkehrsnetzen, eine hohe Verfügbarkeit besitzen. Darüber hinaus sind Kriterien zu beachten, wie sie bei der Planung lokaler Netze auftreten. Typischerweise sollte die Last in einem MAN-Vermittlungssystem möglichst gleichmäßig aufgeteilt oder ein Teil der Last auf *Backbone*-Netze verlagert werden. Ferner sollte die mittlere Transferzeit zwischen Endteilnehmern unterhalb einer gegebenen Schranke liegen. Schließlich weisen die Netzelemente

hinsichtlich ihrer Kosten oder der Kapazität der anzuschließenden Übertragungssysteme eine ganze Reihe von Varianten auf, aus denen die jeweils optimale Komponente gewählt werden sollte.

Diese Anforderungen müssen bei einem MAN-Planungsverfahren berücksichtigt werden mit dem Ziel, eine möglichst kostengünstige Netzstruktur zu entwerfen. In den folgenden Abschnitten werden zunächst die prinzipiellen Modellannahmen und dann das vollständige Optimierungsmodell, dessen Nebenbedingungen die getroffenen Annahmen widerspiegeln, beschrieben.

4.2.1 Netzelemente und Geographie

Grundsätzlich kann ein MAN-Vermittlungssystem (MSS) hierarchisch aufgeteilt werden (vgl. Abschnitt 2.2.4). Ein MAN-Vermittlungssystem umfaßt einen Zugangsnetzbereich und das eigentliche Transportnetz, das wiederum in weitere Hierarchiestufen gegliedert werden kann. Dabei erbringen die Netzelemente je nach Ebene unterschiedliche Funktionen. Die Stationen (*Edge Gateways*), an die die Zugangsnetze angeschlossen sind, konzentrieren den Verkehr aus dem Zugangsnetz und werden als Quellen und Senken des Verkehrs betrachtet. Das Zugangsnetz selbst wird nicht näher einbezogen, da seine Struktur und Konfiguration meist unabhängig vom Transportnetz und nach ökonomischen Kriterien oder Sicherheitsanforderungen festgelegt wird. Daher wird davon ausgegangen, daß die Zuordnung der Teilnehmer zu einem Zugangsnetz und die Zuordnung des Zugangsnetzes zu den Stationen des MSS vorgegeben ist. Als weitere Netzelemente stehen *Router* zur Verfügung, um den Verkehrsfluß weiterzuleiten. Allgemein sind die Netzelemente durch bestimmte Eigenschaften charakterisiert, wie beispielsweise Kapazität (maximaler Durchsatz), Verzögerungsverhalten oder Kosten. Um eine möglichst hohe Verfügbarkeit zu gewährleisten, werden nur Teilnetze in der *Looped-Bus*-Konfiguration betrachtet.

Die Netzelemente werden an bestimmten Standorten plaziert, die in der Regel durch den Netzbetreiber vorgegeben sind, und deren geographische Daten in Form von Koordinaten gegeben sind. Es wird angenommen, daß die Standortkoordinaten bezüglich eines orthogonalen Koordinatensystems festgelegt sind. Die Standortentscheidung kann auch als Optimierungsproblem formuliert werden, um die optimalen Standorte unter Berücksichtigung der räumlichen Verteilung der Teilnehmer und ihres anzunehmenden Verkehrsaufkommens zu bestimmen (s. z. B. [10]). Dieses Problem soll hier nicht näher untersucht werden.

Standorte sind durch Trassen verbunden, die mögliche Wege für die MAN-Übertragungsabschnitte und damit auch Anschlußpunkte an die Übertragungsleitungen repräsentieren (Trassennetzgraph). Der Trassennetzgraph unterliegt Beschränkungen, da nicht alle Standorte direkt verbunden und damit nur bestimmte Wege erlaubt sind (s. Bild 4.1).

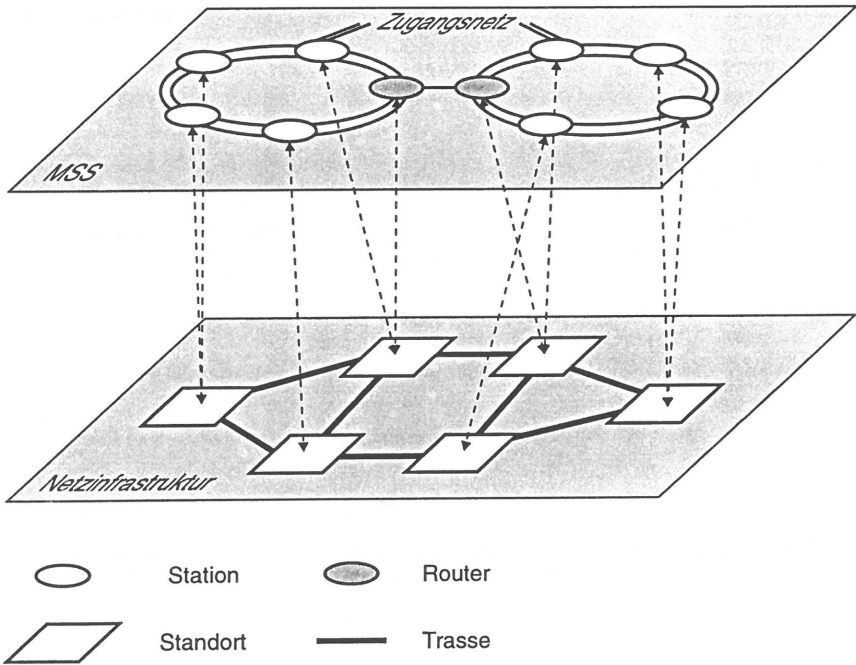


Bild 4.1: Zuordnung zwischen den Komponenten des MSS und der Netzinfrastruktur

MAN-Teilnetze sind in der Regel geographisch ausgedehnt und können nicht als einzelner Knoten modelliert werden, wie das in den meisten Fällen beim Entwurf von LAN-Topologien möglich ist. Daher wird der Abstand zwischen zwei Teilnetzen durch den Abstand zwischen denjenigen Stationen dieser Teilnetze definiert, die die geringste Entfernung zueinander aufweisen. Bei der Modellierung wird das *Network Management Center* (NMC) nur dadurch berücksichtigt, daß die Zahl der Netzelemente in einem MSS, die maximal von einem NMC verwaltet werden können, beschränkt ist.

4.2.2 Kostenmodell

Wichtigen Einfluß auf die Netzstruktur hat das Kostenmodell. Zum einen umfaßt es die Anschaffungs-, Installations- und Betriebskosten für Vermittlungs- und Übertragungseinrichtungen, zum anderen die zeitabhängigen Mietkosten von Übertragungsleitungen, also die Tarife für digitale Festverbindungen. Diese können im allgemeinen durch eine abschnittsweise lineare Funktion in Abhängigkeit von Entfernung und Leitungskapazität dargestellt

werden. Die Anschaffungs- und Installationskosten werden meist durch längen- und kapazitätsunabhängige Grundkosten berücksichtigt [10].

Typischerweise werden zum Aufbau eines MAN digitale Festverbindungen angemietet. Damit überwiegen die laufenden Betriebskosten die einmalig anfallenden Anschaffungs- und Installationskosten erheblich. Daher wird eine Kostenfunktion gewählt, die eine abschnittsweise lineare Tarifierungskurve annähert und damit einfach auszuwerten ist. Für diese Funktion werden lediglich ein Fixkostenanteil K_0 und variable Kosten berücksichtigt, die mit der Entfernung gemäß einer konkaven Funktion zunehmen:

$$K_l = K_l^0 + \kappa_l \cdot \sqrt{d}, \quad (4.1)$$

mit dem Normierungsfaktor κ_l und der Entfernung d . Die Funktion kann für unterschiedliche Leitungstypen l mit gegebener Kapazität (z. B. 34 Mbit/s) auch unterschiedlich parametrisiert sein. Die Fixkosten K_l^0 enthalten einen Sockelbetrag für das Mieten einer Leitung; die variablen Kosten umfassen die entfernungsabhängigen Mietkosten einer digitalen Festverbindung.

Dieses Modell hat den Vorteil, daß es mit relativ wenig Aufwand eine Kostenbewertung während der Optimierung liefert. Gegebenenfalls muß jedoch die entworfene Netzstruktur mit Hilfe der Tarifierungsrichtlinien auf die realen Kosten hin untersucht werden.

4.2.3 Verkehrslenkung und Flußanalyse

Die Ausgangsbasis der Flußanalyse bilden die Verkehrsbeziehungen, die durch eine Verkehrsmatrix beschrieben werden. Ein Element der Verkehrsmatrix repräsentiert den mittleren Verkehrsfluß, beispielsweise die mittlere Paketrage, die zwischen Endteilnehmern zu einer bestimmten Tageszeit übermittelt wird. Üblicherweise wird der Verkehrsfluß zur Hauptverkehrsstunde bestimmt. Die Verkehrsmatrix kann aus Messungen erstellt oder durch Schätzverfahren berechnet werden. Schätzverfahren sind dann notwendig, wenn keine genauen Daten aus Messungen vorliegen oder eine neue Netztechnik eingeführt wird. Liegt eine hierarchische Netzstruktur vor, kann die Verkehrsmatrix komprimiert und für die Netzelemente auf einer höheren Hierarchiestufe dargestellt werden [161].

Um die Last auf den einzelnen Übertragungsabschnitten bzw. in den Netzelementen zu bestimmen, muß die Verteilung der Verkehrsflüsse im MSS ermittelt werden. Dazu werden die Verkehrslenkungsschemata herangezogen, wie sie bereits in Abschnitt 2.2.4.3 beschrieben wurden. Der Flußanalyse liegt die Annahme zu Grunde, daß sich im zeitlichen Mittel für eine gegebene Planungsperiode in einem entsprechend der Planung realisierten Netz ungefähr die Verkehrsverteilung einstellt, die den Planungsvorgaben entspricht, auch wenn im Netzbetrieb ein adaptives Verkehrslenkungsverfahren verwendet wird (vgl. dazu [59]). Das

kurzfristige, dynamische Netzverhalten kann allerdings nur simulativ untersucht werden und wird nachfolgend nicht weiter betrachtet.

Wie bereits in Abschnitt 2.2.4.3 ausgeführt, wird das Verkehrslenkungsschema in einem MSS durch eine Menge von *Shortest Path Trees* realisiert. Diese Bäume werden auch als *Destination Spanning Trees* (DSTs) bezeichnet, da bei einer Darstellung der Teilnetze als Knoten eines Graphen ein DST alle Teilnetze enthält, und die Wurzel des Baums das Zielnetz repräsentiert (vgl. Bild 4.2). Mit Hilfe des DST können dann die Pfade von einer Verkehrsquelle bis zur -senke, d. h. die Quelle-Ziel-Pfade zwischen je zwei Stationen, konstruiert werden. Die Menge aller Quelle-Ziel-Pfade wird mit \mathcal{R} bezeichnet.

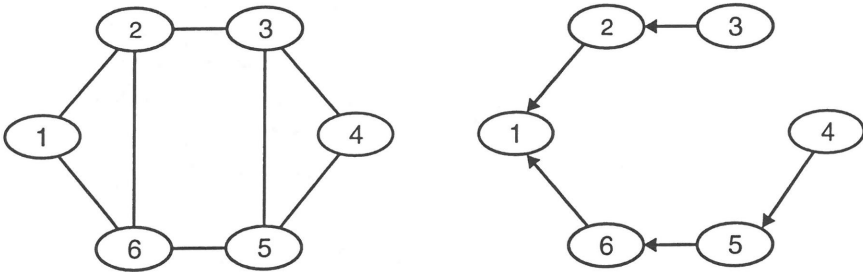


Bild 4.2: Teilnetzgraph und *Destination Spanning Tree* für Knoten 1

Zur Berechnung des DST aus der Netztopologie werden in der Regel statische Größen (Metriken) wie die Anzahl der durchlaufenen Netzknoten (*Hops*), die Entfernung zwischen Teilnetzen oder der Reziprokwert der Kapazität eines Übertragungsabschnittes verwendet. Gesucht sind z. B. DSTs, deren Pfade nur eine minimale Zahl von Teilnetzen enthalten (*Minimum Hop*). Die rechte Darstellung von Bild 4.2 zeigt ein Beispiel für einen DST für Zielknoten 1. In diesem Fall ist auch eine zweite Lösung denkbar, wenn statt der Kante zwischen Knoten 4 und 5 die Kante zwischen Knoten 3 und 4 gewählt wird.

Bei Verwendung der *Minimum-Hop*-Metrik müssen weitere Gewichtungsfaktoren zur Unterscheidung von Pfaden mit der gleichen Zahl von durchlaufenen Teilnetzen herangezogen werden. Ein möglicher Gewichtungsfaktor ist die Anzahl der DSTs, die bereits über ein Teilnetz führen. Dieser Faktor bewertet in einfacher Weise die relative Belastung eines Teilnetzes. Verwendet man ein statisches Verkehrslenkungsverfahren in einem MSS, kann die *Minimum-Hop*-Metrik auch durch zusätzliche Gewichtungsfaktoren auf die geplante Verkehrslast angepasst werden.

Nicht alle Metriken eignen sich für ein Verkehrslenkungsschema in einem MSS. So kann eine Metrik, die von der Entfernung zwischen den Teilnetzen abhängig ist, unerwünschte Ergebnisse hervorrufen. Wird ein Verfahren zur Berechnung der DSTs mit Hilfe dieser Metrik

angewendet, ist es durch die Ausdehnung der Teilnetze möglich, daß die aufsummierte Entfernung zwischen Start- und Zielknoten nur einen Teil der tatsächlichen Pfadlänge darstellt. Wenn die Kapazität der Router in einem MSS gleich groß ist, kommt auch eine kapazitätsabhängige Metrik nicht in Frage. Für die Bestimmung der Verkehrsflüsse wird daher im folgenden von einer *Minimum-Hop*-Metrik ausgegangen.

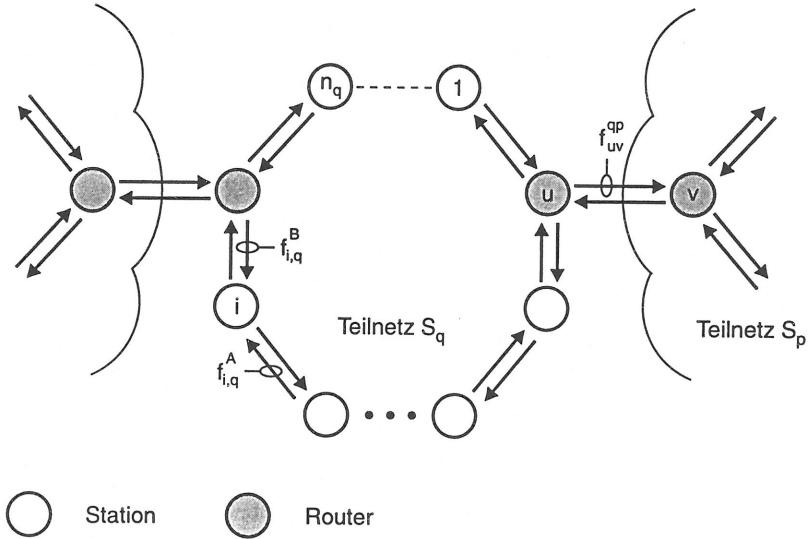


Bild 4.3: Verkehrsfluß im MSS

Sind die Netztopologie und das Verkehrslenkungsschema gegeben, kann der Verkehrsfluß berechnet werden. Dazu soll folgendes Modell betrachtet werden (Bild 4.3). Ein Teilnetz S_q enthält n_q Stationen und Router in der *Looped-Bus*-Konfiguration, wobei nach Bild 4.3 Station 1 bzw. Station n_q den Beginn von Bus A bzw. Bus B markieren. Zwischen diesen beiden Stationen besteht nur eine physikalische Verbindung, die, wie bereits erwähnt, den Bus zu einem Ring schließt, aber für den Verkehrsfluß eine logische Unterbrechung darstellt. Der Verkehrsfluß auf dem Bus kann als gerichteter Graph dargestellt werden, wobei jeder Knoten stromabwärts den Verkehrsfluß auf dem Bus erhöht. Damit lautet der Verkehrsfluß für Bus A bzw. Bus B:

$$f_{i,q}^A = \sum_{j=1}^{i-1} \sum_{k=i}^{n_q} \lambda_{jk}^q \quad (4.2)$$

$$f_{i,q}^B = \sum_{j=i+1}^{n_q} \sum_{k=1}^i \lambda_{jk}^q \quad (4.3)$$

Die Größe λ_{jk}^q kann mit Hilfe der Verkehrsflüsse γ_{ij} , die zwischen zwei Stationen auftreten, und durch Faktoren ausgedrückt werden, die aus der Menge der Pfade \mathcal{R} zu bestimmen sind¹:

$$\lambda_{jk}^q = \begin{cases} \gamma_{jk}, & \text{falls } j, k \in S_q \\ \sum_{n=1}^M \sum_{\forall l \in S_n} \gamma_{jl} \cdot r_{jkl}, & \text{falls } j, k \in S_q, S_n \neq S_q \\ \sum_{n=1}^M \sum_{\forall l \in S_n} \gamma_{lk} \cdot r_{ljk}, & \text{falls } j, k \in S_q, S_n \neq S_q \\ \sum_{n=1}^M \sum_{\forall l \in S_n} \sum_{p=1}^M \sum_{\forall m \in S_p} \gamma_{lm} \cdot r_{jkm}^{lm}, & \text{falls } j, k \in S_q, S_n \neq S_p \neq S_q \end{cases} \quad (4.4)$$

mit $\gamma_{jj} = 0$ und

$$r_{jkl} = \begin{cases} 1, & \text{falls } k \in R_{jl} \text{ und } R_{jl} \in \mathcal{R} \\ 0, & \text{sonst} \end{cases} \quad (4.5)$$

$$r_{jkm}^{lm} = \begin{cases} 1, & \text{falls } j, k \in R_{lm} \text{ und } R_{lm} \in \mathcal{R} \\ 0, & \text{sonst} \end{cases} \quad (4.6)$$

wobei S_n und S_p Teilnetze bezeichnen, und M die Zahl der Teilnetze ist. Die vier Fälle in Gleichung (4.4) stellen der Reihe nach den lokalen Verkehrsfluß in einem Teilnetz, den Verkehrsfluß aus einem bzw. in ein Teilnetz und schließlich den Transitverkehr durch ein Teilnetz dar.

Über eine *Router-Router-Kopplung* (u, v) zwischen zwei Teilnetzen berechnet sich der Verkehrsfluß folgendermaßen:

$$f_{uv}^{qp} = \sum_{m=1}^M \sum_{\forall i \in S_m} \sum_{n=1}^M \sum_{\forall j \in S_n} \gamma_{ij} \cdot r_{uv}^{ij}, \quad \forall u \in S_q, v \in S_p, q \neq p. \quad (4.7)$$

wobei S_m und S_n Teilnetze bezeichnen und r_{ij}^{kl} sich wiederum aus der Menge der Pfade ableitet (vgl. Gleichung (4.6)).

4.2.4 Warteschlangenmodell

Die Netzstruktur kann in ein Warteschlangennetz abgebildet werden, um daraus wichtige Kenngrößen, wie z. B. die Transferzeit für einen Pfad oder für das gesamte Netz, zu berechnen. Da die aus dem Warteschlangenmodell abgeleiteten Berechnungsvorschriften innerhalb eines Optimierungsalgorithmus eingesetzt werden, der gegebenenfalls sehr viele Topologien untersucht, sollte die Berechnung der einzelnen Komponenten nur wenig Zeit beanspruchen. Daher ist ein sinnvoller Kompromiß zwischen einer realistischen Modellierung und dem resultierenden Rechenaufwand zu treffen. Bei allen Modellen wird die Annahme getroffen,

¹Die Indizes, die die Zugehörigkeit einer Station zu einem Teilnetz kennzeichnen, werden aus Gründen der Übersichtlichkeit hier weggelassen.

daß die Verkehrsströme, die an den jeweiligen Netzelementen ankommen, durch Poisson-Ankunftsprozesse approximiert werden können. Dies ist auf Grund der Überlagerung mehrerer unabhängiger Verkehrsströme in einem MAN-Vermittlungssystem gerechtfertigt. Vergleichbare Annahmen werden ganz allgemein bei der Analyse von komplexen Warteschlangennetzen getroffen [98]. Nachfolgend werden zunächst die einzelnen Komponenten und dann das Gesamtmodell beschrieben.

Eine wichtige lastabhängige Komponente ist das Verkehrsmodell, mit dem die Buszugriffszeit einer Station in einem MAN-Teilnetz bestimmt werden kann. Dafür stehen eine Vielzahl von Modellen aus der Literatur zur Verfügung, die zum Teil ausführlich untersucht wurden (vgl. [120]). Für die Einbindung in einen Optimierungsalgorithmus kommen jedoch nur Approximationen in Frage, die wenig rechenintensiv sind. Hier bietet das Modell nach Bisdikian [15] eine brauchbare Approximation. Ein prinzipieller Vergleich verschiedener Analysemodelle mit Simulationsdaten hat gezeigt, daß auch detaillierte Modelle das reale Verhalten eines DQDB-Systems nur näherungsweise nachbilden [114].

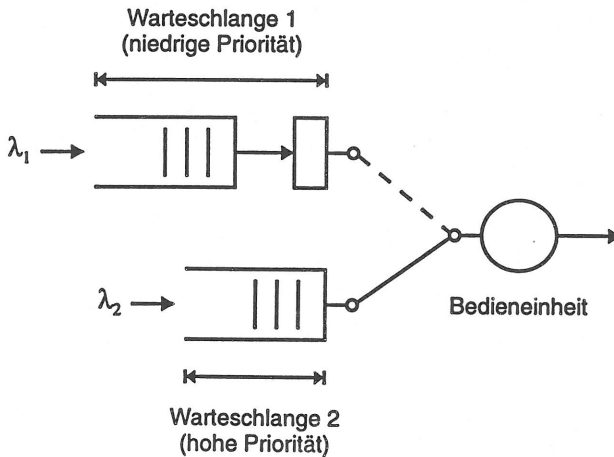


Bild 4.4: Warteschlangenmodell mit zwei Prioritäten

Das Modell nach Bisdikian umfaßt ein Prioritätssystem mit zwei Warteschlangen, die zwei unterschiedliche Verkehrsströme mit hoher bzw. niedriger Priorität zwischenspeichern und von einer Bedieneinheit abgefertigt werden (s. Bild 4.4). Die beiden Warteschlangen besitzen jeweils eine unbegrenzte Zahl an Pufferplätzen.

Die Bedienstrategie ist so gestaltet, daß der Verkehrsstrom, der eine hohe Priorität besitzt, bevorzugt bedient wird, was der Abfertigung aller früher eingetroffenen Anforderungen

gemäß dem DQ-Protokoll entspricht. Die Bedieneinheit schaltet dazu zwischen den beiden Warteschlangen nach folgender Strategie mit vernachlässigbarer Schaltzeit um:

- Die Bedieneinheit schaltet von Warteschlange 1 zu Warteschlange 2 um, nachdem ein Paket bedient wurde. Das ist in Bild 4.4 durch den ersten Pufferplatz der Warteschlange 1 angedeutet.
- Ist Warteschlange 1 noch nicht leer, werden nur die Pakete in Warteschlange 2 bedient, die sich zum Umschaltzeitpunkt in Warteschlange 2 befinden. Danach wird zu Warteschlange 1 zurückgeschaltet.
- Ist Warteschlange 1 leer, werden alle Pakete in Warteschlange 2 bedient, die vor einer erneuten Paketankunft in Warteschlange 1 in Warteschlange 2 eintreffen.

Die beschriebene Abfertigungsdisziplin wird *quasi-gated* genannt. In [15] wird gezeigt, wie die mittleren Wartezeiten des Prioritätssystems approximativ (als obere und untere Schranke) berechnet werden können. Die Bedienzeiten der beiden Wartesysteme können unterschiedlich gewählt werden und nach einem allgemeinen stochastischen Prozeß verteilt sein.

Damit ergibt sich eine Approximation für die Buszugriffszeit einer DQDB-Station, wenn die lokale Warteschlange, die die zu sendenden Segmente einer Station speichert, als Warteschlange 1 mit niedriger Priorität modelliert wird, und alle Reservierungswünsche von anderen Stationen, die durch diese Station registriert werden, durch Warteschlange 2 mit hoher Priorität nachgebildet werden. Als Ankunftsprozeß wird jeweils ein Poisson-Prozeß mit der Paketankunftsrate λ_1 bzw. λ_2 angenommen, die mit Hilfe von Gleichung (4.4) bestimmt werden können. Unter Verwendung der oberen Schranke für die mittlere Transferzeit des Wartesystems 1 (vgl. [15]) kann die mittlere Buszugriffszeit berechnet werden:

$$t_b = \frac{w_0 - (1 - \rho) \frac{\rho_2}{\rho_1} \eta h_1}{(1 - \rho_2)(1 - \rho)} + h_1, \quad (4.8)$$

wobei der Faktor

$$\eta = \left(1 - \frac{(1 - \rho_1)(1 - p_i \exp(-\lambda_1 t_z))}{(1 - p_i) \exp(-\lambda_1 t_z)}\right) \rho_2 \quad (4.9)$$

und

$$w_0 = \frac{1}{2}(\lambda_1 h_1^{(2)} + \lambda_2 h_2^{(2)}) \quad (4.10)$$

so gewählt werden, daß die Bedienzeit für beide Verkehrsströme geometrisch verteilt ist. Die Auslastung durch Klasse 1 sei ρ_1 und die durch Klasse 2 entsprechend ρ_2 . Dann muß die Stabilitätsbedingung $\rho = \rho_1 + \rho_2 < 1$ erfüllt sein. Die Erwartungswerte bzw. zweiten Momente der fiktiven Bedienzeiten sind mit h_1 und h_2 bzw. $h_1^{(2)}$ und $h_2^{(2)}$ bezeichnet; die fiktive Bedienzeit beinhaltet neben der Dauer eines Zeitschlitzes für die eigentliche Bedienung auch die Zeit, die durch das Warten auf einen freien Zeitschlitz entsteht. Die Belegungswahrscheinlichkeit eines Zeitschlitzes an Station i sei p_i , die Dauer eines Zeitschlitzes sei t_z .

Die Annahme einer geometrischen Verteilung für die Bedienzeit einer DQDB-Station wurde bereits von Tran-Gia vorgeschlagen [154]. Die Verteilung erhält man mit

$$P\{T_h = k \cdot t_z\} = p_i^{k-1}(1 - p_i). \quad (4.11)$$

Die Zufallsvariable T_h kann als Zeitintervall zwischen zwei freien Zeitschlitzten interpretiert werden, das an Station i beobachtet wird. Wird nur eine Busrichtung betrachtet, ist die Belegungswahrscheinlichkeit p_i eines Zeitschlitzes an Station i in Teilnetz q durch

$$p_i = \sum_{j=1}^{i-1} \sum_{k=i+1}^n \lambda_{jk}^q t_z, \quad (4.12)$$

gegeben. Die Paketankunftsrate λ_{jk}^q berechnet sich nach Gleichung (4.4). Im Prinzip könnte die Belegungswahrscheinlichkeit auch durch andere Muster vorgegeben werden.

Die Verzögerungszeiten, die ein Paket bei der Bearbeitung durch Protokollmechanismen wie die Segmentierung bzw. das Zusammenfügen von Paketen oder durch das Auswerten von Adressen in höheren Schichten erfährt, können durch ein M/G/1-System approximiert werden. Ähnliche Modelle werden z. B. auch in [9] für die Analyse von Netzkoppeleinheiten verwendet. Die mittlere Protokollbearbeitungszeit t_p ergibt sich nach [100]

$$t_p = \frac{\rho(1 + c_p^2)}{2(1 - \rho)} h_p + h_p, \quad (4.13)$$

mit der mittleren Bedienzeit h_p , dem Variationskoeffizienten der Bedienzeit c_p und der Systemlast ρ . Typischerweise wird die Verzögerungszeit über eine Router-Router-Kopplung (oberhalb der MAC-Teilschicht) mit einem M/G/1-System modelliert.

Schließlich werden Laufzeiten, die sich aus der endlichen Signalausbreitungsgeschwindigkeit auf den Übertragungsabschnitten ergeben, durch Verzögerungsknoten (*Infinite Server*) mit einer konstanten Transferzeit t_l modelliert.

Mit Hilfe des Warteschlangenmodells kann dann die mittlere Transferzeit für ein Paket, das von einer Verkehrsquelle m zur Verkehrssenke n gesendet wird, berechnet werden:

$$t_{mn}^R = \sum_{i \in W_b} t_{b,i} \tilde{r}_{mn}^i + \sum_{j \in W_p} t_{p,j} \tilde{r}_{mn}^j + \sum_{k \in W_l} t_{l,k} \tilde{r}_{mn}^k. \quad (4.14)$$

mit

$$\tilde{r}_{mn}^i = \begin{cases} 1, & \text{falls Knoten } i \text{ auf dem Pfad von } m \text{ nach } n \text{ liegt.} \\ 0, & \text{sonst} \end{cases} \quad (4.15)$$

Die Verzögerungszeit t_{mn}^R wird je nach Komponente aus den Gleichungen (4.8), (4.13) oder als konstante Verzögerungszeit t_l berechnet. Dazu wird aus einer gegebenen Topologie die Menge der Knoten, die die Buszugriffszeit modellieren (W_b), die Menge der Knoten zur Modellierung der Protokollbearbeitung (W_p) und die Menge der Verzögerungszeitknoten (W_l) sowie ihre

Anordnung in einem Warteschlangennetz ermittelt. Die Paketankunftsrate, die an jedem Einzelsystem durch den Eingangsverkehr auftreten, erhält man mit Hilfe der Flußanalyse nach den Gleichungen (4.2 - 4.7).

Die beschriebenen Modellkomponenten können durch genauere Verkehrsmodelle oder durch die Berücksichtigung von mehrfachen Segmentankünften verbessert werden, was aber dann mit erhöhter Rechenzeit während der Optimierung erkauft wird.

4.3 Formulierung des Problems

Ausgehend von den bereits behandelten Modellierungsaspekten läßt sich ein formales Modell aufstellen. Das Planungsziel ist es, die Netzstruktur eines MAN-Vermittlungssystems möglichst wirtschaftlich zu entwerfen unter der Bedingung, daß die Verkehrsanforderungen erfüllt sind, und die Dienstgüteparameter eingehalten werden. Vorgaben für Dienstgüteparameter bei MAN-Vermittlungssystemen, die zur Realisierung von SMDS geeignet sind, finden sich in [12]. Das Optimierungsproblem sollte so formuliert werden, daß Anforderungen, die vorrangig zu erfüllen sind, als Nebenbedingungen auftauchen. Als Zielfunktion werden daher die Gesamtkosten (Leitungskosten), die mit einer bestimmten Netzstruktur verbunden sind, gewählt.

Unter Einbeziehung der in den vorangegangenen Abschnitten beschriebenen Modellierungsaspekte läßt sich das MSS-Entwurfsproblem wie folgt formulieren: Gegeben sei die Zahl der Stationen N sowie die Verkehrsmatrix (γ_{ij}) , $i, j = 1 \dots N$, $\gamma_{ii} = 0$ und die Distanzmatrix (d_{ij}) , $i, j = 1 \dots N$, $d_{ii} = 0$. Falls eine Trasse nicht benutzt werden soll oder zwischen zwei Standorten keine Trasse existiert, wird die Entfernung zwischen diesen Standorten auf einen sehr großen Wert gesetzt, um die Verwendung dieser Trasse in einer optimierten Topologie zu vermeiden. Die Stationen sollen M Teilnetzen S_k , $k = 1 \dots M$ zugeordnet werden, die wiederum über Router geeignet zu verbinden sind. Ein beliebiges Teilnetz S_k enthält sowohl Stationen als auch Router, wobei die Gesamtzahl der Stationen und Router mit n_k bezeichnet wird.

Ein Beispiel für einen Topologiegraphen G_T mit zwei Teilnetzen S_m und S_n , die über die Router k und l gekoppelt sind, zeigt Bild 4.5. Die Doppelbusstruktur wird durch ungerichtete Kanten repräsentiert; die Lage der Kopfstationen ist im Bild nicht eingezeichnet. Ein Router, der zwei Teilnetze koppelt, wird allgemein durch einen Knoten markiert, der drei, mit diesem Knoten inzidente Kanten besitzt. Die Entscheidungsvariablen, die den Topologiegraphen bestimmen, lauten (vgl. Bild 4.5):

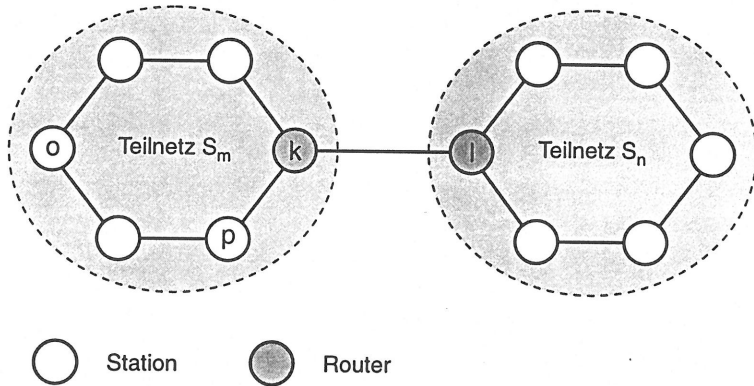


Bild 4.5: Topologiegraph des MSS

1. Variablen, die eine Verbindung zwischen Teilnetzen kennzeichnen:

$$x_{kl}^{mn} = \begin{cases} 1, & \text{falls ein Übertragungsabschnitt zwischen Router } k \in S_n \\ & \text{und Router } l \in S_m, S_n \neq S_m, k < l, \text{ existiert (aus Sym-} \\ & \text{metriegründen sind die Knoten } k > l \text{ damit automatisch} \\ & \text{berücksichtigt).} \\ 0, & \text{sonst.} \end{cases} \quad (4.16)$$

2. Variablen, die eine Verbindung innerhalb von Teilnetzen kennzeichnen:

$$y_{op}^m = \begin{cases} 1, & \text{falls ein Übertragungsabschnitt zwischen Station bzw. Router} \\ & o \in S_m \text{ und Station bzw. Router } p \in S_m, o < p, \text{ existiert (aus} \\ & \text{Symmetriegründen sind die Knoten } o > p \text{ damit automatisch} \\ & \text{berücksichtigt).} \\ 0, & \text{sonst.} \end{cases} \quad (4.17)$$

Damit lautet die Zielfunktion:

$$\min K = \sum_{m=1}^M \sum_{\forall i \in S_m} \sum_{\forall j \in S_m} (K_0 + \kappa \cdot \sqrt{d_{ij}}) y_{ij}^m + \sum_{m=1}^M \sum_{\forall i \in S_m} \sum_{n=1}^M \sum_{\forall j \in S_n} (K_0 + \kappa \cdot \sqrt{d_{ij}}) x_{ij}^{mn}. \quad (4.18)$$

Die Gesamtkosten K berücksichtigen die Einzelkosten einer Leitung nach Gleichung 4.1. Der Leitungstyp wird vorgegeben, so daß der zugehörige Index entfällt.

Die Minimierung der Zielfunktion soll unter den folgenden Nebenbedingungen erreicht werden:

1. Beschränkung der Topologie:

$$\sum_{\forall k \in S_m} y_{kl}^m + \sum_{\forall k \in S_m} y_{lk}^m = 2, \forall l \in S_m, \quad (4.19)$$

$$\sum_{\forall k \in S'_m} \sum_{\forall l \in S'_m} y_{kl}^m \leq |S'_m| - 1, \forall S'_m \subset \{i_2 \dots i_{n_m}\}, 3 \leq |S'_m| < n_m - 1, \quad (4.20)$$

wobei $i_2 \dots i_{n_m} \in \{1 \dots \sum_{m=1}^M n_m\}$ gilt. Diese beiden Nebenbedingungen stellen sicher, daß die Stationen und Router in einem Teilnetz einen Ring bilden (*Looped-Bus-Topologie*). Dabei ist n_m die Anzahl der Netzelemente in Teilnetz S_m . Die Gleichungen (4.19) und (4.20) gelten analog für allgemeine Ringtopologien [135]. Gleichung (4.20) legt fest, daß nur ein einziger Ring pro Teilnetz zulässig ist. Mit i_1 ist die Kopfstation von Bus A bestimmt. Weiterhin gilt:

$$\sum_{\forall k \in S_m} x_{kl}^{mn} \leq 1, \forall l \in S_n, \forall S_n, \quad (4.21)$$

$$\sum_{m=1}^M \sum_{\forall k \in S_m} \sum_{\forall l \in S_m} y_{kl}^m - 2 \cdot \sum_{m=1}^M \sum_{\forall k \in S_m} \sum_{n=1}^M \sum_{\forall l \in S_n} x_{kl}^{mn} = N. \quad (4.22)$$

Die Nebenbedingung nach Gleichung (4.21) führt dazu, daß zwischen zwei beliebigen Teilnetzen höchstens eine einzige *Router-Router-Kopplung* erlaubt ist. Außerdem darf jede Station nur genau einem Teilnetz zugeordnet sein (s. Gleichung (4.22)).

2. Flußerhaltung: Für den Verkehrsfluß f_{ijmn}^{qp} von Station m zu Station n über den Übertragungsabschnitt zwischen Router i , $i \in S_q$, und Router j , $j \in S_p$, gilt:

$$\sum_{\forall j \in S_p} f_{jimm}^{pq} - \sum_{\forall j \in S_p} f_{ijmn}^{qp} = \begin{cases} \gamma_{mn}, & \text{falls } n \in S_q, \\ -\gamma_{mn}, & \text{falls } m \in S_q, \\ 0, & \text{sonst.} \end{cases} \quad (4.23)$$

Diese Nebenbedingung besagt, daß der Verkehr, der in ein Teilnetz hinein fließt, wieder aus dem Teilnetz heraus fließen muß, falls sich nicht die Verkehrsenke in diesem Teilnetz befindet. Um diese Nebenbedingung zu erfüllen, ist es erforderlich, daß das Netz mindestens einen Pfad zwischen jedem Quelle-Ziel-Paar aufweist.

3. Beschränkung der Kapazität:

$$f_{i,q}^A \leq \rho_S C_S, \forall i \in S_q, \forall S_q, \quad (4.24)$$

$$f_{i,q}^B \leq \rho_S C_S, \forall i \in S_q, \forall S_q, \quad (4.25)$$

$$f_{ij}^{mn} \leq \rho_R C_R, \forall i \in S_m, \forall S_m, \forall j \in S_n, \forall S_n. \quad (4.26)$$

Diese Nebenbedingungen garantieren, daß der Verkehrsfluß weder in Teilnetzen (auf Bus A und B) noch über einen Router die zulässige Kapazität $\rho_S C_S$ bzw. $\rho_R C_R$ überschreitet; ρ_S und ρ_R sind die vom Netzplaner vorgegebenen Lastgrenzen. Die Kapazität eines Teilnetzes ist bei MANs auf wenige Werte beschränkt und wird ebenfalls vorgegeben.

4. Beschränkung der mittleren Pfadtransferzeit:

$$t_{mn}^R < t_{max}^R, \forall m, n \in \{1 \dots N\} \quad (4.27)$$

mit t_{mn}^R nach Gleichung (4.14). Mit dieser Nebenbedingung wird zusätzlich zur Beschränkung der Last gefordert, daß für jeden Pfad von Station m bis Station n ein vorgegebener Maximalwert t_{max}^R der mittleren Transferzeit einzuhalten ist.

Für das beschriebene Modell sind verschiedene Erweiterungen möglich. So gibt es bei realen Planungsaufgaben für die Netzelemente meist Varianten bezüglich der Kosten oder der Kapazität. Diese Varianten können berücksichtigt werden, wenn beispielsweise für eine *Router-Router*-Kopplung eine Reihe von *Router*-Typen eingesetzt werden kann. Diese Typen werden im Modell durch eine abzählbare Menge repräsentiert. Folglich müßten auch die Variablen x_{ki}^{mn} angepaßt werden. Darüber hinaus können Anforderungen an die Verfügbarkeit, wie etwa die minimale Zahl knotendisjunkter Pfade zwischen zwei Stationen, in das Modell miteinbezogen werden. Da dies die Komplexität des Modells erhöht, werden solche Anforderungen zunächst vernachlässigt.

Eine duale Formulierung des Modells ergibt sich, wenn die Zielfunktion nach Gleichung (4.18) in eine Nebenbedingung umgewandelt wird, die die Kosten auf einen Maximalwert begrenzt, und als neue Zielfunktion die Minimierung der mittleren Transferzeit des gesamten MSS angestrebt wird. Sinnvollerweise berücksichtigt diese Zielfunktion die Transferzeit aller Pfade, etwa durch Summation, um der Nebenbedingung nach Gleichung (4.27) möglichst nahe zu kommen. Dabei wird vorausgesetzt, daß ein bestimmtes Budget zur Realisierung des Netzes gegeben ist. Da diese Formulierung mit einer zusätzlichen Nebenbedingung verknüpft ist, wird auf die nähere Untersuchung verzichtet.

4.4 Komplexität

Das oben beschriebene Modell stellt ein nichtlineares, kombinatorisches Optimierungsproblem dar. Nichtlinear ist zum einen die Zielfunktion in Abhängigkeit von den Variablen x_{ij}^{mn} und y_{ij}^m und zum anderen die Nebenbedingung, die den Maximalwert der mittleren Transferzeit begrenzt (Gleichung 4.27). Im zweiten Fall ist die mittlere Transferzeit t_{mn}^R ebenfalls von der Netzstruktur, und damit von x_{ij}^{mn} und y_{ij}^m abhängig. Da bereits Probleme der ganzzahligen, linearen Optimierung meist nur approximativ zu lösen sind, enthält das MSS-Entwurfsproblem durch die nichtlinearen Terme einen zusätzlichen Schwierigkeitsgrad. Zu seiner Lösung sind keine Standardverfahren bekannt.

Um zu verdeutlichen, daß das MSS-Entwurfsproblem effizient nur mit heuristischen Verfahren gelöst werden kann, kann einerseits ein reduziertes Modell betrachtet werden und andererseits die Größe des Zustandsraums untersucht werden.

Ein reduziertes Modell erhält man, wenn ein Teil der ursprünglichen Randbedingungen vernachlässigt wird. Zwei Spezialfälle sind offensichtlich. Im ersten Fall sei angenommen, daß alle Stationen in einem Teilnetz plaziert werden können, d. h. die Kapazität der Teilnetze ist ausreichend, um alle Verkehrsanforderungen zu erfüllen, und wird nicht überschritten. Darüber hinaus sei die Beschränkung der maximal zulässigen Transferzeit zwischen je zwei Stationen aufgehoben. Dann ist im wesentlichen eine kostenoptimale Anordnung der Stationen in der *Looped-Bus*-Konfiguration zu finden. Dieses Problem entspricht dem Rundreiseproblem (*Traveling Salesman Problem*) und ist NP-vollständig [158]. Soll darüber hinaus auch eine optimale Anordnung bezüglich der Transferzeiten erreicht werden, ist zusätzlich die Last des Bussystems zu berücksichtigen. Die Last kann aber – im Gegensatz zu den Kosten, die aus der Entfernungsmatrix zwischen jeder Station berechnet werden können – erst nach der Festlegung der Topologie berechnet werden, was eine weitere Erhöhung der Komplexität darstellt.

Im zweiten Fall sei es möglich, jede Station in einem separaten Teilnetz zu plazieren. Dann ergibt sich das Problem des topologischen Entwurfs für ein Paketvermittlungsnetz, unter der Annahme, daß die Stationen auch die Verkehrslenkungsfunktionen übernehmen können. Dieses Problem ist ebenfalls nicht exakt zu lösen [92]. Generell sind jedoch drei Probleme überlagert, die nicht unabhängig voneinander gelöst werden können:

- Ermitteln der Partitionen, die jeweils alle Stationen eines Teilnetzes umfassen.
- Verbindung der Teilnetze.
- Bestimmung der Stationsreihenfolge in den Teilnetzen.

Mit der Festlegung der topologischen Struktur ist auf Grund des statischen Verkehrslenkungsverfahrens auch gleichzeitig das Verkehrslenkungsschema bestimmt.

Die Größe des Zustandsraums kann folgendermaßen abgeschätzt werden. Vereinfachend wird angenommen, daß die N Stationen in M Teilnetze mit jeweils k Stationen aufzuteilen sind, wobei $k \cdot M = N$ gilt. Die Anzahl der möglichen Aufteilungen (M, k) beträgt:

$$Z_P(M, k) = \frac{1}{M!} \prod_{i=0}^{M-1} \frac{(N - ik)!}{k!(N - ik - k)!}. \quad (4.28)$$

Weiterhin sollen zur Verbindung der Teilnetze nur Baumstrukturen betrachtet werden, auch wenn im allgemeinen beliebige Vermaschungsgrade zulässig sind. Die Zahl der Bäume, die M Teilnetze verbinden, erhält man mit der Formel von Cayley [20]:

$$Z_V(M) = M^{M-2}. \quad (4.29)$$

Schließlich kann die Zahl der möglichen Anordnungen in den Teilnetzen approximiert werden, wenn angenommen wird, daß mit der baumförmigen Verknüpfung der Teilnetze mindestens

(M, k)	$Z(M, k)$
(1, 30)	$4.11 \cdot 10^{33}$
(2, 15)	$1.62 \cdot 10^{21}$
(3, 10)	$1.66 \cdot 10^{20}$
(5, 6)	$1.80 \cdot 10^{22}$
(6, 5)	$3.45 \cdot 10^{23}$
(10, 3)	$1.45 \cdot 10^{28}$
(15, 2)	$5.42 \cdot 10^{32}$

Tabelle 4.1: Größe des Zustandsraums für $N = 30$

ein *Router* in jedes Teilnetz zu den bereits vorhandenen k Stationen eingefügt werden muß:

$$Z_B(M, k) = \frac{1}{2} M(k+1)!. \quad (4.30)$$

Daher ergibt sich für eine Aufteilung (M, k) die gesamte Zahl der Möglichkeiten $Z(M, k)$ mit

$$Z(M, k) = Z_P(M, k) \cdot Z_V(M) \cdot Z_B(M, k). \quad (4.31)$$

Selbst für diese vereinfachte Betrachtung kann der Zustandsraum auch bei Stationszahlen mittleren Umfangs eine Größe annehmen, die das vollständige Absuchen des Zustandsraums gänzlich ausschließt. So erhält man beispielsweise für $N = 30$ unter den Voraussetzungen, daß mindestens zwei Stationen in einem Teilnetz sein müssen, und alle Aufteilungen (M, k) betrachtet werden, für die Größe des Zustandsraums unter Auswertung von Gleichung 4.31 die Werte nach Tabelle 4.1. Daher müssen vorrangig Verfahren entwickelt werden, die eine möglichst gute Lösung mit vertretbarem Rechenaufwand finden. Diese Aufgabe ist Gegenstand des folgenden Kapitels.

Kapitel 5

Verfahren für den topologischen Entwurf von MAN-Vermittlungssystemen

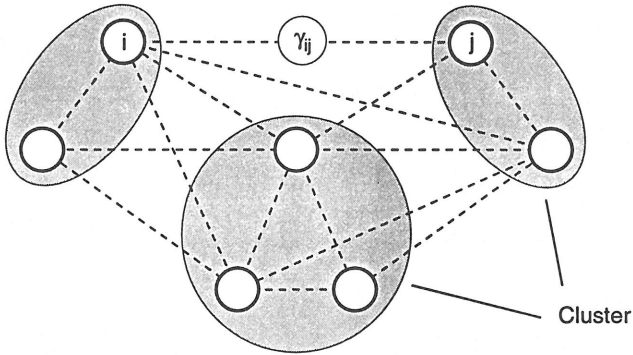
Wie im vorhergehenden Kapitel ausgeführt wurde, ist das MSS-Entwurfsproblem ein komplexes kombinatorisches Optimierungsproblem, das aus der Überlagerung dreier Teilprobleme besteht. Da ein exakter und gleichzeitig effizienter Lösungsalgorithmus nicht angegeben werden kann, stützen sich die MSS-Entwurfsalgorithmen auf heuristische Verfahren. Um die Qualität der Lösungen besser beurteilen zu können, werden dazu drei verschiedene Ansätze gewählt, die in diesem Kapitel beschrieben und einem Vergleich hinsichtlich ihrer Leistungsfähigkeit unterzogen werden. Ein Abschnitt, der die Erweiterungen der Planungsverfahren diskutiert, beschließt das Kapitel.

5.1 Dekompositionsverfahren

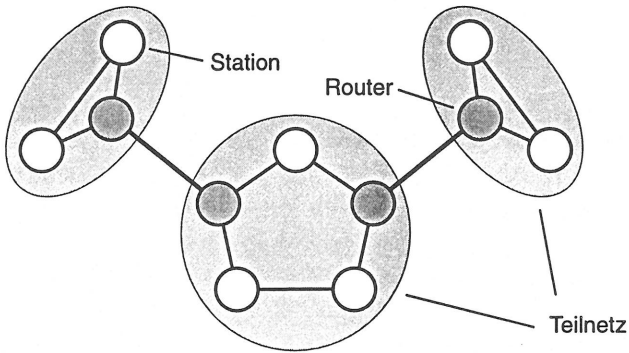
Als erster Ansatz bietet sich eine Aufteilung des Entwurfsproblems an (Dekompositionsverfahren). Das Dekompositionsverfahren verwendet die Zielfunktion nach Gleichung 4.18 nicht direkt, sondern versucht in mehreren Schritten eine gute Lösung zu erzielen. Das wird durch die Anwendung von Entwurfsprinzipien erreicht, die auf eine Minimierung der Kosten abzielen, mit denen zugleich aber auch Netzstrukturen mit einer ausreichenden Netzgüte erzeugt werden können.

Bild 5.1 gibt einen schematischen Überblick über das Verfahren. Im ersten Schritt wird versucht, unter Ausnutzung der geographischen Verteilung der Stationen und der Verkehrsflüsse γ_{ij} sog. *Cluster* zu bilden, die bereits alle Stationen eines Teilnetzes enthalten. Dabei ist ein Kompromiß zwischen der wirtschaftlichen Auslastung und der geographischen Ausdehnung

der *Cluster* zu finden. Der zweite Schritt umfaßt dann die Verknüpfung der Teilnetze durch *Router* und die Festlegung der *Looped-Bus*-Struktur in den Teilnetzen. Damit wird auch gleichzeitig die Verkehrslenkung festgelegt. Die Netztopologie kann abschließend analysiert und bewertet werden. In den folgenden Abschnitten werden die einzelnen Schritte und dann der vollständige Algorithmus ausführlich beschrieben.



Schritt 1



Schritt 2

Bild 5.1: Entwurfsschritte des Dekompositionsverfahrens

5.1.1 Entwurf der Teilnetze

Teilnetze können prinzipiell durch ein *Partitioning*- oder ein *Clustering*-Verfahren gefunden werden. In beiden Fällen ist eine Zuordnung der Stationen zu separaten Teilnetzen zu finden, die eine gegebene Gütefunktion maximiert oder minimiert. Die Gütefunktion hängt typischerweise vom Verkehrsfluß und der Entfernung zwischen den Stationen ab.

Ein *Partitioning*-Verfahren versucht, eine optimale Aufteilung durch aufeinanderfolgende Teilungen einer Anfangsmenge, die alle Elemente enthält, zu finden [90]. Dies gestaltet sich aufwendig, wenn die einzelnen Partitionen nicht nur hinsichtlich der Zahl der Elemente, sondern auch hinsichtlich weiterer Kriterien, wie im vorliegenden Fall durch den maximalen Verkehrsfluß in einem Teilnetz, beschränkt sind. Insbesondere ist es schwierig, Teilungsstrategien zu entwickeln.

Für das vorliegende Planungsproblem ist ein sog. *Clustering*-Verfahren besser geeignet. Beim *Clustering* wird versucht, einen gegebenen Graphen durch schrittweises Zusammenfassen der Knoten zu Teilmengen aufzuteilen. Dabei wird angenommen, daß ein „natürlicher“ Zusammenhang existiert, der durch gewichtete Kanten des Graphen repräsentiert wird. Bezogen auf das MSS-Planungsproblem entsprechen die Knoten den Stationen und die endgültigen *Cluster* den Teilnetzen. Ein Zusammenhang ergibt sich unter der Annahme, daß eine günstige Netzstruktur dann entsteht, wenn sowohl die Lokalität des Verkehrs als auch die geographische Verteilung der Standorte ausgenutzt wird. Das bedeutet, daß vorzugsweise Stationen mit hohem wechselseitigem Verkehrsaufkommen bzw. Stationen in unmittelbarer Nähe zueinander zu einem *Cluster* zusammengefaßt werden sollten. Die Kantengewichte ergeben sich aus der Größe des Verkehrsflusses bzw. der Entfernung zwischen den Stationen.

Clustering-Verfahren zum Entwurf von speziellen zentralen und dezentralen Netzstrukturen wurden bereits in der Literatur untersucht [37, 55, 113, 121, 144]. Sie vernachlässigen jedoch Parameter, die für den Entwurf eines MSS spezifisch sind:

- Da erst im nächsten Schritt der Transitverkehr durch die Teilnetze festgelegt wird, sollte die Last in einem *Cluster* begrenzt werden und deutlich unterhalb der Teilnetzkapazität liegen. Aus Kostengründen sollten die Teilnetze aber auch wirtschaftlich ausgelastet werden und daher eine gewisse Mindestauslastung aufweisen.
- Die Last sollte unter den Teilnetzen möglichst gleichmäßig verteilt werden.
- Um die Kosten, aber auch die Unfairness beim Buszugriff, die sich bei hoher Last und großer Ausdehnung in stark unterschiedlichen Buszugriffszeiten niederschlagen kann, gering zu halten, sollte die *Cluster*-Ausdehnung begrenzt werden.

Das Aufteilungsproblem kann nun wie folgt formuliert werden: Gegeben seien N Stationen, die als Knoten eines Graphen aufgefaßt werden können. Die Kanten des Graphen repräsentieren die Verkehrsmatrix (γ_{ij}) und die Distanzmatrix (d_{ij}) (vgl. Abschnitt 4.3). Gesucht ist

eine Aufteilung in M Cluster, bei der der externe Verkehr möglichst gering ist, die Lastdifferenzen zwischen Teilnetzen nicht zu groß sind und ein bestimmter Cluster-Durchmesser nicht überschritten wird. Diese Ziele werden durch einen Clustering- und Lastteilungs-Algorithmus (CLA) erreicht, der in zwei Phasen vorgeht. Die erste Phase umfaßt das Clustering-Verfahren, die zweite Phase den Lastausgleich zwischen den Teilnetzen.

Für den CLA werden folgende Festlegungen benötigt: Ein Cluster \tilde{S}_i kann durch

- eine Menge von Stationen,
- einen Mittelpunkt (geographischer Schwerpunkt) M_i ,
- den Durchmesser D_i ,
- den internen Verkehrsfluß λ_{int}^i , der innerhalb eines Cluster und damit auch auf dem Dualbussystem auftritt, und
- den externen Verkehrsfluß λ_{ext}^i , der den Verkehrsfluß aus einem bzw. in einen Cluster umfaßt,

definiert werden. Zusätzlich gibt es Nebenbedingungen, die für alle Cluster gelten. Die Kapazität C_S , die aus der Kapazität des Dualbussystems resultiert, und die Auslastungsgrenze ρ_{Cl} begrenzen den Verkehrsfluß in einem Cluster. Der maximal zulässige Durchmesser D_{max} bestimmt die maximale Ausdehnung der Cluster.

Während der Clustering-Phase wird in mehreren Iterationen versucht, jeweils zwei Cluster zusammenzufassen. Ist dies unter Beachtung der Nebenbedingungen möglich, wird ein neuer Cluster gebildet, dessen Elemente sich aus der Vereinigungsmenge der beiden vorhergehenden Cluster zusammensetzen. Die Cluster-Parameter werden dabei neu berechnet. Um zu prüfen, ob der dadurch entstehende Durchmesser D größer ist als D_{max} , wird D wie folgt geschätzt:

$$D = \begin{cases} D_1, & \text{falls } D_1/2 \geq a(M_1, M_2) + D_2/2, \\ a(M_1, M_2) + (D_1 + D_2)/2, & \text{falls } D_1/2 < a(M_1, M_2) + D_2/2. \end{cases} \quad (5.1)$$

wobei $a(M_1, M_2)$ der Abstand zwischen den beiden Mittelpunkten ist, und D_1 bzw. D_2 die jeweiligen Durchmesser der Cluster sind. Gleichung 5.1 berücksichtigt mit der ersten Bedingung, daß Cluster 2 in seiner geographischen Ausdehnung vollständig von Cluster 1 überdeckt werden kann. Als Cluster-Mittelpunkt wird nach dem Zusammenfügen der geographische Mittelpunkt aus allen Stationskoordinaten berechnet. Der größte Abstand einer Station vom Cluster-Mittelpunkt definiert den neuen Cluster-Radius.

Bild 5.2 zeigt den Ablauf des CLA. Die Funktion F_k , die den entstehenden Cluster \tilde{S}_k bewertet, dient der Steuerung des Clustering in der ersten Phase des CLA, um einerseits das Verhältnis zwischen dem internen Verkehrsfluß λ_{int}^k und der Teilnetzkapazität C_S zu maximieren und andererseits das Verhältnis zwischen externem und internem Verkehrsfluß $\lambda_{ext}^k/\lambda_{int}^k$

{ 1. Phase CLA }

- Weise jedem Cluster \tilde{S}_i eine Station zu.
- Bilde Cluster-Liste \tilde{S} mit $\bigcup_{i=1}^M \tilde{S}_i = \tilde{S}$.
- Initialisiere Teilnetzliste S : $S = \emptyset$.
- Wiederhole, solange ($\tilde{S} \neq \emptyset$):

1. Wähle einen Cluster \tilde{S}_i aus und suche einen Partner \tilde{S}_j , der für $\tilde{S}_k = \tilde{S}_i \cup \tilde{S}_j$ die Funktion

$$F_k = \frac{\lambda_{int}^k}{C_S} - \frac{\lambda_{ext}^k}{\lambda_{int}^k},$$

unter den folgenden Bedingungen maximiert: $\lambda_{int}^k < \rho_{Cl} \cdot C_S$ und $D_k < D_{max}$.

2. Falls für \tilde{S}_i kein Partner gefunden wird, entferne \tilde{S}_i aus \tilde{S} und füge $S_i = \tilde{S}_i$ in S ein.
3. Andernfalls, bilde $\tilde{S}_k = \tilde{S}_i \cup \tilde{S}_j$ und aktualisiere die Parameter M_k und D_k sowie die Mengen \tilde{S}_k und \tilde{S} .

{ 2. Phase CLA }

- Ordne Teilnetzliste S nach der Größe von $\lambda_{int}^i, \forall S_i \in S$.
 - Wähle Teilnetze S_j, S_k mit $\lambda_{int}^j = \max(\lambda_{int}^i), \forall S_i \in S$, und $\lambda_{int}^k = \min(\lambda_{int}^i), \forall S_i \in S$.
 - Wiederhole, solange $((\lambda_{int}^j - \lambda_{int}^k)/C_S > \Delta\rho \wedge$ eine Anordnung in S nicht wiederholt wird):
 1. Entferne das Element aus S_j , das den geringsten Abstand zu S_k aufweist, und füge es in S_k ein.
 2. Aktualisiere S_j und S_k sowie λ_{int}^j und λ_{int}^k .
 3. Ordne Teilnetzliste S nach der Größe von $\lambda_{int}^i, \forall S_i \in S$ und wähle Teilnetze S_j und S_k wie oben.
-

Bild 5.2: Clustering- und Lastteilungs-Algorithmus

zu minimieren. Durch die Differenzbildung wird ein annähernd gleichmäßiges Wachstum der Cluster erzielt. Zu Beginn ist zwar die externe Last sehr hoch; diese nimmt jedoch mit der Vergrößerung der Cluster mehr und mehr ab, so daß F_k bei gleichzeitiger Steigerung der internen Last zunimmt. Der interne bzw. externe Verkehrsfluß von \tilde{S}_k ergibt sich mit

$$\lambda_{int}^k = \sum_{i \in \tilde{S}_k} \sum_{j \in \tilde{S}_k} \max(\gamma_{ij}, \gamma_{ji}) + \sum_{i \in \tilde{S}_k} \sum_{j \in S \setminus \tilde{S}_k} \max(\gamma_{ij}, \gamma_{ji}) \quad (5.2)$$

und

$$\lambda_{ext}^k = \sum_{i \in \bar{S}_k} \sum_{j \in \bar{S} \setminus \bar{S}_k} \max(\gamma_{ij}, \gamma_{ji}). \quad (5.3)$$

Für diese Berechnung wird jeweils das Maximum des Verkehrs in beiden Richtungen herangezogen, da zum Zeitpunkt der *Cluster*-Bildung die genaue Struktur des Dualbussystems eines Teilnetzes noch nicht bekannt sein kann. Diese wird aber benötigt, um den exakten Wert zu erhalten. Die Gleichungen (5.2) und (5.3) bilden daher eine obere Schranke für die Last auf einem Bus.

Ziel ist es, die Last bereits während des *Clustering*-Vorgangs gleichmäßig zu verteilen. Dies gelingt jedoch durch die Vorgabe eines maximal zulässigen Durchmessers nicht immer. Daher wird in der zweiten Phase des CLA versucht, die Lastdifferenzen zwischen den Teilnetzen im Rahmen eines erwünschten Bereichs $\Delta\rho$ zu vermindern. Dazu werden die Teilnetze S_i zunächst nach der Größe ihres Verkehrsflusses λ_{int}^i angeordnet. Dann wird eine Station aus dem höher ausgelasteten Teilnetz in das geringer ausgelastete Teilnetz eingefügt, wobei eine geographisch günstig gelegene Station des höher ausgelasteten Teilnetzes ausgewählt wird. Dies geschieht so lange, bis die gewünschte Differenz $\Delta\rho$ erreicht ist oder sich eine Anordnung in der Teilnetzliste S wiederholt. Bei der Wahl von $\Delta\rho$ ist zu beachten, daß sich eine sehr kleine Differenz nachteilig auf die mittels *Clustering* gefundenen geographischen Gruppierungen auswirken kann.

Ist der CLA beendet, werden die Stationen in einem Teilnetz mit Hilfe einer einfachen Heuristik angeordnet, die für Reihenfolgeprobleme relativ gute Ergebnisse erbringt und wenig Rechenzeit in Anspruch nimmt [34]. Es ist zu beachten, daß die so ermittelte Anordnung noch nicht endgültig ist, da zur Verbindung der Teilnetze *Router* in die Teilnetze eingefügt werden müssen. Ein genauerer Algorithmus ist an dieser Stelle nicht notwendig.

Der Algorithmus zur Anordnung der Stationen in einem Teilnetz kann folgendermaßen zusammengefaßt werden: Zunächst wird eine beliebige Station ausgewählt und als erstes Element in die Stationsliste eingefügt. Danach wird sukzessive aus den freien Stationen die Station ausgewählt, die die geringste Entfernung zur letzten Station der Stationsliste besitzt, und an diese angefügt. Das erfolgt so lange, bis alle Stationen in der Stationsliste des Teilnetzes angeordnet sind. Diese Anordnung bestimmt die vorläufige Struktur des Dualbussystems. Das Vorgehen beim Einfügen der *Router* in die Teilnetze wird im nächsten Abschnitt behandelt.

5.1.2 Verbindung der Teilnetze

Um die Netzstruktur zu vervollständigen, müssen die Teilnetze verbunden werden. Die Teilnetze sollten so verbunden werden, daß einerseits nur die zur Führung der Verkehrsflüsse

mindestens notwendigen *Router-Router-Kopplungen* eingerichtet werden, andererseits aber die Anforderungen hinsichtlich der Transferzeit eingehalten werden. Dazu wurde ein Verfahren, das für den topologischen Entwurf beliebig vermaschter Strukturen (typischerweise Paketvermittlungsnetze) geeignet ist (vgl. dazu [91]), so erweitert, daß es die Verbindung von MAN-Teilnetzen ermöglicht. Mit diesem Verfahren wird implizit auch das Verkehrslenkungsschema festgelegt.

Das Verfahren umfaßt zwei Schritte: Der erste Schritt sieht die Berechnung eines erzeugenden Baums (*Spanning Tree, ST*) vor, der alle Teilnetze enthält, d. h. nun werden die Teilnetze als Knoten und die Übertragungsabschnitte zwischen ihnen als Kanten eines Graphen aufgefaßt. Folglich muß aus der Verkehrsmatrix für die Stationen eine Verkehrsmatrix berechnet werden, die den Verkehrsfluß zwischen Teilnetzen repräsentiert. Zu Beginn der ST-Berechnung wird ein Knoten als Zentrum aus der Gewichtung des Verkehrs, der von und zu diesem Knoten fließt, und der Entfernung zu anderen Teilnetzen berechnet. Der Knoten mit dem geringsten Gewicht wird als Zentrum ausgewählt. Dann werden die Teilnetze, in Analogie zum *Minimum-Spanning-Tree-Verfahren*, schrittweise an den Baum angefügt. Allerdings wird als Kantengewicht für einen anzufügenden Knoten i jeweils die modifizierte Länge

$$d'_{ij} = d_{ij} + g_d \cdot d_{jM} \tag{5.4}$$

verwendet. Dabei ist d_{ij} die Entfernung zwischen Teilnetz S_i und Teilnetz S_j , das als möglicher Anschlußort in Frage kommt, und d_{jM} die Entfernung von Teilnetz S_j zu Teilnetz S_M , das den gewählten Mittelpunkt darstellt. Der Gewichtungsfaktor g_d kann Werte zwischen 0 und 1 annehmen.

Die Wirkungsweise verdeutlicht Bild 5.3. Knoten S_1 wird nur dann an Knoten S_2 angeschlossen, wenn der Gewichtungsfaktor g_d den Anteil von d_{2M} so weit verringert, daß die Summe aus d_{12} und d_{2M} kleiner wird als d_{1M} .

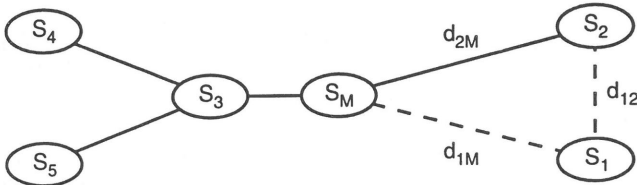


Bild 5.3: Generierung des gewichteten *Spanning Tree*

Diese Berechnungsvorschrift hat zum Ziel, ein Grundgerüst für die Verbindung der Teilnetze zu bilden, das sowohl günstige Eigenschaften hinsichtlich der Netzkosten besitzt als auch geringe Pfadlängen zwischen den Netzknoten aufweist. Dies wird deutlich, wenn man die Extremwerte betrachtet: Falls g_d den Wert 0 annimmt, entsteht ein *Minimum Spanning*

Tree; falls g_d den Wert 1 annimmt, ein sternförmiges Netz, bei dem alle Knoten über ein Zentrum miteinander verbunden sind. Im ersten Fall entsteht ein kostenoptimales Gerüst, im zweiten Fall ein Gerüst, das hinsichtlich der Ende-zu-Ende-Verzögerungszeiten optimal ist. Für die Zwischenwerte ergibt sich eine Mischform.

Der zweite Schritt berücksichtigt speziell die Eigenschaften einer MSS-Topologie. Die Ausgangsbasis bildet die im vorhergehenden Schritt erstellte Baum-Struktur. Beginnend mit den äußeren Knoten – in Bild 5.3 z. B. die Knoten S_4 und S_5 –, wird ein Abhängigkeitsgraph der Verkehrsflüsse zwischen allen Knoten aufgestellt. Eine Abhängigkeit entsteht, wenn für den Verkehrsfluß durch den Baum von Knoten S_i nach Knoten S_j ein weiterer Knoten S_k als Transitknoten genutzt werden kann, der auf dem Weg zum Baummittelpunkt liegt und über den der Verkehrsfluß zwischen S_i und S_j gelenkt werden kann. Damit sind die Kanten (S_i, S_k) und (S_k, S_j) von (S_i, S_j) abhängig. In Bild 5.3 sind beispielsweise (S_4, S_3) und (S_3, S_5) von (S_4, S_5) abhängig. Gibt es zwei Transitknoten, wird der Knoten gewählt, der den kürzeren Weg beinhaltet.

Der Abhängigkeitsgraph wird im nächsten Schritt in eine Liste von Verkehrsflüssen zwischen Knoten umgesetzt, wobei die Reihenfolge in der Liste durch die Reduktion der Abhängigkeiten bestimmt wird. Eine Kante ohne Abhängigkeiten wird sofort in die Liste eingeordnet. Dadurch reduziert sich auch die Abhängigkeit anderer Kanten, die von dieser Kante abhängig sind. Sind alle Kanten in die Liste der Verkehrsflüsse eingetragen, wird diese sequentiell abgearbeitet, um die Verkehrsanforderungen vollständig zu erfüllen.

Generell wird versucht, den Verkehr über den bereits bestehenden Baum bzw. über schon bestehende Kanten zu leiten. Überschreitet jedoch die Last eines Knotens, der als Transitknoten für den Verkehr zwischen einem Quell- und einem Zielknoten dienen soll, die zulässige Auslastungsgrenze ρ_V , wird zwischen dem Quell- und dem Zielknoten eine zusätzliche Kante eingerichtet, ohne diesen und andere potentielle Transitknoten zu belasten. Bild 5.4 zeigt ein Beispiel; die Transitknoten wären S_2 und S_3 . Die Auslastungsgrenze ρ_V , für die die Bedingung $\rho_V \geq \rho_{Cl}$ erfüllt sein muß, erlaubt eine flexible Erhöhung bzw. Verringerung der zusätzlichen Last in den Teilnetzen.

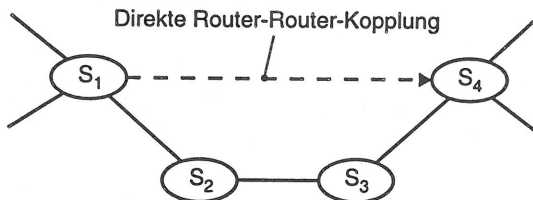


Bild 5.4: Direkte Router-Router-Kopplung

Eine Kante wird auch dann eingefügt, wenn der Schwellwert ρ_g für die Last auf einer Kante überschritten wird, und eine hinreichende Auslastung des korrespondierenden Übertragungsabschnittes gewährleistet ist. Andernfalls wird der Verkehrsfluß über das bereits bestehende Netz gelenkt und der Verkehrsfluß der entsprechenden Kante erhöht. Da der Algorithmus für ungerichtete Verkehrsflüsse konzipiert ist, wird nur das Maximum aus beiden Richtungen des Verkehrsflusses als Abschätzung für den tatsächlichen Verkehrsfluß verwendet. Der Algorithmus erstellt darüber hinaus auch ein Verkehrslenkungsschema, das aus einer Menge von *Destination Spanning Trees* besteht und für MAN-Vermittlungssysteme geeignet ist.

Abschließend werden die *Router* in den Teilnetzen an den Standorten plaziert, die zu den Teilnetzen, mit denen die *Router* jeweils verbunden sind, die geringste Entfernung aufweisen. Damit ist auch die Struktur der Dualbussysteme vollständig bestimmt. Da der Netztyp und die Kapazität des Teilnetzes durch die Stationen vorgegeben ist, muß ein *Router* gewählt werden, der die Anschlußbedingungen (gleicher Leitungstyp) erfüllt.

Das geschilderte Verfahren zur Verbindung der Teilnetze zielt nicht direkt auf eine Minimierung der Zielfunktion nach Gleichung 4.18 ab. Folglich kann durch eine lokale Optimierung der Topologie das Ergebnis weiter verbessert werden. Eine Möglichkeit ist die Nutzung freier Teilnetzkapazität. Im einfachsten Fall werden *Router-Router*-Kopplungen zwischen Teilnetzen entfernt, wenn sie nur gering ausgelastet sind. Das erfolgt solange, bis die maximal zulässige Transferzeit für ein beliebiges Quelle-Ziel-Paar überschritten wird. Meist genügt bereits die Entfernung weniger *Router-Router*-Kopplungen, um eine signifikante Kostenminderung zu erreichen.

5.1.3 Ablaufschema

Nun kann der vollständige Ablauf des Dekompositionsverfahrens erfaßt werden. Die bereits geschilderten Schritte, also der Entwurf und die Verbindung der Teilnetze, werden sequentiell ausgeführt. Damit ist die Netzstruktur und das Verkehrslenkungsschema des MAN-Vermittlungssystems, das mit der Netzstruktur implizit festgelegt wird, bekannt. Es können aber auch andere Verkehrslenkungsschemata berechnet werden. Abschließend erfolgt die Analyse und Bewertung bezüglich der Kosten, der Auslastung und der mittleren Transferzeit der einzelnen Pfade bzw. des gesamten Systems. Werden die geforderten Randbedingungen nicht eingehalten, muß das Netz mit geänderten Parametern erneut entworfen werden.

Das Verfahren besitzt eine Reihe von Parametern, deren Einstellung problemabhängig ist. Die Erfahrung hat gezeigt, daß es in gewissen Parameterbereichen robust ist und mit Grundeinstellungen gute Ergebnisse erzielt (s. Abschnitt 5.4). Darüber hinaus unterstützt das Verfahren die schnelle Erzeugung und Bewertung verschiedener MSS-Topologien und kann

daher den Ausgangspunkt für Verfahren zur genaueren, aber deutlich rechenintensiveren Planung bilden. Wie das geschieht, wird in den nächsten beiden Abschnitten gezeigt.

5.2 Verfahren basierend auf Simulated Annealing

5.2.1 Struktur der Nachbarschaft

Die Lösung eines kombinatorischen Optimierungsproblems mit Hilfe eines Optimierungsverfahrens, das auf *Simulated Annealing* basiert, erfordert den Entwurf eines lokalen Suchverfahrens. Ein solches Suchverfahren beruht, wie bereits ausgeführt, auf der Festlegung einer Nachbarschaft und geeigneter Transformationen, die der Erzeugung neuer Nachbarschaftsstrukturen dienen. Die Festlegung der Transformationen stützt sich auf zwei Kriterien:

- Ein Suchalgorithmus mit einer Nachbarschaft, die nur einen geringen Umfang hat, besitzt günstige Eigenschaften hinsichtlich der Rechenzeit.
- Durch die Transformationen sollte ein lokales Optimum und damit im günstigsten Fall auch das globale Optimum erreichbar sein.

Um einen effizienten Algorithmus zu erhalten, ist auf Grund der Größe des Zustandsraumes eine Problemreduktion vor der eigentlichen Phase des *Simulated Annealing* notwendig. Der CLA, der in Abschnitt 5.1.1 vorgestellt wurde, bildet dafür eine gute Ausgangsbasis. Mit dem CLA kann eine günstige Aufteilung der Stationen in Teilnetze gefunden und der Zustandsraum erheblich eingeschränkt werden. Um eine vollständige Topologie zu erhalten, muß daran anschließend eine beliebige, aber zulässige Verbindungsstruktur der Teilnetze festgelegt werden. Ausgehend davon kann die Topologie schrittweise verändert werden. Dazu wird der Topologiegraph G_T (s. Abschnitt 4.3) betrachtet. Bei einer Transformation von G_T werden prinzipiell Knoten bzw. Kanten hinzugefügt oder entfernt. Um ein lokales Optimum zu erreichen, sind die folgenden Transformationen erforderlich:

- (a) Hinzufügen einer *Router-Router*-Kopplung zwischen zwei Teilnetzen, d. h. in den beteiligten Teilnetzen wird jeweils ein Knoten eingefügt; die beiden Knoten werden mit einer Kante verknüpft.
- (b) Entfernen einer *Router-Router*-Kopplung zwischen zwei Teilnetzen. Diese Transformation ist die Umkehroperation von (a).
- (c) Entfernen einer Station aus einem Teilnetz und Einfügen in ein anderes Teilnetz an einer zufällig gewählten Position. Diese Transformation erfordert lediglich das Entfernen und Hinzufügen von Kanten.
- (d) Invertieren der Reihenfolge zweier Netzelemente innerhalb eines Teilnetzes. Dazu müssen die beiden angrenzenden Kanten vertauscht werden (s. Bild 5.5).

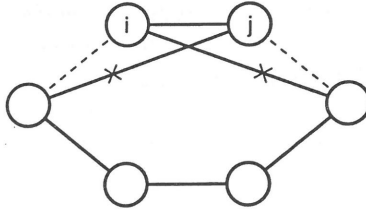


Bild 5.5: Invertierung der Reihenfolge in einem Teilnetz

Die Transformationen (a) und (b) sind erforderlich, um die Verbindungsstruktur zwischen den Teilnetzen zu optimieren. Die zugehörige Nachbarschaft enthält höchstens $M(M-1)/2$ Konfigurationen, d. h. die maximale Anzahl möglicher *Router-Router*-Kopplungen. Im Fall von Transformation (a) muß je ein *Router* in den beteiligten Teilnetzen eingefügt werden. Die Platzierung der *Router* erfolgt so, daß die Standorte mit der geringsten Entfernung gewählt werden. Bei Transformation (b) findet der umgekehrte Vorgang statt. Für den Ausgleich ungünstiger Lastverhältnisse oder eine Verbesserung der Struktur des Dualbussystems hinsichtlich der Kosten wird Transformation (c) benötigt. Dabei wird eine Station einem anderen Teilnetz nur logisch zugeordnet; der Standort der Station bleibt unverändert. Diese Nachbarschaft enthält N Konfigurationen.

Durch Transformation (d) in Kombination mit Transformation (c) ist es möglich, die optimale Reihenfolge der Stationen und *Router* für den *Looped Bus* eines Teilnetzes zu finden. Prinzipiell kommt hierfür auch ein Vertauschungsmechanismus in Frage, wie er für das *Traveling Salesman Problem* vorgeschlagen wurde [109]. Dabei werden jeweils zwei beliebige Kanten des Graphen, der die Reiseroute durch die Städte darstellt, ausgetauscht. Die Kanten dürfen jedoch nur gegen Kanten ausgetauscht werden, die ebenfalls zu einer gültigen Reiseroute gehören. Diese Vertauschungsoperation kann auch auf jeden Teilgraphen von G_T , der ein Teilnetz enthält, angewandt werden. Die Kardinalität der damit verbundenen Nachbarschaft beträgt $n_i(n_i-3)/2$, wenn n_i die Zahl der Stationen und *Router* in Teilnetz S_i ist. Diese Zahl wächst quadratisch mit n_i und kann damit bei großer Dimension des Problems zu hohen Rechenzeiten führen. Eine weniger umfangreiche Nachbarschaft wird dagegen mit Transformation (d) erreicht, die eine Untermenge der oben genannten Transformation ist. Da die Positionen der Stationen bzw. der *Router* sich auch durch die anderen Transformationen ändern können, ist eine hinreichende Erreichbarkeit des lokalen Optimums gesichert. Die Kardinalität der Transformation (d) ist n_i .

Eine Transformation, die zusätzlich das Aufteilen eines Teilnetzes bzw. das Zusammenfügen zweier oder mehrerer Teilnetze umfaßt, weist keine weiteren Vorteile auf. Obwohl die Nachbarschaft einen scheinbar geringen Umfang hat – sie umfaßt alle M Teilnetze –, sind beim Teilen eines Teilnetzes bzw. beim Zusammenfügen zweier Teilnetze zum einen die Reihen-

folge der Stationen und zum anderen die Verbindungen zu den anderen Teilnetzen und die Position der *Router* zu bestimmen. Je nach Größe des zu teilenden bzw. des neu entstandenen Teilnetzes kann das sehr rechenaufwendig sein. Hinzu kommt, daß beispielsweise durch die Teilung eine hohe Wahrscheinlichkeit besteht, bereits gefundene und möglicherweise optimale Teilstrukturen wieder aufzulösen. Dies gilt vor allem dann, wenn die Transformationen bzw. die Teilnetze und Übertragungsabschnitte, die verändert werden sollen, zufällig ausgewählt werden.

Wie in Abschnitt 3.2.3 ausgeführt wurde, kann aus der Zahl der möglichen Konfigurationen einer Nachbarschaft ein Richtwert für die Zahl der Schritte pro Temperaturänderung Θ_i abgeschätzt werden. Die maximale Zahl der Konfigurationen, die mit den Transformationen (a) - (d) erreicht werden können, erhält man mit

$$n_T = N \cdot \frac{1}{2} M(M-1) \cdot \sum_{v_i} n_i. \quad (5.5)$$

Da n_i im Verlauf des Optimierungsprozesses durch das Entfernen und Einfügen der *Router* schwanken kann, ist eine genaue Angabe nicht möglich. Tendenziell wächst n_T quadratisch mit steigendem N . Bei den Untersuchungen hat sich aber herausgestellt, daß auch gute Ergebnisse erzielt werden, wenn n_T deutlich kleiner gewählt wird.

Zwei Spezialfälle erfordern eine etwas andere Vorgehensweise. Falls durch den CLA nur ein Teilnetz generiert wird, genügt die Vertauschungsoperation, die auch beim *Traveling Salesman Problem* Verwendung findet. Besteht die Starttopologie lediglich aus zwei Teilnetzen, sollte statt Transformation (a), die nicht benötigt wird, eine Transformation verwendet werden, die die beiden *Router* an zufällig gewählte Standorte verschiebt.

5.2.2 Ablaufschema

Ein Übersichtsschema des Verfahrens zeigt Bild 5.6. Eine Starttopologie wird mit Hilfe des *Clustering*-Verfahrens und einer beliebigen Vernetzungsstruktur, die die Nebenbedingungen erfüllt, entworfen. Für das *Cooling Schedule* ist ein Anfangswert der Zielfunktion sowie der Temperatur Θ erforderlich (s. Abschnitt 3.2.3). Der Anfangswert von Θ wird für eine gegebene Akzeptanzwahrscheinlichkeit aus der Bewertung von 100 Topologien, die zufällig aus der Starttopologie erzeugt werden, bestimmt.

Sind die Startparameter ermittelt, generieren die Mechanismen des lokalen Suchverfahrens sukzessive neue Topologien. Für jede der generierten Topologien wird das Verkehrslenkungsschema berechnet. Die daran anschließende Fluß- bzw. Transferzeitanalyse dient der Überprüfung der Nebenbedingungen. Sind die Nebenbedingungen für eine neu erzeugte Topologie verletzt, wird die bereits vorhandene Topologie beibehalten. Sind die Nebenbedingungen erfüllt, und wird außerdem die geänderte Topologie durch den *Annealing*-Mechanismus akzeptiert, wird diese Topologie als nun gültige Topologie angenommen.

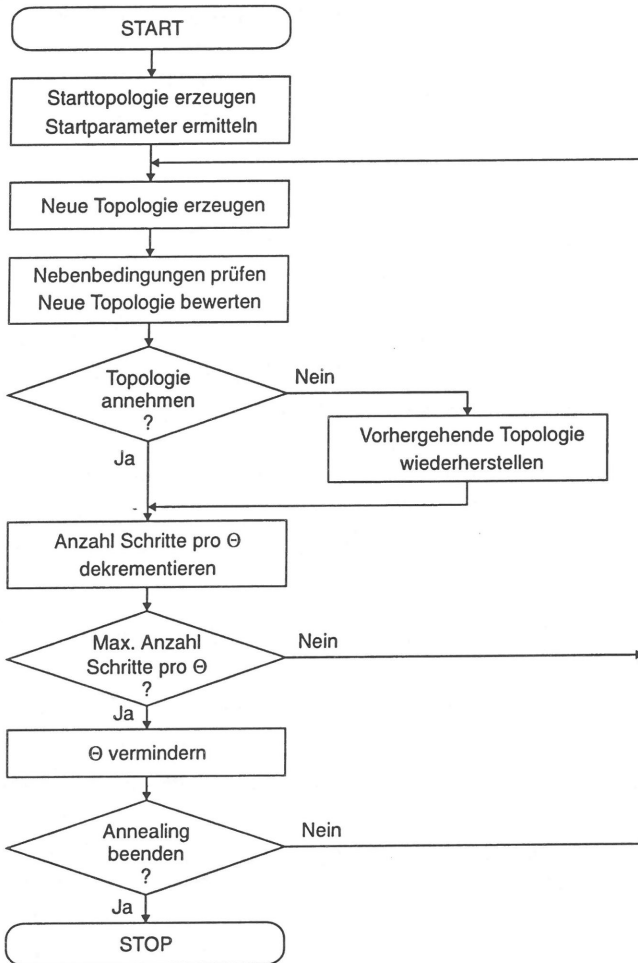


Bild 5.6: Optimierung der Netztopologie durch Simulated Annealing

Eine Absenkung der Temperatur erfolgt, wenn am Ende einer Iteration die maximale Zahl von Iterationen pro Temperaturschritt erreicht ist. Ein Abbruch des Verfahrens erfolgt, wenn sich keine signifikanten Verbesserungen der Zielfunktion mehr ergeben oder wenn eine vorher definierte Zahl von Temperaturschritten erreicht ist.

Um das Verfahren flexibel zu halten, können die Transformationen, die das lokale Suchverfahren bereitstellt, miteinander kombiniert werden. So kann es beispielsweise sinnvoll sein,

bei bestimmten Konfigurationen, die durch den CLA erzeugt wurden, nur noch Transformationen zur Verbesserung der Reihenfolge und zur Bestimmung des Vermaschungsgrades der Teilnetze zu verwenden.

5.3 Verfahren basierend auf Genetischen Algorithmen

Genetische Algorithmen sind ein weiterer Ansatz, um das MSS-Entwurfsproblem zu lösen. Voraussetzung ist, daß eine geeignete Abbildung des Problems in Form von Zeichenketten gefunden wird, und Genetische Operatoren zur Verfügung stehen, die für diese Problemrepräsentation geeignet sind. Dabei gibt es beim Entwurf eines spezifischen Algorithmus eine ganze Reihe von Freiheitsgraden, die die Leistungsfähigkeit Genetischer Algorithmen beeinflussen. Folgende Gesichtspunkte, die anhand von einfachen Planungsproblemen näher untersucht wurden (vgl. [86]), sollten beim Entwurf eines speziellen Optimierungsverfahrens beachtet werden:

- Nicht alle Probleme können „natürlicherweise“ durch binäre Zeichenketten codiert werden. In manchen Fällen führt die binäre Codierung zu sehr langen Zeichenketten, die die Genetischen Algorithmen ineffektiv werden lassen.
- Falls die Nebenbedingungen nicht in der Codierung berücksichtigt werden, können auch Individuen entstehen, die die Nebenbedingungen verletzen. Zwei Strategien sind denkbar: Entfernen der Individuen aus der Population bzw. Bewertung der Individuen mit zusätzlichen Strafkosten (*Penalty*-Verfahren). Im ersten Fall sind die Nachteile ersichtlich: Zum einen wird durch das Erzeugen ungültiger Individuen die Rechenzeit verlängert und zum anderen gehen Informationen, die für das globale Optimum wichtig sein können, verloren. Daher ist ein Strafkostenverfahren zu bevorzugen, bei dem alle Individuen berücksichtigt werden.
- Die Skalierungsfunktion ist mit besonderer Sorgfalt auszuwählen, da zwischen den verschiedenen Funktionen (linear, nichtlinear) deutliche Unterschiede bestehen.
- Auf den Mutationsoperator sollte nicht verzichtet werden, weil damit verlorengegangene Informationen wieder erzeugt werden können. Bei einer Codierung mit natürlichen Zahlen ist der Mutationsoperator geeignet anzupassen.

Anhand dieser Richtlinien wurde ein Verfahren, das auf Genetischen Algorithmen beruht, entworfen.

5.3.1 Codierung und Genetische Operatoren

Wesentliches Ziel der Codierung ist es, eine möglichst einfache Abbildung der Problemstruktur auf Zeichenketten zu finden, die in einer effizienten Form durch Genetische Operatoren bearbeitet werden können. Da die Nebenbedingungen nicht vollständig in codierter Form darzustellen sind, – das betrifft vor allem die Nebenbedingungen für den Verkehrsfluß und die maximal zulässige Transferzeit für eine Kommunikationsbeziehung – orientiert sich das Codierungsschema am Topologiegraphen.

Das Codierungsschema folgt zwei Prinzipien. Zum einen soll es möglich sein, die Zahl der Teilnetze zu variieren, um durch den Suchmechanismus der Genetischen Algorithmen die günstigste Konfiguration eines MSS zu ermitteln. Dies wurde bei *Simulated Annealing* nur indirekt durch den vorgeschalteten CLA ermöglicht. Zum anderen sollte die Optimierung der Reihenfolge der Stationen und *Router* in den Teilnetzen unterstützt werden. Dafür erscheint der PMX-Operator in Kombination mit einem Mutationsoperator geeignet. Der PMX-Operator ist als sog. Ordnungsoperator schon bei ähnlich gelagerten Problemen erfolgreich eingesetzt worden (vgl. [115]). Um ihn zu verwenden, ist eine Codierung mit natürlichen Zahlen notwendig.

Prinzipiell müssen Stationen, *Router* und Teilnetze sowie die Zuordnungen zwischen diesen Netzelementen repräsentiert werden. Jedes Netzelement ist eindeutig zu kennzeichnen. Die Zuordnung von Netzelementen zu Teilnetzen und die Reihenfolge ihrer Verknüpfung, d. h. letztendlich die Kennzeichnung von Übertragungsabschnitten, wird durch den Wirkungsort festgelegt. Das Codierungsschema umfaßt drei Teile: Stationen werden mit den Elementen der Menge $\{1 \dots N\}$ codiert. Um die Grenzen zwischen Teilnetzen zu markieren, werden sog. Separatoren eingesetzt, die Elemente der Menge $\{N + 1 \dots N + M\}$ sind. Schließlich werden *Router* mit Elementen der Menge $\{N + M + 1 \dots N + M(M - 1)\}$ codiert. Die Zahl der Separatoren, die als Optimierungsparameter vorgegeben wird, begrenzt die maximale Zahl von Teilnetzen.

Bild 5.7 zeigt beispielhaft einen Topologiegraphen mit drei Teilnetzen, die vollständig vermascht sind. Das korrespondierende Codierungsschema lautet:

$$\{1 \ 10 \ 12 \ 2 \ |7| \ 11 \ 3 \ 4 \ 14 \ |8| \ 15 \ 5 \ 6 \ 13 \ |9|\}$$

Die Elemente $\{7, 8, 9\}$ stellen die Separatoren für die drei Teilnetze dar. Die Elemente $\{10 \dots 15\}$ kennzeichnen die *Router*. Ihre Codierung folgt den aufsteigenden Nummern, die eine Verbindung zwischen Teilnetzen festlegen. So wird beispielsweise die Verbindung (Teilnetz 1 - Teilnetz 2) durch die *Router* 10 und 11 markiert. Anzumerken bleibt, daß die erste Position in einem Teilnetz immer den Beginn von Bus A kennzeichnet.

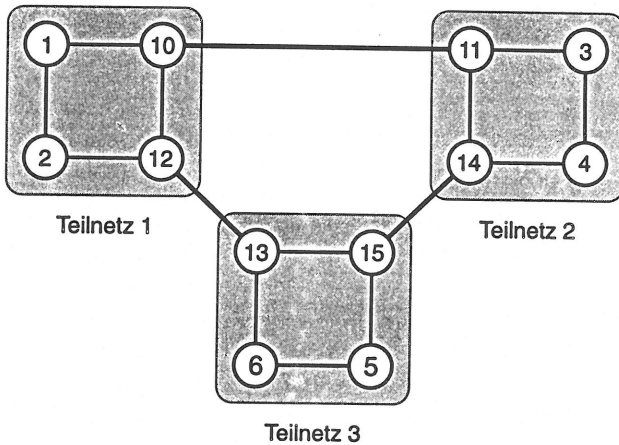


Bild 5.7: Beispiel zur Codierung einer MSS-Topologie

Gegenüber der binären Darstellung, bei der die Existenz einer Kante des Topologiegraphen mit einer „1“ markiert wird, weist dieses Schema zwei Vorteile auf. Zum einen ist die Zeichenkette in kompakter Form codiert. Damit wird vermieden, daß bereits bei zweistelligen Stationszahlen sehr lange Zeichenketten auftreten, die wiederum zu erhöhten Rechenzeiten führen. Zum anderen können die Nebenbedingungen, die die Topologie betreffen, durch problemspezifische Anpassungen geeignet berücksichtigt werden. So kann die Wahrscheinlichkeit für das Auftreten unzulässiger Topologien gesenkt werden. Ein Beispiel ist die Anpassung des Mutationsoperators (s. unten).

Als genetische Operatoren werden der PMX-Operator und eine abgewandelte Form der Mutation verwendet. Da die zufällige Änderung eines beliebigen Gens in ein anderes Gen zu seiner Duplizierung und damit zu einem ungültigen Individuum führt, wird als Mutationsoperator ein spezieller Vertauschungsmechanismus unterstützt. Dabei wird ein Gen, das eine Station bzw. einen Separator repräsentiert, mit einer zufällig gewählten Station bzw. einem zufällig gewählten Separator vertauscht. Bei der Auswertung der Individuen werden außerdem Gene, die Router darstellen, nur dann berücksichtigt, wenn sich eine zulässige Verbindung zwischen den zugehörigen Teilnetzen ergibt.

Wie schon dargelegt, ist für das MSS-Entwurfsproblem eine vollständige Berücksichtigung aller Nebenbedingungen in der Codierung nicht möglich. Deshalb wird ein Strafkostenverfahren zur Bewertung ungültiger Individuen verwendet, wobei sowohl Individuen, die einen ungültigen Topologiegraphen repräsentieren, als auch Individuen, die andere Nebenbedingungen verletzen, mit zusätzlichen Netzkosten beaufschlagt werden. Dementsprechend verringert sich auch ihre Fitness.

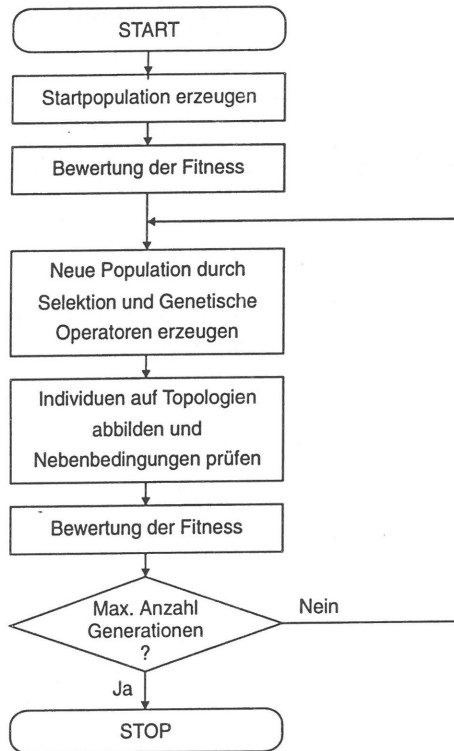


Bild 5.8: Optimierung der Netztopologie durch Genetische Algorithmen

5.3.2 Ablaufschema

Der Ablauf des Verfahrens, der in Bild 5.8 schematisch gezeigt ist, wird im wesentlichen vom Erzeugungsmechanismus der Genetischen Algorithmen bestimmt. Zu Beginn wird eine Startpopulation generiert, die aus einer bestimmten Zahl von zulässigen Zeichenketten (Individuen) besteht, und ihre Fitness bewertet. Die Startpopulation hat einen wesentlichen Einfluß auf die Leistungsfähigkeit der Genetischen Algorithmen. Das Codierungsschema kann am effektivsten ausgenutzt werden, wenn die Startpopulation bereits Individuen mit günstigen Eigenschaften bezüglich der Aufteilung in Teilnetze und einer genügenden Variationsvielfalt enthält. Zur Erzeugung solcher Individuen dient das Dekompositionsverfahren. Durch die zufällige Variation seiner Parameter kann ein breites Spektrum von Individuen mit einer relativ hohen Fitness erzeugt werden. Die Kombination ihrer Eigenschaften im Selektionsprozeß und die Optimierung der Reihenfolge in den Teilnetzen sollte zu günsti-

gen Ergebnissen führen. Gegebenenfalls können auch zufällig erzeugte Individuen in die Startpopulation eingefügt werden.

Durch die iterative Manipulation der Startpopulation mit Hilfe der Genetischen Operatoren entwickeln sich neue Populationen. Bei jeder Generation werden die Zeichenketten in Topologien umgesetzt, die Einhaltung der Nebenbedingungen geprüft und die Topologien mit Hilfe der Zielfunktion bewertet. Aus den Werten der Zielfunktion wird wiederum die Fitness der einzelnen Individuen abgeleitet. Ein Abbruch erfolgt, wenn eine bestimmte Zahl von Generationen erreicht ist.

Für den Entwurf des Algorithmus wurden mehrere Skalierungsfunktionen an einfachen Planungsproblemen getestet (s. [86]). Als geeignete Skalierungsfunktion hat sich eine Skalierung herausgestellt, die aus der Zielfunktion nach Gleichung 4.18 mit

$$\tilde{f}_i = \frac{1}{K^n} \quad (5.6)$$

die Fitness eines Individuums i berechnet. Die Fitness ist demnach umso höher, je geringer die Kosten einer Topologie sind. Als Erfahrungswert eignet sich $n = 3$.

Dieses Verfahren besitzt neben den Parametern der Genetischen Algorithmen, d. h. neben der Populationsgröße (Zahl der Individuen), der PMX- und der Mutationswahrscheinlichkeit sowie der Zahl der Generationen, mit der Zahl der Separatoren eine weitere Einflußgröße. Damit ist ein zusätzlicher Freiheitsgrad bei der Suche im Zustandsraum verbunden.

5.4 Vergleich der Verfahren

In diesem Abschnitt sollen die vorgestellten Verfahren an charakteristischen Beispielen, die konkrete Ausformungen des allgemeinen Entwurfproblems repräsentieren, untersucht und ihre Eigenschaften verglichen werden. Üblicherweise sind die Qualität der erzielten Lösungen und die Laufzeit der Algorithmen von vorrangigem Interesse. Darüber hinaus werden aber auch Zusammenhänge zwischen den Netzkosten und den Dienstgüteparametern aufgezeigt, die der Netzplaner beim Entwurf eines MAN-Vermittlungssystems berücksichtigen sollte.

5.4.1 Parameter der Netzelemente

Zunächst wird von einer vermaschten MSS-Topologie ausgegangen, deren Teilnetze eine nominale Kapazität von 34 Mbit/s aufweisen. Damit beträgt die Nettokapazität, die auf Grund der Rahmenstruktur des Übertragungssystems der MAC-Teilschicht zur Verfügung steht, 30,528 Mbit/s. Der maximale Durchsatz der Router wird mit 30 Mbit/s in der gleichen

Größenordnung angenommen. Die Lastgrenzen ρ_S und ρ_R (s. Abschnitt 4.3) werden gleich 1 gewählt.

Die Paketlänge beträgt 53 Byte, so daß ein Paket vollständig in einem DQDB-Zeitschlitz übertragen werden kann. Für die Berechnung der mittleren Transferzeiten stellt das eine Näherung dar, da die mittleren Paketgrößen in lokalen Netzen deutlich über 100 Byte liegen (vgl. [27]), was jedoch für den Vergleich, bei dem primär das prinzipielle Verhalten interessiert, von untergeordneter Bedeutung ist. Wie in Abschnitt 4.2.4 ausgeführt wurde, kann die Protokollbearbeitungszeit in einem *Router* durch ein M/G/1-System modelliert werden. Dazu muß die mittlere Bedienzeit bestimmt werden. Diese ergibt sich aus dem Verhältnis von Paketlänge zu maximalem *Router*-Durchsatz; der Variationskoeffizient sei gleich 1. Die Protokollbearbeitungszeit in einer Station sei hingegen vernachlässigbar, d. h. für eine Station wird lediglich die Buszugriffszeit für die Transferzeiten berücksichtigt.

Als Verkehrslenkungsschema wird das bereits in Abschnitt 4.2.3 beschriebene *Minimum-Hop*-Verfahren gewählt. Dieses Schema wird auch zur Bestimmung der Pfade in denjenigen Topologien verwendet, die mit Hilfe des Dekompositionsverfahrens ermittelt werden. Für die Netzstruktur ergibt sich außerdem die Beschränkung, daß ein Teilnetz mindestens zwei Stationen enthalten muß. Die Kosten eines Übertragungsabschnitts werden nach Gleichung 4.1 ermittelt und auf eine „Kosteneinheit“ (KE) normiert. Zu den Leitungskosten tragen die Fixkosten mit 1 KE für jeden Übertragungsabschnitt bei. Zwischen den Standorten sind alle möglichen Trassen zugelassen, so daß keine Beschränkungen durch die übertragungstechnische Infrastruktur bestehen.

5.4.2 Vergleichsstrategie

Die Untersuchungen stützen sich auf möglichst realitätsnahe Annahmen für die Verkehrsmatrizen und die geographische Verteilung der Stationen bzw. ihrer Standorte, da Messungen bisher nur an MANs mit einer geringen Zahl von Teilnehmern vorgenommen wurden (s. z. B. [11]) und somit nicht in ausreichendem Umfang zur Verfügung stehen.

Exemplarisch werden für den Vergleich drei Probleme ausgewählt. Die ersten beiden Probleme, Beispiel B_1 und Beispiel B_2 , (s. Tabelle 5.1) sollen das Verhalten der Algorithmen bei einer zufälligen Wahl der Verkehrsmatrix und der geographischen Verteilung der Stationen zeigen, um die Algorithmen unabhängig von spezifischen Eigenschaften der Verkehrs- und Entfernungsmatrix zu testen. Dazu werden die Standorte zufällig in einem Gebiet von 100×100 km Ausdehnung verteilt. Die Verkehrsmatrix ist gleichförmig, zufällig verteilt, d. h. jedes Matricelement, das den Verkehrsfluß einer Verkehrsbeziehung charakterisiert, besitzt einen Zufallswert aus dem Intervall $(0, \gamma_{max}]$.

Das dritte Problem (B_3) stellt ein Modell für Meßergebnisse dar, wie sie in Datennetzen im Weitverkehrsbereich gewonnen wurden [27]. Die Messungen zeigen, daß die Verkehrsflüsse

Parameter	Problem		
	B_1	B_2	B_3
Zahl der Stationen	30	50	50
Zahl der Zentren	0	0	5
Geographische Ausdehnung [km]	100 × 100	100 × 100	100 × 100
Verkehrsfluß	zufällig	zufällig	$\sim 1/d_{ij}$
γ_{max} [Mbit/s]	0.3	0.1	-
t_{max}^R [ms]	4.0	5.5	4.5

Tabelle 5.1: Allgemeine Parameter der charakteristischen Probleme

in ihrer geographischen Verteilung typischerweise nicht gleichförmig sind, sondern daß ein Großteil des Verkehrs innerhalb weniger Verkehrsbeziehungen erzeugt wird (Lokalität des Verkehrs). Dieser Effekt, der auch im regionalen Bereich auftritt, kann beim Entwurf von MAN-Vermittlungssystemen genutzt werden, um Stationen mit hohem wechselseitigen Verkehrsaufkommen an ein gemeinsames Teilnetz anzuschließen. Damit wird der Transitverkehr deutlich reduziert.

Zur Modellierung wird die Verkehrsmatrix durch einen Graphen mit N Knoten dargestellt, dessen Kanten die Verkehrsbeziehungen zwischen den Stationen repräsentieren und zufällig bestimmt werden. Die Wahrscheinlichkeit P_E , daß eine Verkehrsbeziehung zwischen zwei Stationen i und j besteht, erhält man mit

$$P_E = a \cdot \exp\left(\frac{-d_{ij}}{b \cdot d_m}\right), \quad (5.7)$$

wobei a und b Gewichtungsfaktoren zwischen $0 \dots 1$ sind, d_{ij} die Entfernung zwischen den beiden Knoten und d_m eine Entfernungskonstante bezeichnet. Mit Hilfe der Faktoren a und b in Gleichung (5.7) kann sowohl die Zahl der Verkehrsbeziehungen einer Station als auch deren „Reichweite“ begrenzt werden. Für Problem B_3 wird $a = b = 0.5$ und $d_m = 40$ km gewählt. Der Verkehrsfluß ist umgekehrt proportional zur Entfernung der Stationen. Ein ähnliches Modell wurde bereits bei der Simulation von Netzgraphen verwendet [89].

Zur Erzeugung geographischer Gruppierungen werden zunächst fünf Standorte ausgewählt und gleichmäßig in einem quadratischen Gebiet verteilt. Die übrigen Standorte werden in fünf Teilen zufällig in einem bestimmten Radius um diese Zentren angeordnet. Die Gruppierungen können im weitesten Sinne als Modell für Ballungszentren dienen. Innerhalb der Gruppierungen ist der Verkehr relativ hoch, zu anderen Gruppierungen nimmt der Verkehr und die Zahl der Verkehrsbeziehungen deutlich ab.

Üblicherweise wird die Qualität eines heuristischen Verfahrens durch die Differenz zwischen der besten zulässigen Lösung und einer unteren Schranke gemessen. Für das MSS-Entwurfsproblem ergibt sich eine untere Schranke, wenn alle Stationen durch einen *Minimum*

Spanning Tree verbunden werden, dessen Kanten durch die Leitungskosten gewichtet werden. Eine solche Baumstruktur weist die geringsten Kosten aller möglichen Netzstrukturen auf (vgl. [92]). Allerdings kann die untere Schranke sehr ungenau sein, weil die Verkehrsmatrix, die die Größe und Zahl der Teilnetze stark beeinflusst, nicht berücksichtigt wird.

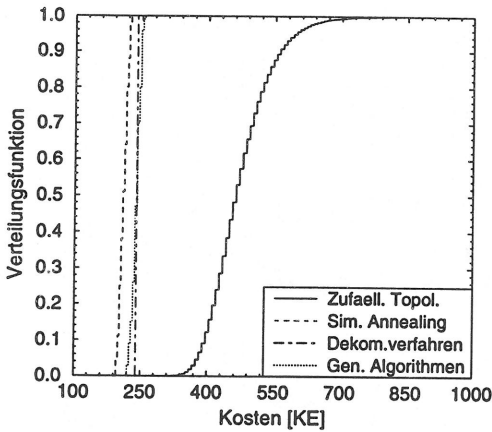
Daher werden die beschriebenen Verfahren in erster Linie mit einer repräsentativen Stichprobe aus der Menge aller Topologien verglichen, die die Nebenbedingungen (Gleichungen 4.19 - 4.27) erfüllen, und deren Teilnetze mindestens zwei Stationen enthalten. Die Kosten der Topologien werden nach Gleichung (4.18) bewertet. Die Stichproben umfassen jeweils 100 000 Werte und werden in 100 Klassen eingeteilt.

Die Verfahren des *Simulated Annealing* bzw. der Genetischen Algorithmen können auf Grund ihrer statistischen Suchmechanismen unterschiedliche Lösungen finden. Um von den Ergebnissen einzelner Läufe unabhängig zu sein, werden jeweils 100 Lösungen ermittelt, wobei der Zufallszahlengenerator mit unterschiedlichen Startwerten initialisiert wird. Die Kenngrößen dieser Stichproben dienen zum Vergleich mit den Werten der zufällig ermittelten Topologien. Soweit nicht anders angegeben, werden die Parameter für *Simulated Annealing* wie folgt eingestellt: Der Startwert der Akzeptanzwahrscheinlichkeit P_A beträgt 0.3, der Faktor τ sei 0.95 und die Zahl der Temperaturschritte 150. Für die Genetischen Algorithmen lauten die Parametereinstellungen: Die Wahrscheinlichkeit für den PMX-Operator beträgt 0.55, die Mutationswahrscheinlichkeit 10^{-3} ; die Populationsgröße umfaßt in der Regel 200 Individuen, die Zahl der Generationen schwankt zwischen 100 und 150. Die Zahl der Separatoren wird problemabhängig angepaßt.

Für das Dekompositionsverfahren wird der Wert gewählt, der bei der günstigsten Parametereinstellung auftritt. Hier gilt, daß Parameter, die die Kosten direkt beeinflussen, wie etwa der Gewichtungsfaktor g_d oder der Schwellwert ρ_g , für alle Beispiele die gleichen Werte aufweisen ($g_d = 0$ bzw. $\rho_g = 0.6 \dots 0.9$), die anderen Parameter aber problemabhängig anzupassen sind.

5.4.3 Vergleich mit zufällig ermittelten Topologien

Bild 5.9 zeigt die Verteilungen und deren Maßzahlen für das erste Problem. Die Werte der Topologien, die mit den drei Entwurfsverfahren erreicht werden, liegen deutlich unter denjenigen, die zufällig erzeugt werden. Gegenüber dem Minimum der Verteilung der zufällig erzeugten Topologien ergibt sich eine Verbesserung zwischen 8 Prozent (95. Perzentil, Genetische Algorithmen) und 27 Prozent (5. Perzentil, *Simulated Annealing*). Das Verfahren mit *Simulated Annealing* erzielt für dieses Beispiel die besten Werte. Auf Grund der Größe des Zustandsraums ist die Wahrscheinlichkeit, zufällig eine gute Topologie zu erzeugen, sehr gering. Eine Überlappung der Verteilungen der Entwurfsverfahren mit der Verteilung, die



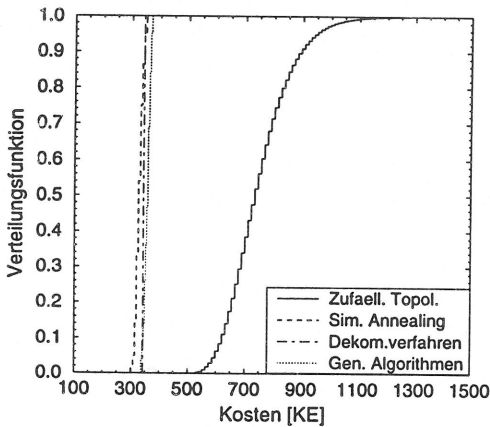
	Simul. Anneal.	Genet. Algor.
Mittelwert	208 KE	237 KE
5. Perzentil	196 KE	221 KE
95. Perzentil	220 KE	248 KE

Dekomp.verfahren	239 KE
Min. zufällige Topol.	270 KE
Untere Schranke	121 KE

Bild 5.9: Verteilungsfunktionen und Maßzahlen der Kosten für Problem B_1

die zufällig erzeugten Topologien repräsentiert, ist daher nicht zu erwarten. Das Dekompositionsverfahren ist mit einem Wert von 239 KE nicht wesentlich schlechter als der Mittelwert, der mit Genetischen Algorithmen erzielt werden kann.

Darüber hinaus weisen die Verteilungen für die Verfahren des *Simulated Annealing* bzw. der Genetischen Algorithmen lediglich eine geringe Schwankungsbreite auf. Das bedeutet, daß für einen guten Entwurf bereits wenige Optimierungsläufe genügen. Der Vergleich mit der unteren Schranke läßt eine deutliche Entfernung zum globalen Optimum vermuten. Wird



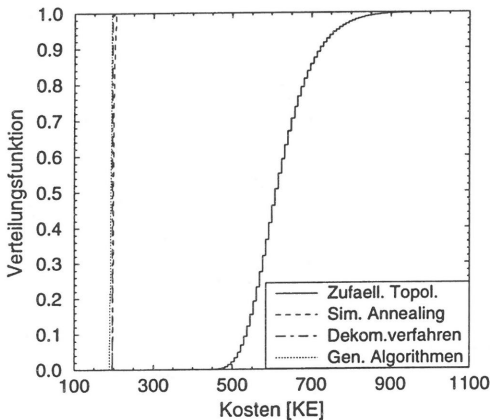
	Simul. Anneal.	Genet. Algor.
Mittelwert	325 KE	354 KE
5. Perzentil	308 KE	337 KE
95. Perzentil	342 KE	365 KE

Dekomp.verfahren	343 KE
Min. zufällige Topol.	460 KE
Untere Schranke	196 KE

Bild 5.10: Verteilungsfunktionen und Maßzahlen der Kosten für Problem B_2

jedoch für das Problem B_1 der Wert γ_{max} auf 0.05 gesenkt, sind wesentlich niedrigere Netzkosten zu erzielen, z. B. ein Mittelwert von 161 KE in 10 Optimierungsläufen von *Simulated Annealing*. Für das Verkehrsvolumen, das bei B_1 auftritt, werden folglich mehr Netzelemente und Übertragungsleitungen benötigt als bei einem vergleichbaren Problem, das ein geringeres Verkehrsvolumen als Eingangsgröße hat.

Ein ähnliches Resultat zeigt Bild 5.10, wenn die Zahl der Stationen von 30 auf 50 erhöht wird. Dabei nimmt der gesamte Verkehrsfluß zu, so daß mehr Teilnetze und Übertragungsabschnitte benötigt werden, und damit die Kosten steigen. Auch hier erzielt *Simulated Annealing* die besseren Ergebnisse, während die besten Werte, die mit Genetischen Algorithmen und mit dem Dekompositionsverfahren erreicht werden, dichter an den 90-Prozent-Bereich des *Simulated-Annealing*-Verfahrens heranrücken. Gegenüber dem Minimum der zufälligen Verteilung sind nun Verbesserungen zwischen 21 Prozent (95. Perzentil, Genetische Algorithmen) und 33 Prozent (5. Perzentil, *Simulated Annealing*) zu erreichen. Der Abstand zu den Werten der zufällig bestimmten Topologien nimmt im Vergleich mit B_1 weiter zu.



	Simul. Anneal.	Genet. Algor.
Mittelwert	199 KE	192 KE
5. Perzentil	196 KE	189 KE
95. Perzentil	203 KE	194 KE

Dekomp.verfahren	197 KE
Min. zufällige Topol.	409 KE
Untere Schranke	158 KE

Bild 5.11: Verteilungsfunktionen und Maßzahlen der Kosten für Problem B_3

Die Lösungen für Problem B_3 hingegen weisen eine weitgehende Übereinstimmung bei allen drei Verfahren auf (vgl. Bild 5.11). Die Verteilungen liegen eng beieinander, wobei nun das Verfahren mit Genetischen Algorithmen die besten Lösungen erbringt. Da durch den *Clustering*-Prozeß die geographischen Gruppierungen und die Struktur der Verkehrsmatrix vorteilhaft zum Aufbau einer Topologie genutzt werden kann, genügen für das Verfahren mit *Simulated Annealing* die Transformationen, die die Verbindungsstruktur zwischen den Teilnetzen und innerhalb der Teilnetze optimieren. Aus dem gleichen Grund erzielt auch das Dekompositionsverfahren ein gutes Resultat.

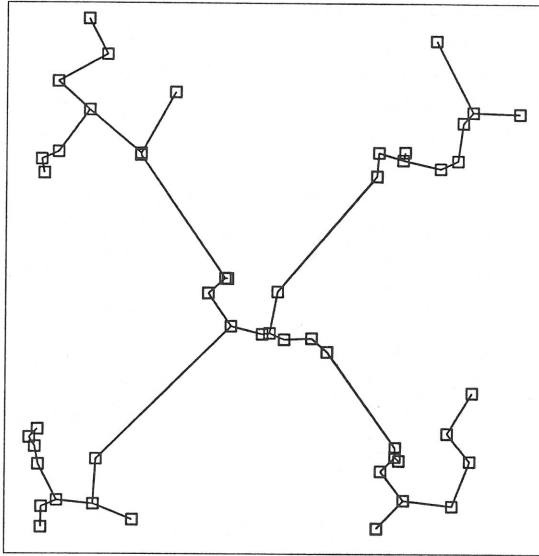


Bild 5.12: Untere Schranke für Problem B_3

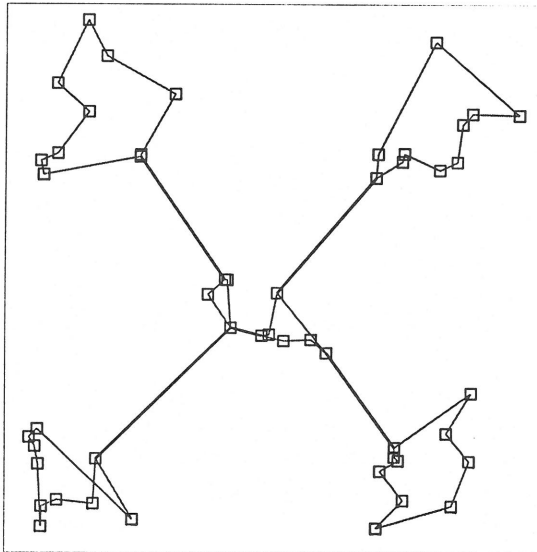


Bild 5.13: Exemplarische Topologie (Problem B_3)

Die Verbesserung gegenüber dem Minimum der zufällig ermittelten Topologien beträgt im günstigsten Fall 54 Prozent (5. Perzentil, Genetische Algorithmen) und im ungünstigsten 50 Prozent (95. Perzentil, *Simulated Annealing*). Das läßt die Aussage zu, daß alle drei Verfahren in der Lage sind, sowohl die Lokalität des Verkehrs als auch günstige geographische Verteilungen vorteilhaft für eine Optimierung der Netztopologie auszunutzen. Auch bei etwas geänderten geographischen Grupperungen konnten die Ergebnisse bestätigt werden.

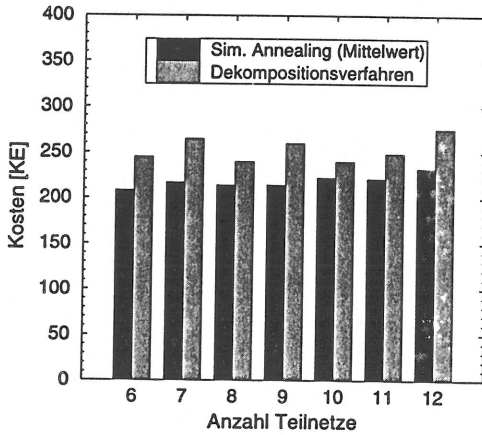
Der Vergleich mit der unteren Schranke legt die Vermutung nahe, daß die besten Ergebnisse, die um 19 Prozent über dem Wert der unteren Schranke liegen, nur geringfügig vom Optimum entfernt sind. Die graphische Darstellung des *Minimum Spanning Tree* (Bild 5.12), der die untere Schranke für Problem B_3 bildet, bestätigt das, wenn der *Minimum Spanning Tree* mit einer exemplarischen Topologie, die mit Genetischen Algorithmen erzeugt wurde (Bild 5.13) verglichen wird¹. Die Topologie in Bild 5.13 umfaßt sechs Teilnetze und wird mit 190 KE bewertet.

5.4.4 Einfluß der Startbedingungen

Netzstrukturen, die durch das Dekompositionsverfahren oder mit dem CLA-Algorithmus in Kombination mit *Simulated Annealing* ermittelt werden, sind von der Zahl der Teilnetze abhängig. Es stellt sich daher die Frage, ob die Qualität dieser beiden Verfahren starken Schwankungen ausgesetzt ist, wenn die Zahl der Teilnetze variiert wird. Für das Problem B_1 wurden dazu verschiedene Konfigurationen untersucht (s. Bild 5.14), wobei die Zahl der Teilnetze durch geeignete Parametereinstellungen des CLA-Algorithmus (D_{max} und ρ_{CI}) verändert wurde. Für das Verfahren mit *Simulated Annealing* wird der Mittelwert sowie 5. und 95. Perzentil aus 100 Lösungen ermittelt.

Der Vergleich bestätigt zunächst die signifikante Verbesserung von Topologien, die mit *Simulated Annealing* generiert werden, gegenüber dem Dekompositionsverfahren. Dies ist im wesentlichen auf die Fähigkeit von *Simulated Annealing* zurückzuführen, die Anordnung in den Teilnetzen global zu optimieren. Das ist auch daran erkennbar, daß die besten Werte des Dekompositionsverfahrens bei 8 bzw. 10 Teilnetzen erzielt werden, d. h. wenn tendenziell weniger Stationen in einem Teilnetz sind. Während das *Simulated-Annealing*-Verfahren über den gesamten Bereich sich nur um maximal 11.5 Prozent verschlechtert, zeigt das Dekompositionsverfahren etwas größere Schwankungen: Das Maximum liegt um 15 Prozent über dem Minimum. Anzumerken ist, daß alle Lösungen bis auf die Topologie, die das Dekompositionsverfahren bei 12 Teilnetzen erzeugt, unter dem Minimum der zufälligen Verteilung bleiben. Beide Verfahren sind damit relativ robust gegenüber Änderungen der oben genannten Parameter.

¹In Bild 5.13 sind Router-Router-Kopplungen durch breite Linien dargestellt, Übertragungsleitungen in Teilnetzen durch schmale Linien.



Anzahl Teilnetze	Kosten [KE]	
	5. Perz.	95. Perz.
6	196	220
7	202	230
8	201	232
9	202	229
10	205	240
11	210	236
12	220	248

Bild 5.14: Netzkosten in Abhängigkeit von der Zahl der Teilnetze (B_1)

Der Anstieg der Werte bei einer Steigerung der Anzahl der Teilnetze ist aus dem zusätzlich erzeugten Transitverkehr zu erklären. Die Ergebnisse deuten darauf hin, daß das „Auffüllen“ von Teilnetzen mit möglichst vielen Stationen zu kostengünstigen Netzstrukturen führt, da dann der Transitverkehr gering bleibt und damit weniger *Router-Router-Kopplungen* benötigt werden. Allerdings ist dabei auch eine hohe mittlere Last auf einem Bus zu erwarten, was wiederum zu erhöhten Transferzeiten führt. Ein exemplarisches Beispiel zeigt Tabelle 5.2 für jeweils eine Lösung von B_1 mit 6 Teilnetzen. Die Lösung, die mit *Simulated Annealing* erzielt wurde, weist eine deutlich höhere mittlere Auslastung auf. Der Zusammenhang zwischen den Netzkosten und den Transferzeiten einer Topologie wird in Abschnitt 5.4.5 genauer untersucht.

	Anzahl Teilnetze	Kosten [KE]	Mittlere Auslastung	Mittlere Transferzeit [ms]
Dekompositionsverfahren	6	245	0.70	1.81
Simulated Annealing	6	209	0.84	3.82

Tabelle 5.2: Exemplarischer Vergleich von zwei Lösungen für Problem B_1

Auch die Starttopologien, die für die Verfahren des *Simulated Annealing* bzw. der Genetischen Algorithmen erzeugt werden, haben Einfluß auf das erzielte Ergebnis. Für die Genetischen Algorithmen wird die Startpopulation mit Individuen, die durch das Dekompositionsverfahren erzeugt werden, initialisiert (s. Abschnitt 5.3.2). Auch *Simulated Annealing* erhält durch den CLA eine vorteilhafte Starttopologie. Um den Einfluß der Starttopologie

zu verdeutlichen, zeigt Bild 5.15 beispielhaft ein Histogramm aus 100 Lösungen für Problem B_3 , die für *Simulated Annealing* mit jeweils einer zufälligen Starttopologie und einer Starttopologie, die mit dem CLA generiert wurden. Jede Netzstruktur enthält sechs Teilnetze.

In beiden Fällen wird das Minimum der Verteilung der zufälligen Topologien deutlich unterboten. Darüber hinaus erzielt das *Simulated-Annealing*-Verfahren in Kombination mit dem CLA eine weitere, signifikante Verbesserung gegenüber einem zufälligen Startwert. Das gilt auch dann, wenn nur das 95. Perzentil betrachtet wird. Dabei muß der Rechenaufwand bei einer zufälligen Starttopologie erheblich gesteigert werden: Der Faktor τ wurde von 0.95 auf 0.96 erhöht, und für eine Lösung wurde im Mittel um den Faktor 2.3 mehr CPU-Zeit als mit *Simulated Annealing* mit CLA benötigt. Mit dem CLA ist also auch eine Verringerung der Rechenzeiten verbunden.

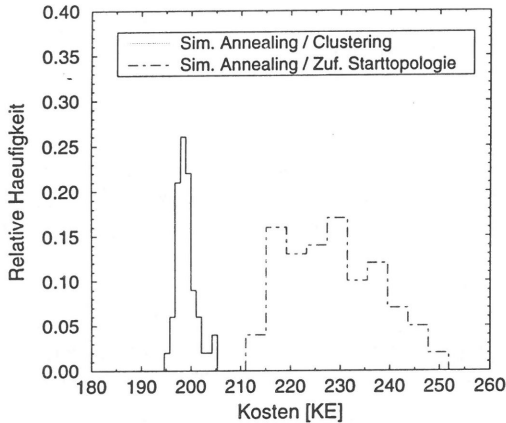


Bild 5.15: Vergleich unterschiedlicher Starttopologien für Problem B_3

Für die Probleme B_1 bzw. B_2 ist dieses Verhalten nur schwach ausgeprägt. Die Ursache liegt im wesentlichen in den Eigenschaften des lokalen Suchverfahrens begründet, da bei den Transformationen keinerlei problemspezifische Information (Verkehrsfluß, Geographie) genutzt wird, und die Transformationen mit gleicher Wahrscheinlichkeit ausgewählt werden. Der CLA kann somit vor allem bei Strukturen, wie sie Problem B_3 darstellt, gewinnbringend eingesetzt werden.

Der Entwurf eines MAN-Vermittlungssystems kann sich nicht ausschließlic an einem einzigen Kriterium orientieren. Meist ist es möglich, mit nur geringfügig mehr Netzelementen und Übertragungsleitungen eine beträchtliche Verbesserung der Dienstgüte zu erzielen. Das betrifft beispielsweise die Auslastung in einem Teilnetz oder die Transferzeit einer Kom-

munikationsbeziehung. Das Verhalten der Algorithmen in dieser Hinsicht soll im nächsten Abschnitt untersucht werden.

5.4.5 Einfluß der maximal zulässigen Transferzeit

Ein wichtiges Kriterium für die Dienstgüte in einem MAN-Vermittlungssystem ist die Transferzeit, die eine Paketübermittlung von der Ursprungsstation i bis zur Zielstation j benötigt. Zur Beurteilung der Dienstgüte wird hier die mittlere Transferzeit t_{ij}^R verwendet (vgl. Abschnitt 4.2.4). Zur Bewertung der Topologien werden zwei Größen ermittelt: Der Mittelwert von t_{ij}^R , gemittelt über alle Quelle-Ziel-Paare (i, j) , mit

$$t_{mit} = \frac{1}{N^2 - N} \sum_{i=1}^N \sum_{j=1}^N t_{ij}^R, \quad (5.8)$$

und der Maximalwert der Transferzeit mit

$$t_{max,ms}^R = \max_{ij} t_{ij}^R, \quad (5.9)$$

der als Maximum der Transferzeit aller Quelle-Ziel-Paare (i, j) gemessen wird. Zur Bestimmung dieser Größen werden für Problem B_1 100 Lösungen der Verfahren, die auf *Simulated Annealing* bzw. Genetischen Algorithmen beruhen, herangezogen und die gemessenen Größen entsprechend gemittelt.

Von Interesse ist das Verhalten der Algorithmen, wenn die maximal zulässige Pfadtransferzeit t_{max}^R verändert wird, insbesondere wenn t_{max}^R abgesenkt wird. In der Realität entspricht das schärferen Anforderungen an die Netzgüte, die beim Netzentwurf einzuhalten sind. Beginnend mit einem Wert von 4 ms wird t_{max}^R auf 1.5 ms abgesenkt. Die Zwischenwerte sind 3 ms, 2.5 ms, 2 ms und 1.8 ms. Bild 5.16 zeigt den Zusammenhang zwischen den Kosten, die bei Absenken von t_{max}^R im Mittel bei 100 Lösungen von *Simulated Annealing* erreicht werden, und der mittleren Transferzeit t_{mit} . Zum Vergleich mit dem Dekompositionsverfahren kann die Lösung für Problem B_1 (s. dazu Abschnitt 5.4.3) herangezogen werden: Das Dekompositionsverfahren erzielt Netzkosten von 239 KE, wobei $t_{mit} = 0.53$ ms und $t_{max,ms}^R = 1.2$ ms ist. Dabei wird deutlich, daß das Dekompositionsverfahren einen guten Kompromiß zwischen Netzgüte und Kostenoptimierung erreicht. Soll das Resultat bei *Simulated Annealing* vergleichbar sein, muß t_{max}^R signifikant verringert werden. Damit bestätigt sich die für das Dekompositionsverfahren gesetzte Zielvorgabe.

Wie sich die Vorgabe von t_{max}^R auf die tatsächliche gemessene maximale Transferzeit $t_{max,ms}^R$ auswirkt, zeigt Bild 5.17. Die Bilddiagonale begrenzt den zulässigen Bereich nach oben. Ein „+“ markiert in Bild 5.17 parallel zur Ordinate jeweils einen Wert von $t_{max,ms}^R$. Bei zunehmend kleineren Werten von t_{max}^R schränkt sich die Variationsbreite ein, da unzulässige

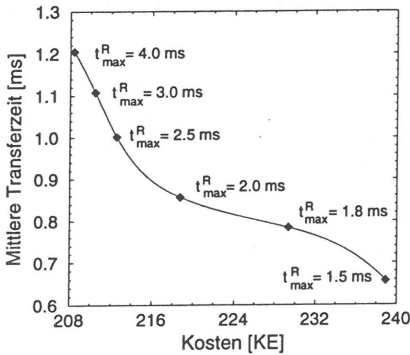


Bild 5.16: Einfluß von t_{max}^R (Simulated Annealing)

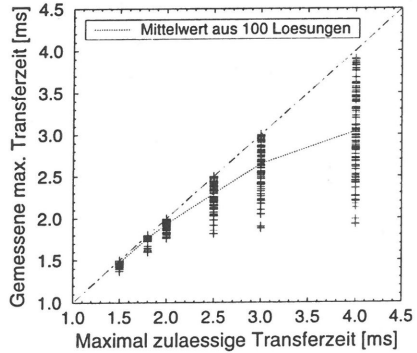


Bild 5.17: Gemessene max. Transferzeit (Simulated Annealing)

Topologien durch den Algorithmus abgelehnt werden und durch die physikalischen Laufzeiten stets ein Mindestwert von $t_{max,ms}^R$ vorgegeben ist. Umgekehrt ist es möglich, auch bei großem t_{max}^R günstige Topologien sowohl hinsichtlich der Kosten als auch der Transferzeit zu finden. Der Bereich der maximal zulässigen Transferzeit wird demnach nicht immer voll ausgeschöpft.

Die Bilder 5.18 und 5.19 zeigen die analogen Ergebnisse für die Genetischen Algorithmen. Auch hier führt, wenn auch auf höherem Niveau, das Absenken von t_{max}^R zu erhöhten Kosten. Die Verteilung von $t_{max,ms}^R$ liegt wesentlich unter dem zulässigen Bereich.

Die Kenntnis dieser Zusammenhänge kann dem Netzplaner die Auswahl verschiedener Netz-

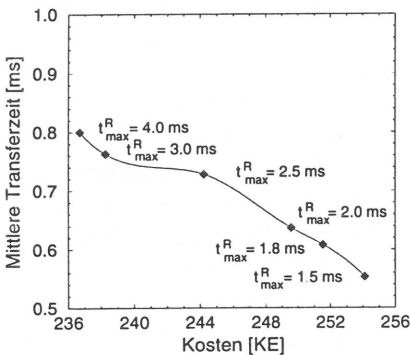


Bild 5.18: Einfluß von t_{max}^R (Gen. Algorithmen)

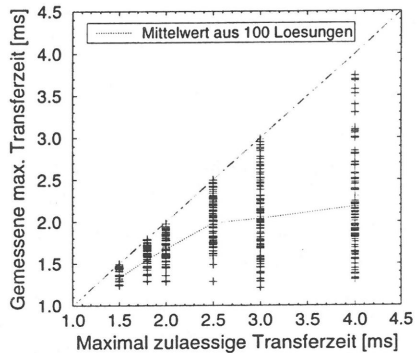


Bild 5.19: Gemessene max. Transferzeit (Gen. Algorithmen)

topologien erleichtern, da bei realen Planungsproblemen oftmals weitere, nur schwer quantifizierbare Einflußgrößen eine Rolle spielen und damit ein zusätzliches Kriterium beim Netzentwurf darstellen. Deshalb sollten mit Hilfe der Planungsverfahren stets mehrere Alternativen betrachtet werden. In diesem Zusammenhang sei erwähnt, daß das Dekompositionsverfahren auch die direkte Begrenzung der Buszugriffszeit in einem Teilnetz ermöglicht, wenn die Lastgrenzen ρ_{CI} bzw. ρ_V auf einen entsprechenden Wert gesetzt werden.

5.4.6 Abschätzung der Rechenzeiten

Ein weiteres Kriterium zur Bewertung von Planungsverfahren ist die Rechenzeit (CPU-Zeit), die für einen Algorithmus aufgewendet werden muß, um eine Lösung zu finden. Eine Abschätzung der Größenordnung für die Zahl der Schritte, die jeweils benötigt werden, ist für die statistischen Optimierungsverfahren nicht möglich. Daher werden die mittleren Programmlaufzeiten der Implementierungen verglichen.

Vergleiche der Programmlaufzeiten zwischen unterschiedlichen Verfahren sind meist stark von der Systemplattform, dem *Compiler* und nicht zuletzt der Art und Weise, wie der Algorithmus codiert ist, abhängig. Im vorliegenden Fall sind zumindest einige Voraussetzungen gegeben, die einen sinnvollen Vergleich zulassen. Alle Verfahren wurden auf leistungsfähigen Arbeitsplatzrechnern der Serie HP 9000/700 unter dem Betriebssystem HP-UX implementiert. Die problemabhängigen Programmteile, d. h. Programmteile, die die Repräsentation des MSS oder die Berechnung des Verkehrslenkungsschemas betreffen, basieren auf der gleichen Implementierung (vgl. dazu Kapitel 6). Die Programme wurden mit einem *Compiler* auf der höchsten Optimierungsstufe, die der *Compiler* erreichen kann, übersetzt.

Die Programmlaufzeiten werden in Abhängigkeit von der Zahl der Stationen gemessen. Die Beispiele, die hier verwendet werden, sind entsprechend Problem B_2 gewählt, d. h. der maximale Verkehrsfluß einer Verkehrsbeziehung (γ_{max}) beträgt 0.1 Mbit/s. Die maximale Anzahl der Schritte pro Temperaturschritt Θ_i (*Simulated Annealing*) wurde so ausgelegt, daß die Nachbarschaft hinreichend abgesucht wird. Um die Anzahl der untersuchten Topologien in eine vergleichbare Größenordnung zu bringen, wurde die Gesamtzahl der Schritte des *Simulated-Annealing*-Verfahrens genau so groß wie das Produkt aus der Zahl der Generationen und der Zahl der Individuen bei den Genetischen Algorithmen gewählt. Bei 60 Stationen beträgt die maximale Anzahl untersuchter Topologien 60 000. Die CPU-Zeit wird als Mittelwert aus 10 Ergebnisläufen berechnet.

Die Ergebnisse sind in Bild 5.20 aufgetragen. Beide Verfahren zeigen ein deutlich überproportionales Wachstum der Laufzeiten, wobei *Simulated Annealing* im Mittel deutlich weniger Rechenzeit benötigt. Der Vergleich sollte aber auch unter einem anderen Blickwinkel betrachtet werden. In erster Linie ist nicht die reine Rechenzeit interessant, sondern die in einer

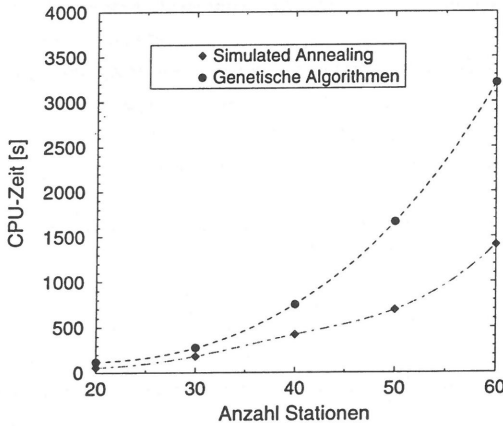


Bild 5.20: Vergleich der CPU-Zeiten (Mittelwerte)

bestimmten Rechenzeit „produzierte“ Qualität. Die korrespondierenden Werte der Zielfunktion zeigt Tabelle 5.3: Ist die Zahl der Stationen gering, sind die Genetischen Algorithmen bei vergleichbarer Rechenzeit leicht besser. Bei höheren Stationszahlen, d. h. damit auch bei steigender Last pro Teilnetz, ist das *Simulated-Annealing*-Verfahren überlegen. Zum Vergleich sei erwähnt, daß zur Bestimmung der Verteilungen der zufällig ermittelten Topologien (s. oben) jeweils mehrere Stunden CPU-Zeit benötigt werden.

Bei *Simulated Annealing* kann die Laufzeit durch das Absenken der Akzeptanzwahrscheinlichkeit P_A von 0.99 auf 0.3 weiter verringert werden. Wie die Fallstudien in den vorangegangenen Abschnitten gezeigt haben, ist damit auch eine, zum Teil beträchtliche Verbesserung der Zielfunktion möglich, da das Verfahren in Kombination mit dem CLA verwendet wird. Bei den Genetischen Algorithmen kann zwar die Zahl der Individuen und damit die Lauf-

Anzahl Stationen	Kosten [KE]	
	Simulated Annealing	Genetische Algorithmen
20	113	106
30	163	158
40	229	246
50	319	358
60	436	479

Tabelle 5.3: Kostenvergleich (Mittelwert aus 10 Lösungen)

zeit verringert werden. Dann sollten jedoch die Wahrscheinlichkeit für das *Crossover* durch den PMX-Operator und die Mutationswahrscheinlichkeit erhöht werden, um eine genügende Variationsbreite an Individuen zu garantieren. In den Experimenten hat sich jedoch herausgestellt, daß die Qualität der Lösungen dadurch eher geringer wird. Die besten Ergebnisse werden erzielt, wenn die Wahrscheinlichkeit für *Crossover* zwischen 0.5 und 0.55 eingestellt und die Mutationswahrscheinlichkeit auf 0.001 gesetzt wird.

Eine weitere Verbesserung der Lösungen kann im allgemeinen durch erhöhten Rechenaufwand erreicht werden. Werden hochqualitative Lösungen gefordert, ist es sinnvoll, bei *Simulated Annealing* den Faktor τ bis auf 0.99 und gleichzeitig die Zahl der Schritte pro Temperaturschritt Θ_i zu erhöhen. Bei den Genetischen Algorithmen ist die Erhöhung der Zahl der Individuen oder der Generationen eine vergleichbare Strategie. So wurde beispielsweise für 60 Stationen ein Zielfunktionswert von 447 KE erzielt, als die Maximalzahl der untersuchten Topologien durch die Steigerung der Zahl der Individuen und Generationen auf 125 000 erhöht wurde.

Das Dekompositionsverfahren bleibt über den ganzen Bereich unter einer Laufzeit von 4 s und ist damit in der Rechenzeit das weitaus schnellste Verfahren. Es ist sowohl zur raschen Bewertung vieler verschiedener Entwurfsalternativen als auch zur Erzeugung guter Starttopologien für die Optimierungsverfahren mit globalen Suchstrategien geeignet. Insbesondere bei der Kombination des CLA mit *Simulated Annealing* ist der Anteil des CLA an der Rechenzeit vernachlässigbar.

5.4.7 Diskussion

In den vorangegangenen Abschnitten wurden die drei Planungsverfahren anhand von charakteristischen Beispielen untersucht. Vergleicht man die Ergebnisse, so läßt sich zusammenfassend feststellen:

1. Die drei Planungsverfahren erzielen im Vergleich mit repräsentativen Stichproben zufälliger Topologien signifikant bessere Ergebnisse als das Minimum aller zufälligen Verteilungen, die in der gleichen CPU-Zeit berechnet werden können. In einigen Fällen erreichen die Planungsverfahren eine untere Schranke des Problems bis auf 19 Prozent. Prinzipiell sind demnach alle Verfahren in der Lage, für das MSS-Entwurfsproblem Lösungen von hoher Qualität zu ermitteln.
2. Das Verfahren mit *Simulated Annealing* erzeugt für Entwurfsprobleme mit gleichverteilter Verkehrsmatrix und zufälliger Verteilung der Standorte die kostengünstigsten Lösungen. Für Problem B_3 ist unter Ausnutzung des CLA eine gute Lösung zu erreichen.

3. Das Dekompositionsverfahren bzw. die Genetischen Algorithmen weichen bei den Problemen B_1 bzw. B_2 von den besten Ergebnissen, die man mit *Simulated Annealing* erhält, um etwa 10 - 20 Prozent ab. Dies liegt beim Dekompositionsverfahren an der geringeren Fähigkeit, optimale Anordnungen in einem Teilnetz zu erzeugen. Bei den Genetischen Algorithmen sind die *Building Blocks* in diesen Fällen nicht ausreichend signifikant. Beide Verfahren finden jedoch sehr gute Topologien für Problem B_3 und ähnlich strukturierte Probleme.
4. Darüber hinaus sind alle drei Verfahren auch in der Lage, Topologien zu finden, die einen Kompromiß zwischen geringen Netzkosten bei gleichzeitig ausreichender Netzgüte darstellen.
5. Die statistischen Optimierungsverfahren erfordern in der Regel einen relativ hohen Rechenaufwand. Durch die Kombination mit dem CLA bzw. durch die Erzeugung günstiger Individuen (Genetische Algorithmen) kann dieser jedoch reduziert werden. Das Dekompositionsverfahren ist außerdem zur raschen Bewertung verschiedener Entwurfsalgorithmen geeignet, was im wesentlichen durch die gleichzeitige Festlegung der Topologie und des Verkehrslenkungsschemas ermöglicht wird. Die Anwendung der einzelnen Verfahren richtet sich nach den Anforderungen an die Güte bei einer Netzplanungsaufgabe.

Eine Anwendung der Verfahren für andere Netztypen ist denkbar, insbesondere bei Netzen, die ringförmige Strukturen als wesentliches Element ihrer Architektur besitzen, wie z. B. FDDI- oder zum Teil auch SDH-Netze. Dies gilt in erster Linie für die Netzstrukturänderungen; die Berechnung der Verkehrsflüsse ist entsprechend anzupassen.

5.5 Erweiterungen

Für die entwickelten Planungsverfahren sind verschiedene Erweiterungen denkbar. Die Verfahren sind zum einen zur Untersuchung alternativer Netzstrukturen und zur Neuplanung eines MAN-Vermittlungssystems und zum anderen zur Erweiterung bereits bestehender Netze geeignet. Der letztgenannte Aspekt wird im nächsten Abschnitt untersucht.

5.5.1 Ausbau bestehender MAN-Vermittlungssysteme

Die bisher beschriebenen Verfahren und Fallstudien zielen auf die Planung eines neu zu erstellenden MAN-Vermittlungssystems ab. Während des Netzbetriebs können sich jedoch die Voraussetzungen ändern, die bei einer Neuplanung gegeben sind. So können durch steigendes Verkehrsaufkommen oder durch zusätzliche Anschlüsse von Teilnehmern Maßnahmen zur

Netzerweiterung notwendig werden. Meist wird dabei die weitgehende Erhaltung bestehender Strukturen und Konfigurationen gewünscht. Bei MAN-Vermittlungssystemen kommen daher folgende Maßnahmen zur Netzerweiterung in Frage, die in ihren Wirkungen von „lokal beschränkt“ bis „netzweit wirksam“ eingestuft werden können:

- Bestehende *Router-Router*-Kopplungen können durch leistungsfähigere *Router* in ihrer Kapazität gesteigert werden. Diese Maßnahme bleibt lokal beschränkt, sofern die Verkehrslenkung nicht geändert wird.
- Durch zusätzliche *Router-Router*-Kopplungen kann ein Teil des Transitverkehrs in Teilnetzen auf direkte Verbindungen zwischen Teilnetzen verlagert werden. In diesem Fall entstehen neue Kommunikationspfade und die Verkehrslenkung muß geändert werden.
- Mit der Einführung eines *Backbone*-Netzes, das mehrere Teilnetze verbindet und nur *Router* enthält, wird der gleiche Effekt erzielt, da *Backbone*-Netze ihre Kapazität ausschließlich für Transitverkehr bereitstellen. Auch hier muß die Verkehrslenkung geändert werden.

In den genannten Fällen gelten die Nebenbedingungen und die Zielfunktion, wie sie für das Optimierungsmodell in Abschnitt 4.3 formuliert wurden. Als zusätzliche Einschränkung sind aber ein Teil der Variablen x_{ki}^{mn} und y_{op}^m bereits mit festen Werten belegt, die die bestehende Topologie kennzeichnen.

Während die Erhöhung der Kapazität von Übertragungsabschnitten, auch in Teilnetzen, meist manuell durchgeführt werden kann, ist in den anderen Fällen die Anwendung von automatisierten Planungsverfahren erforderlich. Von den oben genannten Verfahren sind zwei dazu insbesondere geeignet: Das Dekompositionsverfahren und *Simulated Annealing*. Diese beiden Verfahren können einfach erweitert werden.

Das erste Verfahren zur Netzerweiterung basiert auf dem in Abschnitt 5.2 beschriebenen *Simulated-Annealing*-Verfahren. Die wesentliche Änderung betrifft die Starttopologie, die durch das bestehende MSS bereits vorgegeben ist. Strukturen wie z. B. Übertragungsabschnitte, die nicht geändert werden sollen, werden durch das Setzen der Entscheidungsvariablen fixiert. Danach werden fehlende Verbindungen zwischen Teilnetzen zu einer vollständigen Vermaschung ergänzt. Schließlich kann *Simulated Annealing* in reduzierter Form mit den Transformationen, die eine *Router-Router*-Kopplung zwischen Teilnetzen hinzufügen oder entfernen, die optimale Topologie ermitteln. Um die Einhaltung der zusätzlichen, durch die bestehende Topologie vorgegebenen Nebenbedingungen zu gewährleisten, werden nur Änderungen der „freien“ *Router-Router*-Kopplungen akzeptiert.

Als Beispiel sei die Erweiterung der Topologie geplant, die in Abschnitt 5.4.3 für Problem B_3 erzeugt wurde (vgl. Bild 5.13). Dabei wird angenommen, daß die Anzahl der Verkehrsbeziehungen um einen bestimmten Prozentsatz wächst und die Verkehrsintensität pro

Steigerungsrate Verkehr (in Prozent)	Erhöhung der Kosten (in Prozent)	Erhöhung der mittleren Transferzeit (in Prozent)	Erhöhung der maximalen Transferzeit (in Prozent)
20	0	40.3	18.9
25	0	71.7	35.3
30	3.2	75.9	55.7
35	6.3	100	56.7
40	66.8	98.3	73.2

Tabelle 5.4: Sensitivität der Netzkosten bei Anstieg des Verkehrsvolumens

Verkehrsbeziehung im gleichen Maße zunimmt. Da Verkehrsschätzungen mit einer gewissen Ungenauigkeit behaftet sein können, wird die Steigerungsrate des Verkehrsvolumens variiert. Dieses Vorgehen empfiehlt sich, um die Sensitivität der Kenngrößen einer Topologie zu prüfen. Die maximal zulässige Transferzeit t_{max}^R beträgt 2 ms. Die Kenngrößen der ursprünglichen Topologie lauten: Die Kosten betragen 190 KE, die mittlere Transferzeit t_{mt} 0.48 ms und die gemessene maximale Transferzeit $t_{max,ms}^R$ 1.11 ms.

Die Ergebnisse zeigt Tabelle 5.4. Die Kosten nehmen bei 20 bzw. 25 Prozent erhöhtem Verkehrsvolumen nicht zu, d. h. die Topologie ist unverändert. Das erhöhte Verkehrsvolumen schlägt sich aber in den steigenden Transferzeiten nieder. Auffallend ist das Verhalten bei einer Steigerungsrate von 40 Prozent, bei der die Kosten sprunghaft zunehmen. In diesem Fall sind die Teilnetze fast vollständig vermascht. Das bedeutet, daß das Netz die Gren-

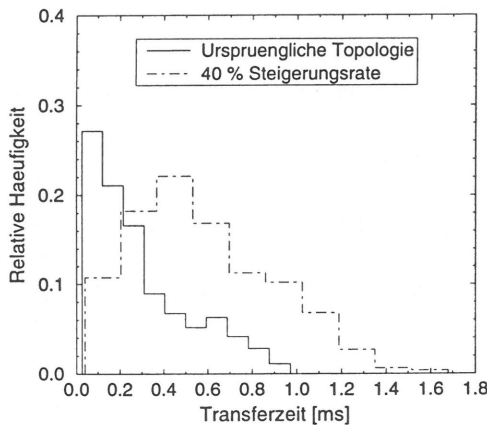


Bild 5.21: Histogramm der Transferzeiten t_{ij}^R

zen seiner Gesamtkapazität erreicht hat und weitere Steigerungen des Verkehrsvolumens nur durch das Aufteilen von Teilnetzen zu bewältigen sind. Das bestätigt auch der Vergleich der Transferzeitverteilungen über alle Kommunikationsbeziehungen (Bild 5.21). Gegenüber der ursprünglichen Topologie erhöhen sich die Transferzeiten auf beinahe allen Kommunikationsbeziehungen. Wie das Beispiel zeigt, kann mit einer Sensitivitätsanalyse der Spielraum für eine bestimmte Erweiterungsstrategie einfach geprüft werden.

Eine alternative Möglichkeit zur Erweiterungsplanung ergibt sich mit dem Dekompositionsverfahren, wenn ausschließlich das Verfahren zur Verbindung der Teilnetze verwendet wird. Dieses Verfahren setzt auf einer gegebenen Zuordnung der Stationen zu Teilnetzen auf. Es wird so lange ausgeführt, bis eine Netzstruktur gefunden ist, die die bereits bestehende vollständig enthält. Diese Vorgehensweise eignet sich besonders dann, wenn die Zahl der Teilnetze nicht sehr groß ist. Darüber hinaus kann die Infrastrukturmatrix genutzt werden, um die Benutzung bereits bestehender Übertragungsabschnitte zu unterbinden.

Die Einführung eines *Backbone*-Netzes ist eine weitere Alternative zur Netzgestaltung, insbesondere dann, wenn die Vernetzung mit direkten *Router-Router*-Kopplungen durch den hohen Vermaschungsgrad nicht mehr kostengünstig ist. *Backbone*-Netze werden im folgenden als Ebene-2-Netze bezeichnet, um sie von Ebene-1-Netzen, die auch Stationen enthalten, zu unterscheiden. Ebene-2-Netze werden aus der Sicht der Planungsverfahren wie andere Teilnetze behandelt. Geeignet ist dafür das *Simulated-Annealing*-Verfahren. Als Vorgabe wird dabei aber die Anzahl der Ebene-2-Netze benötigt. Von den Transformationen, die das lokale Suchverfahren bereitstellt, werden wiederum nur diejenigen benötigt, die die Verbindungsstruktur der Teilnetze oder die Reihenfolge der Netzelemente in den Teilnetzen ändern. Die Mindestzahl von Ebene-1-Netzen, die an ein Ebene-2-Netz angeschlossen werden sollte, beträgt drei.

Falls mehrere *Backbone*-Netze die Ebene-1-Netze verbinden sollen, ergibt sich eine weitere Nebenbedingung für eine zulässige Topologie (s. Bild 5.22). Das Verkehrslenkungsschema ist dann so zu wählen, daß bei mehreren physikalischen Pfaden, die das Ebene-2-Netz einschließen, der Verkehr direkt über das Ebene-2-Netz geführt wird (hierarchische Verkehrslenkung). Nur bei einer geringeren Anzahl von Netzkoppeleinheiten wird der Verkehr direkt in ein Ebene-1-Netz geführt. Die Einhaltung dieser Nebenbedingung sollte jedoch erst nach dem Entwurf der physikalischen Topologie geprüft werden, da dies während der Optimierung zur weiteren Erhöhung der Rechenzeiten beiträgt und meist nur wenige Pfade betroffen sind.

5.5.2 Erhöhung der Verfügbarkeit

Ein wichtiges Kriterium beim Netzentwurf ist die Netzverfügbarkeit. Im Zusammenhang mit dem MSS-Entwurfproblem sollen nur Planungsmaßnahmen betrachtet werden, die dazu beitragen können, eine ausreichende Funktionsfähigkeit des MSS zu gewährleisten. Dazu sind

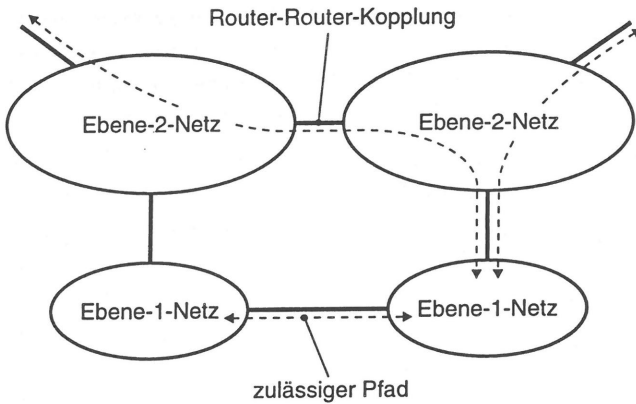


Bild 5.22: Hierarchische Verkehrslenkung bei *Backbone*-Strukturen

zwei Maßzahlen von besonderem Interesse: Die Zahl der knotendisjunkten Pfade zwischen je zwei Netzknuten und die freie Kapazität, die im Fehlerfall bereitsteht, um den Verkehrsfluß zusätzlich zu übermitteln, der z. B. durch den Ausfall eines Übertragungsabschnittes umgelenkt werden muß. Zur Bestimmung von weiteren Zuverlässigkeitskenngrößen einer Rechnernetztopologie sei auf die Literatur verwiesen [67].

Zunächst werden Maßnahmen zur Erhöhung der Zahl der knotendisjunkten Pfade betrachtet. Durch die *Looped-Bus*-Struktur der Teilnetze ist bereits eine Absicherung in den Teilnetzen für einen einmaligen Ausfall eines Übertragungsabschnittes vorhanden. Somit bestehen zwischen allen Stationen in einem Teilnetz zwei knotendisjunkte Pfade. Als einfachste Möglichkeit, zwei knotendisjunkte Pfade im gesamten MSS zu garantieren, bietet sich die Erweiterung des Dekompositionsverfahrens an. Dabei liegt der Ansatzpunkt bei der Ermittlung des Grundgerüsts zur Verbindung der Teilnetze (s. Abschnitt 5.1.2). Zusätzlich zu der Erzeugung des *Spanning Tree* wird aus den Knoten, die nur durch eine Kante an den Baum angebunden sind, ein zweiter *Spanning Tree* konstruiert, der einen *Minimum Spanning Tree* bezüglich der Entfernungen bildet. Diese Konstruktion führt stets auf eine Topologie mit zwei knotendisjunkten Pfaden zwischen je zwei Knoten [118]. Der übrige Teil des Algorithmus zur Verbindung der Teilnetze ist dann aber nicht mehr verwendbar. Somit muß die Einhaltung der Kapazitäts- bzw. Auslastungsgrenzen nachträglich geprüft werden.

Sind die Anforderungen geringer, kann auch eine Struktur mit Ebene-2-Netzen verwendet werden. In diesem Fall werden die Ebene-1-Netze an ein oder mehrere Ebene-2-Netze angeschlossen. Das MSS ist nur dann getrennt, wenn die *Router* zu den Ebene-2-Netzen ausfallen. Eine Verdoppelung der *Router* an den Koppelstellen kann das Risiko weiter verringern.

Maßnahmen zur Verbesserung der Funktionsfähigkeit, die die bestehenden Mechanismen der

Planungsverfahren benutzen, sind weniger geeignet. Durch sehr niedrige Werte für t_{max}^R kann auch gleichzeitig eine hohe Verfügbarkeit erreicht werden. Die Topologien sind dann meist fast vollständig vermascht. Sinnvoll ist es jedoch, die freie Kapazität durch Begrenzung der Auslastung zu erhöhen, so daß diese im Fehlerfall für umzulenkenden Verkehr bereit steht.

Wenn die Zahl der Teilnetze in einem MSS nicht sehr hoch ist, können auch einfache Näherungslösungen ausreichen, bei denen ein, auf die Netzkosten optimiertes MSS mittels Erfahrungsregeln so ergänzt wird, daß eine ausreichende Verfügbarkeit erreicht wird.

Kapitel 6

Entwurf und Realisierung eines Planungswerkzeugs

Verfahren zur rechnergestützten Netzplanung und -dimensionierung sind von erheblicher Bedeutung für einen Netzplaner, um in systematischer Weise eine kostenoptimale Netzstruktur, die die Verkehrs- und Dienstgüteanforderungen erfüllt, zu entwerfen und, darauf aufbauend, einzelne Netz- oder auch Dienstelemente geeignet zu dimensionieren. Dieses Kapitel beschreibt die Architektur und einige Implementierungsaspekte eines Softwarewerkzeugs zur Planung von MAN-Vermittlungssystemen. Darüber hinaus wird die Anwendung des Planungswerkzeugs für einen typischen Entwurfsprozeß gezeigt.

6.1 Anforderungen

Werkzeuge zur Netzplanung und -dimensionierung gewinnen mit der fortschreitenden Entwicklung auf dem Gebiet der Kommunikationstechnik eine stetig wachsende Bedeutung für einen effizient gestalteten Netzplanungsprozeß [9, 19, 21, 22, 87, 125]. Trotzdem scheinen sich Werkzeuge, die den Planungsprozeß in verschiedenen Phasen unterstützen, erst allmählich durchzusetzen [21]. Daher sollte ein Netzplanungswerkzeug ein klares Anforderungsprofil erfüllen, das sich aus den zu lösenden Planungsproblemen ergibt. Dazu sind einige prinzipielle Anforderungen zu klären:

- Für welche Aufgaben soll das Netzplanungswerkzeug eingesetzt werden?

Für ein Netzplanungswerkzeug sind verschiedene Zielsetzungen denkbar. So kann es zum einen nur für theoretische Voruntersuchungen zur Bewertung einer künftigen Netzstruktur eingesetzt werden. Dann genügt ein Softwarewerkzeug, das mit effizienten Algorithmen auf unterschiedlichen Genauigkeitsstufen Netzstrukturen entwerfen und

bewerten kann. Andererseits kann ein Einsatz in einem integrierten Netzplanungs- und Netzmanagementwerkzeug erwünscht sein. Dann spielen auch Fragen der Datenorganisation eine wichtige Rolle. Dazu zählt z. B. die Frage, ob eine Netzdatenbank zur Beschreibung und Bestandsführung von Netz- oder Dienstelementen vorgesehen ist, oder wie Daten über standardisierte Schnittstellen des Netzmanagements gewonnen werden. Da bei großen Netzen ein überaus umfangreiches Datenvolumen anfallen kann – die Tarifierungsdaten von Teilnehmern bieten ein gutes Beispiel –, sind diese Schnittstellen sorgfältig zu entwerfen. In jedem Fall muß geklärt werden, welche Daten, wie z. B. Verkehrsvolumen, Standortdaten oder Tarifstrukturen, überhaupt notwendig sind.

- Welche Methoden der Netzplanung sind notwendig?

Netzplanungsverfahren lassen sich in Analyse- und Entwurfsverfahren aufteilen. Analyseverfahren umfassen beispielsweise verkehrstheoretische Analysen oder die Bewertung von Netzstrukturen hinsichtlich der Kosten. Zu den Entwurfsverfahren zählen die Optimierungsalgorithmen. Ein Netzplanungswerkzeug sollte beides enthalten. Die Funktionalität wird sinnvollerweise in Form mehrerer Grundbausteine erbracht, die miteinander kombiniert werden können, um komplexere Planungsverfahren zu ermöglichen. Der für eine bestimmte Netzart spezifische Teil sollte eher gering ausfallen, um allgemeine und damit erprobte Verfahren gewinnbringend einzusetzen. Ferner sind Verfahren mit unterschiedlichem Detaillierungsgrad erforderlich, um einerseits eine mehr makroskopische Sicht auf die Netzstruktur zu ermöglichen, aber andererseits auch die verfeinerte Planung und Dimensionierung der Netzelemente nach der Festlegung der prinzipiellen Netztopologie zu unterstützen.

- Wie wird die Interaktion mit dem Planungswerkzeug unterstützt?

Mit der zunehmenden Leistungsfähigkeit moderner Arbeitsplatzrechner wird eine immer detailliertere und umfassendere Modellierung auch umfangreicher Planungsprobleme möglich. Dennoch sind viele Nebenbedingungen, die im realen Netzbetrieb auftreten, nur umständlich zu modellieren, so daß sie rechnergestützt nicht mehr effizient zu bearbeiten sind. Ein erfahrener Netzplaner wird dann die angebotenen Lösungen als Entscheidungshilfe benutzen. Er sollte daher die Möglichkeit besitzen, in den Planungsablauf an definierten Punkten einzugreifen. Hierzu ist eine grafische Darstellung für den Anwender überaus nützlich; Ergebnisdaten sollten aber auch in Textform vorliegen, um sie dem Netzplaner zur weiteren Verarbeitung zugänglich zu machen.

- Wie wird das Planungswerkzeug realisiert?

Ein weiterer Aspekt ist die softwaretechnische Realisierung des Netzplanungswerkzeugs. Neben den typischen Problemen einer Softwarearchitektur, die vor allem die Flexibilität und die Erweiterbarkeit betreffen, gibt es auch Gesichtspunkte, die spezifisch für ein Netzplanungswerkzeug sind. Ein Kernpunkt ist die Gestaltung von

Datenstrukturen. Hier gilt, daß die Netztopologie soweit wie möglich auf allgemeine Datenstrukturen und Objekte, wie etwa Matrizen, Ring- oder Baumstrukturen, abgebildet werden sollte, die der Behandlung durch effiziente Algorithmen besonders zugänglich sind.

Im Rahmen dieser Arbeit wurde ein Netzplanungswerkzeug für MAN-Vermittlungssysteme entworfen und implementiert, das auf den in Kapitel 5 entwickelten Netzplanungsverfahren beruht. Die dabei entwickelten Architekturkonzepte sind allgemein anwendbar und können als Basis für künftige Netzplanungswerkzeuge dienen.

6.2 Funktionalität und Architektur

Die Funktionalität des Planungswerkzeugs zeigt Bild 6.1 in einer schematischen Übersicht. Das Planungswerkzeug umfaßt neben Funktionen zum Einlesen der Eingabedaten die Planungsverfahren, die dem Entwurf einer Netzstruktur dienen, sowie Funktionen zur Ergebnisaufbereitung und zur Ausgabe der Ergebnisse. Eingabedaten sind die Daten der Netzinfrastruktur einschließlich der potentiellen Standorte für die Netzelemente, die Verkehrsmatrix, die Beschreibung der Netzelemente sowie gegebenenfalls die vollständige Netzstruktur eines MAN-Vermittlungssystems. Sie werden mit Hilfe einer Netzbeschreibungssprache beschrieben und durch einen *Parser* eingelesen. Diese Daten werden zusammen mit den Optimierungsparametern während des Planungsprozesses durch interne Datenstrukturen repräsentiert.

Der Kern des Netzplanungswerkzeugs umfaßt unterschiedliche Planungsverfahren, mit denen je nach Aufgabenstellung eine Netzstruktur neu entworfen oder eine bereits bestehende Netzstruktur erweitert werden kann. Die Planungsverfahren, die in Kapitel 5 vorgestellt wurden, sowie weitere Analysemodule unterstützen den Netzplaner beim Entwurf.

Als wichtigstes Ergebnis erhält der Netzplaner die entworfene Netztopologie. Sie wird durch die Netzelemente und ihre logischen Zuordnungen zu Teilnetzen bzw. durch die Verbindungsstruktur zwischen Teilnetzen repräsentiert. Die Netzstruktur wird auch in graphischer Form ausgegeben. Mit Hilfe der Netzanalyse können Daten über die Netzkosten und die Dienstgüteparameter ermittelt werden. Ferner wird auch eine schematische Darstellung der Verkehrslenkungstabellen unterstützt, um für den Netzplaner, der verschiedene Verkehrslenkungsverfahren untersuchen möchte, den Verlauf der Verkehrsflüsse in einem MAN-Vermittlungssystem darzustellen.

Um die Bedienung des Netzplanungswerkzeugs für den Benutzer komfortabel zu gestalten, gibt es überdies eine Benutzeroberfläche, die der Auswahl der verschiedenen Planungsfunktionen dient.

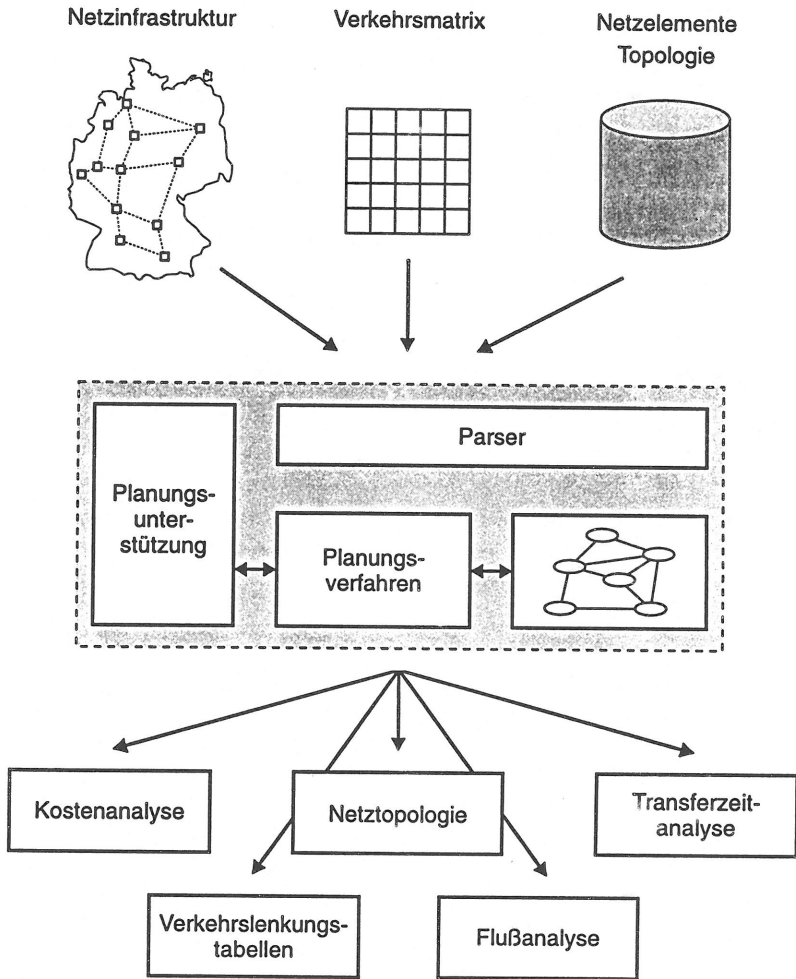


Bild 6.1: Funktionalität des Netzplanungswerkzeugs

Die genannte Funktionalität wird durch verschiedene Architekturkonzepte bereitgestellt, die im folgenden näher erläutert werden sollen.

6.2.1 Hierarchiebildung

Die Struktur des Planungswerkzeugs orientiert sich nur teilweise an den hierarchisch unterteilten Netzebenen des MAN-Vermittlungssystems. Im Gegensatz zur Aufteilung in Ebene-1-

und Ebene-2-Netze (*Backbone*-Netze) ist für die Organisation der Datenstrukturen und die Effizienz der Planungsverfahren eine Aufteilung in Zugangsnetze, Teilnetze und die Verbindungsstruktur der Teilnetze günstiger. Um das Verkehrslenkungsschema zu bestimmen, ist beispielsweise die genaue Anordnung der Netzelemente in den Teilnetzen nicht relevant. Im folgenden wird daher von Planungsebenen gesprochen. Die Funktionalität jeder Planungsebene wird wiederum durch mehrere Softwaremodule erbracht.

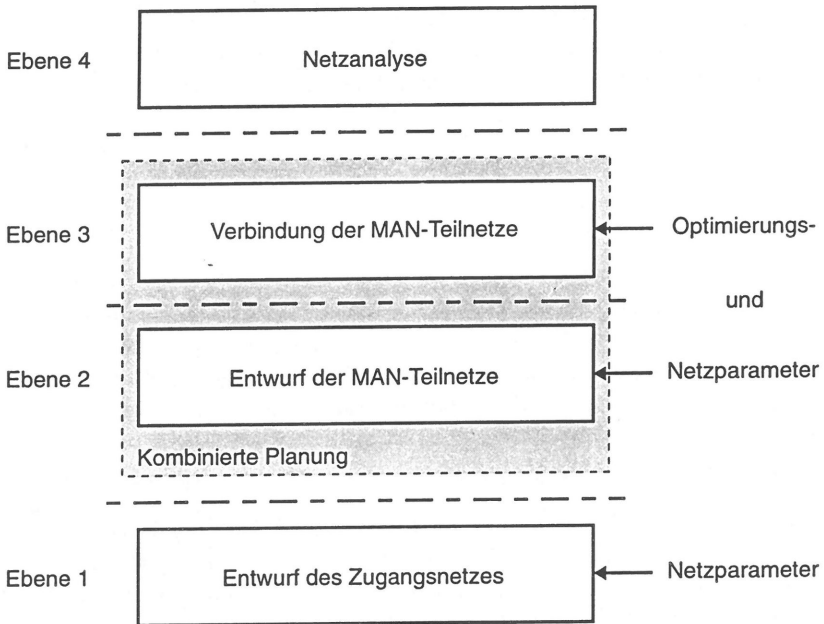


Bild 6.2: Hierarchische Planung

Bild 6.2 verdeutlicht, wie die Ebenen logisch unterteilt werden. Die Planung von Netzen bei den Teilnehmern selbst, z. B. die Planung lokaler Datennetze, wird nicht unterstützt. Somit beschreibt die erste Planungsebene die Struktur des Zugangsnetzes (*Customer Access Network*), an die die Teilnehmer angeschlossen sind. Wie bereits in Kapitel 2 erwähnt, ist diese Struktur entweder schon bekannt, oder sie kann interaktiv entworfen werden. Wahlweise kann mit Hilfe des Planungswerkzeugs eine automatische Zuordnung der *Customer Gateways* zu den potentiellen Standorten der *Edge Gateways* erfolgen, wobei die *Customer Gateways* auf der Basis der kürzesten Entfernung zugeordnet werden. Ihre Reihenfolge in einem Dualbussystem muß der Netzplaner bestimmen. Die Festlegung des Zugangsnetzes ist völlig unabhängig vom eigentlichen Transportnetz. Die nächsten beiden Ebenen können

entweder schrittweise mit dem Dekompositionsverfahren oder kombiniert geplant werden. Eine kombinierte Planung ist mit den Verfahren des *Simulated Annealing* bzw. der Genetischen Algorithmen möglich. Auf jeder dieser Ebenen ist auch der Eingriff des Netzplaners nach Abschluß eines Planungsprozesses zulässig, beispielsweise um neue Netzelemente einzufügen oder ihre Eigenschaften (Netzparameter) zu ändern. Die oberste Ebene umfaßt die Netzanalyseverfahren.

Die hierarchisch gegliederte Softwarestruktur bietet den Vorteil, daß auf den verschiedenen Planungsebenen gleichartige Datenstrukturen und Zugriffsfunktionen verwendet werden können, was den Softwareentwurf stark vereinfacht. Die Abbildung der Netzstruktur auf eine hierarchische Softwarestruktur führt deshalb zu einer sinnvollen Modularisierung der Software.

6.2.2 Netzdatenbasis

Um eine flexible und effektive Eingabe der Netzdaten zu ermöglichen, ist eine Netzbeschreibungssprache notwendig, die, wie bereits erwähnt, mit Hilfe eines *Parsers* in Datenstrukturen umgesetzt wird. Der *Parser* erkennt die benötigten Eingabedaten anhand von Schlüsselwörtern. Die Beschreibungssprache erlaubt beliebige benutzerdefinierte Typen, die ein bestimmtes Netzelement oder einen Standort kennzeichnen, und ist dadurch einfach erweiterbar. Durch die Verwendung eines *Parsers* können die Eingabedaten außerdem auf ihre Konsistenz geprüft werden. Die Eingabedaten lassen sich folgendermaßen klassifizieren:

- **Netzinfrastruktur:** Diese Daten umfassen zum einen die Lage des Teilnehmeranschlußpunktes an das Zugangnetz und zum anderen die potentiellen Standorte für die Netzelemente. Ferner zählt dazu der Trassennetzgraph, der in Form einer Matrix repräsentiert wird. Fehlt eine Trasse zwischen zwei Standorten, wird die zugehörige Entfernung in der Entfernungsmatrix auf einen sehr großen Wert gesetzt.
- **Verkehrsmatrizen:** Daten, die aus Verkehrsmessungen gewonnen werden, können relativ umfangreich sein. Daher ist die Definition der vollständigen Verkehrsmatrix durch die Festlegung beliebiger Teilmatrizen möglich. Mit diesem Konzept können auch geschlossene Benutzergruppen (s. Abschnitt 2.2.3) und die innerhalb dieser Benutzergruppen auftretenden Verkehrsfüße definiert werden. Da ein MAN auch durchschaltvermittelten Verkehr (CS-Verkehr) übermitteln kann, wird dieser mit einer konstanten Datenrate in der Verkehrsmatrix und bei der Flußanalyse berücksichtigt. Diese Annahme ist gerechtfertigt, weil CS-Verkehr über semipermanente Verbindungen übermittelt wird, die durch das Netzmanagement eingerichtet werden (vgl. dazu [69]).
- **Eigenschaften der Netzelemente:** Diese Daten umfassen sämtliche Eigenschaften der Übertragungsleitungen, der *Edge Gateways* und der *Router*, soweit sie für die Planung

benötigt werden. Das schließt beispielsweise Tarifstrukturen, Parameter der Verkehrsmodelle oder den maximalen Durchsatz mit ein. Tarifstrukturen sind stets mit einer Übertragungsleitung verknüpft.

- Topologie des MSS: Um die Netzanalyse und die Erweiterungsplanung zu unterstützen, ist auch die Eingabe bereits vorhandener Netzstrukturen durch das *Parser*-Konzept möglich.
- Optimierungsparameter: Die für die Planungsverfahren benötigten Optimierungsparameter werden separat eingelesen.

Die Netzdatenbasis ist in Form von Dateien realisiert, so daß die Planungsdaten auch außerhalb des Werkzeugs weiterverwendet werden können. Die Planungsdaten selbst sind in mehreren Teildatenbasen organisiert. Die Eigenschaften der Netzelemente bilden die erste Teildatenbasis. Sie wird typischerweise dann geändert, wenn neue Netzelemente hinzugefügt werden oder sich die Tarifierungsrichtlinien ändern. Die zweite Teildatenbasis umfaßt die Standortkoordinaten, die Verkehrsmatrix und die Matrix, die die Infrastrukturbeschränkungen beschreibt. Diese Daten sind für jedes spezifische Planungsproblem anzupassen. Das trifft auch auf die Optimierungsparameter zu, die die dritte Teildatenbasis bilden. Die Topologiedaten bilden weitere Teildatenbasen, die zu den oben genannten Planungsebenen korrespondieren. Die internen Datenstrukturen, die sowohl die Ausgangsdaten als auch die entworfene Netzstruktur repräsentieren, sind in ähnlicher Weise aufgebaut.

6.2.3 Planungsverfahren und Netzanalyse

Die in Kapitel 5 beschriebenen Verfahren dienen dem Netzplaner zum Entwurf neuer bzw. zur Änderung und Erweiterung bestehender MAN-Vermittlungssysteme. Neben der Gliederung in Planungsebenen, die als horizontale Aufteilung angesehen werden kann und die sich vor allem am Datenfluß orientiert, weist das Netzplanungswerkzeug auch eine vertikale Gliederung auf. Diese orientiert sich an den zur Verfügung stehenden Planungs- und Analyseverfahren. Einzelne Verfahren können getrennt realisiert werden, was auch die Erweiterbarkeit der Software unterstützt. Die Verfahren setzen auf einer Planungsebene auf und schließen zur nächsten Ebene hin ab. Parallel können so das Dekompositionsverfahren und die Verfahren des *Simulated Annealing* bzw. der Genetischen Algorithmen verwendet werden. Auf der obersten Planungsebene sind z. B. unterschiedliche Verkehrslenkungsverfahren einsetzbar.

Alle Planungsverfahren arbeiten mit der gleichen Datenstruktur, die die Topologie und die Eigenschaften der Netzelemente eines MAN-Vermittlungssystems beschreibt. Die Genetischen Algorithmen benötigen dazu eine Zwischenschicht, die durch spezielle Abbildungsfunktionen realisiert wird. Die Abbildungsfunktionen stellen zwischen der Codierung in

Zeichenketten und der Netztopologie eine Verknüpfung her. Das Dekompositionsverfahren und das Verfahren des *Simulated Annealing* können die Topologie direkt verändern.

Neben den Planungsverfahren wird die Netzanalyse bestehender Netze unterstützt. Wie schon erwähnt, ist damit auch die Erstellung von Verkehrslenkungstabellen verbunden. Das *Minimum-Hop*-Verfahren kann zusätzlich erweitert werden, um durch den Netzplaner vorgegebene Gewichte für *Router-Router*-Kopplungen bei der Berechnung der *Destination Spanning Trees* mit einzubeziehen. Ferner kann eine bestehende Topologie durch das Entfernen von Übertragungsabschnitten auf eventuelle Konsequenzen von Leitungsausfällen überprüft werden.

6.2.4 Benutzerschnittstelle

Die Interaktion mit dem Netzplaner wird durch eine menü-gesteuerte Benutzerschnittstelle ermöglicht. Ein wesentlicher Teil dieser Schnittstelle ist die Netzbeschreibungssprache, die bereits in Abschnitt 6.2.2 erwähnt wurde. Diese kann mit Hilfe eines Editors eingegeben werden.

Die Ergebnisdaten umfassen

- die Daten zur Netzstruktur einschließlich der gewählten Netzelemente und ihrer logischen Verknüpfungen,
- die Verkehrslenkungstabellen der *Router*,
- eine Aufschlüsselung der Kosten nach Netzelementen,
- die Daten zur Auslastung der Doppelbussysteme, der EGWs und der *Router* sowie
- die Verteilung der Transferzeiten auf den einzelnen Pfaden.

Diese Daten sind zur weiteren Verarbeitung durch Auswertungssoftware, wie z. B. Statistiksoftware, geeignet. Darüber hinaus wird die graphische Darstellung der Netzstruktur auf verschiedenen Hierarchieebenen unterstützt. Die Ausgabedaten sind auf jeder Hierarchiestufe des Planungswerkzeugs so organisiert, daß sie bei Erweiterungen der Netzstruktur oder bei Netzanalysen unabhängig bearbeitet werden können. Die ermittelten Daten zur Netzstruktur können stets erneut als Eingabedaten genutzt werden.

6.3 Implementierungsaspekte

Die Umsetzung der Konzepte in das Softwarewerkzeug erfordert auch die Beachtung von Implementierungsaspekten. Um die Effizienz der Algorithmen in der Implementierung zu

sichern, wurde auf die Gestaltung der Datenstrukturen besonderes Gewicht gelegt (s. auch [155]). Zwei Anforderungen stehen dabei im Vordergrund. Zum einen sollten die Datenstrukturen möglichst einfach aufgebaut sein, um Änderungen bei einer Erweiterung des Planungswerkzeugs zu ermöglichen. Zum anderen sollte ein schneller Zugriff auf die Daten während der Optimierungsphase möglich sein. Deshalb muß die Verwendung dynamischer Datenstrukturen, wie z. B. verkettete Listen, sorgfältig geprüft werden, da Suchoperationen in dynamischen Datenstrukturen unter Umständen die Rechenzeit stark erhöhen können. Andererseits sind rein statische Datenstrukturen ab einer gewissen Größe aus Speicherplatzgründen ungeeignet. Daher ist ein Kompromiß zwischen den beiden Positionen anzustreben.

Ein typisches Beispiel sind Matrizen und ihre Zugriffsfunktionen. Matrizen sind deshalb so wichtig, weil sie auf allen Netzebenen für den Verkehrsfluß, die Beschreibung von Distanzen und auch für das Verkehrslenkungsschema benötigt werden. Je nach Planungsproblem können sparsam oder dicht besetzte Matrizen auftreten. In der Implementierung ist das durch die Bereitstellung unterschiedlicher Matrixtypen berücksichtigt. Für dicht besetzte Matrizen wird ein Vektor pro Matrixzeile verwendet, für sparsam besetzte Matrizen hingegen verkettete Listen.

Für die Verkehrsmatrizen, die die Verkehrsbeziehungen zwischen CGWs bzw. EGWs beschreiben, werden durchgängig verkettete Listen verwendet, so daß der Speicherplatzbedarf nur linear mit der Zahl der Verkehrsbeziehungen steigt. Dadurch kann aber auch der Rechenzeitbedarf verringert werden, weil die häufig benötigte Verkehrsflußberechnung das wiederholte Aufsummieren der Verkehrswerte nur auf tatsächlich vorhandene Verkehrsbeziehungen beschränkt. Darüber hinaus wurden die Berechnungsalgorithmen so gestaltet, daß ein wahlfreier Zugriff auf Elemente der Verkehrsmatrix, der bei verketteten Listen nicht möglich ist, nicht benötigt wird.

Auch bei den hierarchisch angeordneten Netzelementen (CGWs, EGWs, Teilnetze und *Router*) wurde auf eine möglichst redundanzsparende Repräsentation geachtet. Hier wurden im wesentlichen verkettete Listen verwendet, um gleichartige Elemente in einer Liste zusammenzufassen. Eine Ausnahme ist die Verbindungsstruktur mit *Router-Router*-Kopplungen, die am effektivsten durch eine Matrix dargestellt wird. Das Vorhandensein eines Elementes dieser Matrix bedeutet, daß auch die korrespondierende *Router-Router*-Kopplung existiert. Referenzen, die zum Auffinden einzelner Elemente der Datenstruktur benötigt werden, sind im wesentlichen so gestaltet, daß Verweise von hierarchisch höheren Planungsebenen auf hierarchisch tiefer liegende Planungsebenen erfolgen. Ferner wird die Entkopplung der Planungsebenen vor allem dadurch erreicht, daß die Verkehrsmatrizen auf jeder Ebene separat berechnet werden. Der Rechenaufwand ist gegenüber dem Gewinn an Modularität vernachlässigbar.

Ein weiterer Aspekt, der besondere Aufmerksamkeit verdient, ist der Algorithmus zur Be-

rechnung des Verkehrslenkungsschemas. Dieser Algorithmus wird bei den statistischen Suchverfahren wiederholt benötigt und sollte daher effizient implementiert werden. Die Implementierung ist in [163] beschrieben. Das Hauptmerkmal ist die temporäre Umsetzung der Daten, die für den Verkehrslenkungsalgorithmus benötigt werden, in eine nach aufsteigender Verbindungskennung geordneten Liste. Auf diese Liste kann durch spezielle Funktionen schnell zugegriffen werden. Eine Erweiterung des Verkehrslenkungsmoduls, um parallele Router-Router-Kopplungen zu berücksichtigen, ist möglich.

Das Planungswerkzeug ist in der Programmiersprache ANSI-C implementiert. Es wurde auf zwei Hardwareplattformen mit unterschiedlichen Varianten des Betriebssystems UNIX entwickelt (HP 9000/700 unter HP-UX, PC unter Linux). Ferner konnte die Portabilität des Planungswerkzeugs auf eine weitere Hardwareplattform nachgewiesen werden (SUN SPARC).

6.4 Entwurfsablauf

Abschließend soll der Ablauf des Planungsprozesses mit Hilfe des Planungswerkzeugs schematisch verdeutlicht werden.

Der Netzentwurf (vgl. Bild 6.3) beginnt damit, daß die Anforderungen der Teilnehmer mit Hilfe einer Verkehrsmatrix spezifiziert werden. So können aus den Anwendungen der Teilnehmer, wie z. B. verteilte Datenbanken, die Verkehrsflüsse abgeschätzt oder durch Messungen überprüft werden. Ferner sind die Standorte (und Koordinaten) der Teilnehmer, d. h. der *Customer Gateways* (CGWs), und der Netzelemente des MSS zu bestimmen. Parallel dazu müssen die Eigenschaften der Netzelemente und die zulässigen Verknüpfungen zwischen ihnen festgelegt werden. So müssen beispielsweise spezielle Router-Typen definiert werden, wenn die Verbindung eines Teilnetzes mit einer Datenrate von 34 Mbit/s mit einem Teilnetz mit einer Datenrate von 140 Mbit/s erforderlich ist, da bei einer Verknüpfung zweier Netzelemente auch Anschlußbedingungen, d. h. die Verwendung des gleichen Leitungstyps, geprüft werden müssen.

Sind die Standorte für das MSS bekannt, kann nach der Eingabe der Infrastrukturmatrix, die aus unterschiedlichen Ursachen beschränkt sein kann, die Distanzmatrix berechnet werden. Darüber hinaus muß die Reihenfolge der CGWs im Dualbusssystem des Zugangnetzes (CAN) bestimmt werden. Jedes Zugangnetz wird an ein *Edge Gateway* angeschlossen. Sind diese Schritte ausgeführt, können die Planungsverfahren angewendet werden (s. Kapitel 5). Der Netzplaner prüft die Ergebnisse und entscheidet, ob Änderungen der Optimierungsparameter notwendig sind, oder etwa ein anderes Verfahren gewählt werden soll. Ist das nicht der Fall, ist der Planungsprozeß abgeschlossen.

Nummehr müssen die Planungsergebnisse in die physikalischen Komponenten des MSS umgesetzt werden. Das beinhaltet die Ermittlung der notwendigen Hard- und Softwarekom-

ponenten der Netzelemente, der Übertragungsleitungen und eventueller implementierungsspezifischer Parameter. Ferner muß der Adressierungsplan, der ein wesentlicher Bestandteil der Konfigurationsparameter ist, bestimmt werden. Prinzipiell kann nun das MSS installiert und in Betrieb genommen werden. Diese Schritte sind in Bild 6.3 lediglich angedeutet.

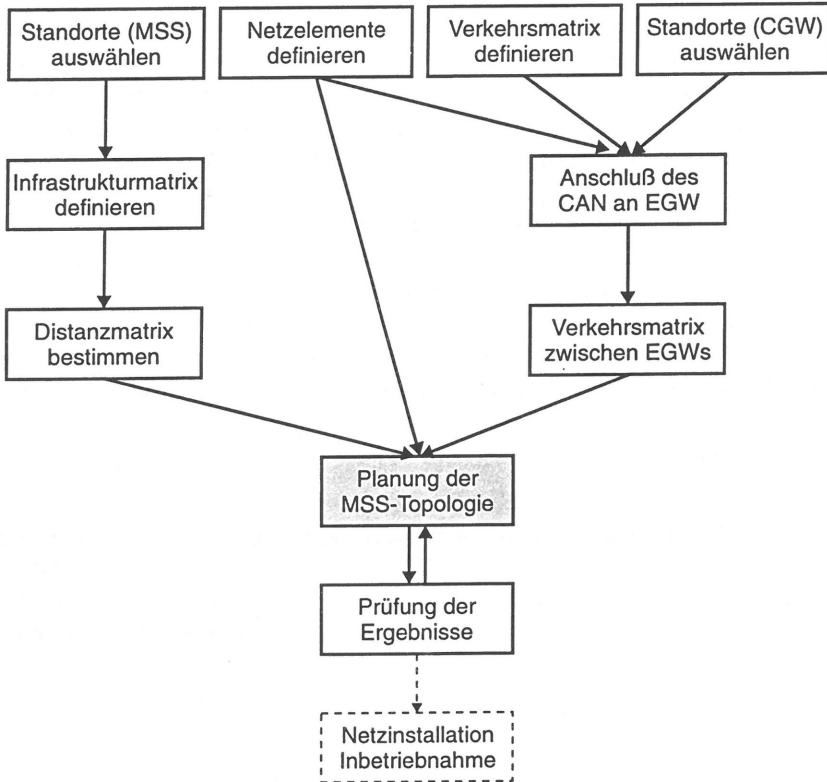


Bild 6.3: Entwurfsablauf

6.5 Bewertung und Einsatzmöglichkeiten

Teile des Netzplanungswerkzeugs wurden im Rahmen einer Industriekooperation mit der Firma Alcatel SEL entwickelt. Es ist vorgesehen, das Werkzeug dort zur Unterstützung bei der Planung und Projektierung einzusetzen. Obwohl erste Installationen von MAN-Vermittlungssystemen, die in Betrieb gegangen sind, noch eine relativ geringe Zahl an Netz-

elementen (*Edge Gateways, Router*) aufweisen – hier sei stellvertretend das DATEX-M-Netz in Deutschland genannt (momentan 12 *Edge Gateways* [149]) –, wird mit weitaus größeren MAN-Vermittlungssystemen gerechnet. So befindet sich in Großbritannien das sog. Super-JANET (*UK Joint Academic NETwork*) mit über 50 Netzknoten im Aufbau [29], das einen Großteil der Universitäten verbindet.

Das Netzplanungswerkzeug ist daher so ausgelegt, daß auch MAN-Vermittlungssysteme mit deutlich mehr als 100 Netzelementen geplant werden können. Da die Software redundanzsparend implementiert wurde, hält sich der Speicheraufwand bei Problemen dieser Dimension in Grenzen. Dabei sind für jeden Eintrag in den Verkehrsmatrizen, die den größten Teil des Speicherplatzes benötigen, lediglich 48 Bytes erforderlich. Wie die Untersuchungen in Kapitel 5 bereits andeuten, stoßen die statistischen Optimierungsverfahren auf Grund der Rechenzeit allmählich an ihre Grenzen. Ein MAN-Vermittlungssystem mit 100 bis 120 Netzelementen kann aber noch mit Hilfe von *Simulated Annealing* geplant werden. Die Anwendung des Dekompositionsverfahrens ist bis zu 500 Netzelementen problemlos möglich. Anzumerken ist, daß bei dieser Größenordnung meist administrative Beschränkungen entscheidend sind und der Umfang des MAN-Vermittlungssystems dadurch beschränkt wird. Beispielsweise kann der Aufwand zur Adreßverwaltung so groß werden, daß die Stationen vor Beginn der Optimierungsphase unterschiedlichen MAN-Vermittlungssystemen (vgl. Abschnitt 2.2.4) zugeordnet werden.

Problematisch wird bei diesen Größenordnungen die Beschreibung der Verkehrsmatrix für den Benutzer. Da die Eingabe einzelner Verkehrsbeziehungen aufwendig ist, wird, wie schon erwähnt, die Beschreibung von Teilmatrizen unterstützt, wobei alle Verkehrsbeziehungen den gleichen Wert für den Verkehrsfluß erhalten. Diese Vereinfachung stellt aber eine potentielle Fehlerquelle bei der Planung dar, so daß die geplanten Topologien dann sorgfältig analysiert werden sollten.

Zusammenfassend läßt sich feststellen, daß die Anwendung des Planungswerkzeugs für den Netzplaner einige Erfahrung erfordert, nicht zuletzt auf Grund der notwendigen Kenntnisse über die Wirkung der Planungsparameter. Dafür erhält der Netzplaner jedoch eine Reihe von Freiheitsgraden, die er zur Ermittlung verschiedener Netzalternativen und zur Optimierung der Topologie nutzen kann.

Kapitel 7

Zusammenfassung und Ausblick

Der Schwerpunkt der vorliegenden Arbeit lag auf der Entwicklung und Untersuchung von Verfahren, die zur Planung von regionalen Hochgeschwindigkeitsnetzen geeignet sind. Mit Hilfe der Verfahren können Breitband-Netzinfrastrukturen, die auf MAN-Vermittlungssystemen basieren, rechnergestützt entworfen werden.

Obwohl für MANs auf der Basis von DQDB zahlreiche, meist approximative Analyseverfahren in der Literatur bekannt sind, wurden im Bereich der Dimensionierung und des Entwurfs von MANs bisher nur ansatzweise Methoden entwickelt. Diese beschränken sich auf Teilaspekte, wie z. B. die Ermittlung der optimalen Position einer Netzkoppeleinheit in einem MAN-Teilnetz und, damit zusammenhängend, die Bestimmung optimaler Betriebspunkte bezüglich der Systemauslastung [122].

Ziel der Arbeit war es daher, Planungs- und Entwurfsverfahren für MAN-Vermittlungssysteme zu entwickeln, die verschiedene Teilaspekte berücksichtigen, und mit deren Hilfe ein MAN-Vermittlungssystem unter Einhaltung unterschiedlicher Dienstgüteparameter entworfen werden kann. Insbesondere sind dazu Anforderungen für die schnelle Datenkommunikation zu beachten, die mit Diensten wie SMDS und CBDS realisiert wird, um lokale Netze über das öffentliche Netz miteinander zu koppeln.

Zur näheren Untersuchung der Problematik wurde in Kapitel 4 ein Optimierungsmodell aufgestellt, das die aus den Beschränkungen des Entwurfsproblems resultierenden Nebenbedingungen umfaßt. Die Nebenbedingungen wurden mit Hilfe graphen- und verkehrstheoretischer Modelle abgeleitet. Da die wesentlichen Netzkosten durch das Mieten von Übertragungsleitungen verursacht werden, wurde eine kapazitäts- und entfernungsabhängige Kostenfunktion als Zielfunktion verwendet. Um eine eventuelle Verwendung von Standardverfahren zu prüfen, wurde auch die Komplexität des Problems untersucht. Dabei stellte sich heraus, daß das vorliegende Entwurfsproblem etliche Schwierigkeitsgrade aufweist. So sind bereits Teilprobleme NP-vollständig und damit nur durch Näherungsverfahren zu lösen. Auch ein

Lösungsverfahren, das den Zustandsraum durch teilweise oder vollständige Enumeration absucht, scheidet auf Grund der Größe des Zustandsraums aus.

Da für dieses Planungsproblem bisher keine geeigneten Planungsverfahren bekannt sind, wurden im Rahmen dieser Arbeit drei neue heuristische Verfahren für den topologischen Entwurf von MAN-Vermittlungssystemen entwickelt: Ein deterministisches Dekompositionsverfahren sowie Verfahren, die auf der Kombination von *Simulated Annealing* mit einem *Clustering*- und Lastteilungsalgorithmus bzw. auf Genetischen Algorithmen beruhen. Diese drei Verfahren, die in Kapitel 5 detailliert beschrieben sind, wurden anhand mehrerer charakteristischer Beispiele bezüglich ihrer Leistungsfähigkeit verglichen. Dazu wurden auch Verteilungen der Netzkosten zufällig ermittelter Topologien sowie eine untere Schranke des MSS-Entwurfsproblems herangezogen. Die in Kapitel 5 dokumentierten Ergebnisse zeigen, daß alle drei Verfahren Lösungen von hoher Qualität für durchaus unterschiedliche Probleme erzeugen. Diese Aussage stützt sich auf die Erkenntnis, daß die Lösungen das Minimum der oben genannten Verteilungen deutlich unterbieten, und in manchen Fällen der Abstand zur unteren Schranke relativ gering ist. Mit dem Dekompositionsverfahren können rasch Entwurfsalternativen bestimmt werden, die durch die statistischen Optimierungsverfahren bei erhöhtem Rechenaufwand noch verbessert werden können. Der Netzplaner besitzt damit verschiedene Möglichkeiten, eine günstige Netzstruktur zu ermitteln. Wie ebenfalls gezeigt wurde, kann bei geeigneter Variation der Nebenbedingungen und der Optimierungsparameter ein guter Kompromiß zwischen der Netzgüte und den Netzkosten erzielt werden.

Das Dekompositionsverfahren und das Verfahren, das auf *Simulated Annealing* beruht, lassen sich auch auf Probleme der Netzerweiterung anwenden. In Kapitel 5 wurde deutlich gemacht, wie diese Verfahren bei einer erforderlichen Erhöhung der Systemkapazität eingesetzt werden können. Als weitere Hierarchieebene in einem MAN-Vermittlungssystem sind beispielsweise *Backbone*-Netze in der Lage, einen Teil des Transitverkehrs aufzunehmen. Die Verfügbarkeit kann durch eine einfache Erweiterung des Dekompositionsverfahrens oder durch die Verwendung von *Backbone*-Netzen gesteigert werden.

Um die Verfahren möglichst effektiv für einen Netzplaner zugänglich zu machen, wurde ein Planungswerkzeug entwickelt, dessen Konzepte und Softwarearchitektur in Kapitel 6 skizziert sind. Dieses Werkzeug umfaßt neben den Verfahren zur Planung von MAN-Vermittlungssystemen auch die Möglichkeit, bestehende MAN-Vermittlungssysteme zu analysieren. Zusätzlich kann die Struktur der Zugangsnetze festgelegt werden. Besonderes Gewicht wurde bei der Realisierung auf eine modulare Architektur sowie auf die effiziente Gestaltung der Datenstrukturen und deren Zugriffsfunktionen gelegt, die eine Voraussetzung für günstige Programmlaufzeiten sind. Auf Grund des modularen Aufbaus kann das Werkzeug um weitere Netzplanungs- und Analyseverfahren erweitert werden.

Betrachtet man die entworfenen Planungsverfahren näher, lassen sich zwei Ansätze erkennen, die in die Entwicklung zukünftiger Planungsverfahren einfließen können. Zum einen sollte

eine Anwendung der Transformationen, die für das Verfahren des *Simulated Annealing* entwickelt wurden, bei der Planung anderer, ringförmiger Netze untersucht werden. Hier sind in erster Linie gekoppelte FDDI-Netze zu nennen, die im Bereich der privaten Datenkommunikation vorherrschend sind. Interessant wäre aber auch ein Einsatz bei der Planung von ringförmigen SDH-Netzen, die untereinander vermascht sind. Das gilt analog auch für das in Kapitel 5 entworfene Codierungsschema, das für ein Verfahren mit Genetischen Algorithmen benötigt wird.

Zum anderen können die Planungsverfahren so erweitert werden, daß auch das zukünftige B-ISDN als Transitnetz zwischen MAN-Vermittlungssystemen einbezogen wird. Zunächst ist vor allem an die Einführungsphase des B-ISDN gedacht, bei der ausschließlich semi-permanente VP- oder VC-Verbindungen zur Verfügung stehen werden, um MANs zu verbinden. Langfristig wird erwartet, daß das B-ISDN dominieren wird und MANs die Funktion von Zugangsnetzen übernehmen. Integrierte Planungsverfahren, die für beide Netzarten geeignet sind, könnten dabei den Prozeß der Netzevolution unterstützen und zu Lösungen beitragen, die für einen solchen Netzverbund optimal sind.

Literaturverzeichnis

- [1] E. Aarts, J. Korst: *Simulated Annealing and Boltzmann Machines*, J. Wiley, 1989.
- [2] A. F. Albers, S. Zachok, J. D. Lapidès: "Bell Atlantic/QPSX Public MAN Trial", *Proceedings of the 14th International Switching Symposium (ISS)*, Yokohama, Japan, Oktober 1992, Paper A2.4.
- [3] K. Altinkemer, Z. Yu: "Topological Design of Wide Area Communication Networks", *Annals of Operations Research*, Vol. 36, No. 1-4, Mai 1992, S. 365-381.
- [4] American National Standard: "FDDI Token Ring Media Access Control (MAC)", Doc. ANSI X3.139-1987, 1987.
- [5] F. Antonelli, M. Bianco, E. Peretti, A. Proietti: "SMDS: The Italian Network Operator's Point of View", *Proceedings of the 11th International Conference on Computer Communication (ICCC)*, Genova, Italy, September 1992, S. 223-228.
- [6] H. Armbrüster, H. J. Rothamel: "Breitbandanwendungen und -dienste: Qualitative und quantitative Anforderungen an künftige Netze", *Nachrichtentechnische Zeitschrift*, 43 (1990) Heft 3, S. 150-159.
- [7] H. R. van As: "The Evolution Towards High-Performance LANs and MANs", *Informationstechnik und Technische Informatik*, Jg. 35, Nr. 4, (1993), S. 9-19.
- [8] H. R. van As: "Media Access Techniques: The evolution towards terabit/s LANs and MANs", *Computer Networks and ISDN Systems*, Vol. 26, No. 6-8, März 1994, S. 603-656.
- [9] D. W. Bachmann, et. al.: "NetMod: A Design Tool for Large-Scale Heterogeneous Campus Networks", *IEEE Journal on Selected Areas in Communications*, Vol. 9, No. 1, Januar 1991, S. 15-24.
- [10] R. Bäßler, A. Deutsch: *Nachrichtennetze*, 2. Auflage, VEB Verlag Technik, Berlin, 1989.
- [11] M. Bayreuther: *Messung und Auswertung von Verkehrsprofilen in Metropolitan Area Networks*, Diplomarbeit Nr. 1192, Universität Stuttgart, Institut für Nachrichtenvermittlung und Datenverarbeitung, November 1992.

- [12] Bellcore: "Generic System Requirements in Support of Switched Multi-Megabit Data Service", *Technical Reference TR-TSV-000772*, Issue 1, Mai 1991.
- [13] D. Bertsekas, R. Gallager: *Data Networks*, Prentice Hall, Eaglewood Cliffs, N.J., USA, 1987.
- [14] S. Bessler, K. Nüssler: "Gateway Planning in the DATEX-P Packet Network", *Proceedings of the 10th International Conference on Computer Communication (ICCC)*, New Dehli, India, November 1990, S. 118-123.
- [15] C. Bisdikian: "A Queueing Model with Applications to Bridges and the DQDB (IEEE 802.6) MAN", *Computer Networks and ISDN Systems*, 25 (1993), S. 1279-1289.
- [16] M. Bonatti et.al.: "B-ISDN Economical Evaluation in Metropolitan Areas", *IEEE Journal on Selected Areas in Communications*, Vol. 10, No. 9, Dezember 1992, S. 1489-1502.
- [17] R. R. Boorstyn, H. Frank: "Large-Scale Network Topological Optimization", *IEEE Transactions on Communications*, Vol. 25, No. 1, Januar 1977, S. 29-47.
- [18] M. Burak: "Intelligent Bridges and Routers", *Proceedings of the GLOBECOM '92*, Orlando, USA, Dezember 1992, S. 1234-1239.
- [19] L. Burger: "Architektur eines wissensbasierten Netzwerkplanungssystems", *Tagungsband Kommunikation in verteilten Systemen*, ITG/GI-Fachtagung, Stuttgart, Februar 1989, S. 818-833.
- [20] R. G. Busacker, T. L. Saaty: *Endliche Graphen und Netzwerke*, R. Oldenbourg Verlag, München, 1968.
- [21] R. S. Cahn, P. Chang, P. Kermani, A. Kershenbaum: "INTREPID: An Integrated Network Tool for Routing, Evaluation of Performance, and Interactive Design", *IEEE Communications Magazine*, Vol. 29, No. 7, Juli 1991, S. 40-47.
- [22] G. Cappellini, A. Massari: "DOPNET: A PC Tool for Design and Planning Packet Switching Networks" *Teletraffic and Datatraffic - Socio-Economic Aspects ITC-13*, A. Jensen, V.B. Iversen, Editors, Elsevier Science Publishers B.V. (North-Holland) 1992, S. 439-444.
- [23] M. Capurro, M. Coronaro, D. Giuli: "Metropolitan Area Network Planning and Design", *European Transactions on Telecommunications and related Technologies*, Vol. 2, No. 4, Juli-August 1991, S. 423-432.
- [24] M. Capurro: "Planning and Design Criteria for Metropolitan Area Networks", *Proceedings EFOC/LAN '92*, Paris, France, Juni 1991, S. 184-189.
- [25] V. Catania, M. Gerla, C. Pavanelli: "A Routing Strategy for MAN Interconnection", *Proceedings of the IEEE INFOCOM '91*, Bal Harbour, USA, April 9-11 1991, S. 1353-1359.

- [26] V. Catania, A. Puliafito, L. Vita: "High-speed Data Service in Distributed Systems based on SMDS", *Computer Communications*, Vol. 16, No. 7, Juli 1993, S. 394-402.
- [27] K. C. Claffy, G. C. Polyzos, H. Braun: "Traffic Characteristics of the T1 NSFNET Backbone", *Proceedings of the IEEE INFOCOM '93*, San Francisco, USA, März 1993, S. 885-892.
- [28] G. H. Clapp: "LAN Interconnection Across SMDS", *IEEE Network*, Vol. 5, No. 5, September 1991, S. 25-32.
- [29] L. Clyne: "SuperJANET Update", *Computer Networks and ISDN Systems*, Vol. 25, Suppl. 3, 1994, S. S111-S116.
- [30] S. Cosares, I. Saniee: "An Optimization Problem Related to Balancing Loads on SONET Rings", *Telecommunication Systems*, Vol. 3 (1994), S. 165-181.
- [31] R. Dittmann, T. Stock, P. Tran-Gia: "Das DQDB-Zugriffsprotokoll in Hochgeschwindigkeitsnetzen und der IEEE-Standard 802.6", *Informatik-Spektrum*, (1993) 16, S. 143-158.
- [32] F. R. Dix, M. Kelley, R. W. Klessig: "Access to a Public Switched Multi-Megabit Data Service Offering", *ACM SIGCOMM Computer Communications Review*, Vol. 20, No. 3, Juli 1990, S. 46-61.
- [33] R. C. Dixon, D. A. Pitt: "Addressing, Bridging and Source Routing", *IEEE Network Magazine*, Vol. 2, No. 1, Januar 1988, S. 25-32.
- [34] G. Dueck, T. Scheuer, H.-M. Wallmeier: "Toleranzschwelle und Sintflut: neue Ideen zur Optimierung", *Spektrum der Wissenschaft*, März 1993, S. 42-51.
- [35] A. Dutta, J. Lim: "A Multiperiod Capacity Planning Model for Backbone Computer Communication Networks", *Operations Research*, Vol. 40, No. 4, Juli-August 1992, S. 689-705.
- [36] A. Dutta: "Capacity Planning of Private Networks Using DCS Under Multibusy-Hour Traffic", *IEEE Transactions on Communications*, Vol. 42, No. 7, Juli 1994, S. 2371-2374.
- [37] H. G. Dysart, N. D. Georganas: "NEWCLUST: An Algorithm for the Topological Design of Two-Level, Multidrop Teleprocessing Networks", *IEEE Transactions on Communications*, Vol. 26, No. 1, Januar 1978, S. 55-62.
- [38] J. Eberspächer: "Netztechnologien für Corporate Networks: Stand und Tendenzen", *Praxis der Informationsverarbeitung und Kommunikation (PIK)*, 18. Jahrgang, Heft 1/95, Januar - März 1995, S. 4-10.
- [39] A. Engel, H. Orlamünder: "MANs - The First Step to B-ISDN", *Proceedings of the 11th International Conference on Computer Communication (ICCC)*, Genova, Italy, September 1992, S. 103-108.

- [40] R. Elbaum, M. Sidi: "Topological Design of Local Area Networks Using Genetic Algorithms", *Proceedings of the IEEE INFOCOM '95*, Boston, USA, April 1995, S. 64-71.
- [41] C. Ersoy, S. S. Panwar: "Topological Design of Interconnected LAN/MAN Networks", *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 8, Oktober 1993, S. 1172-1182.
- [42] ETSI European Telecommunication Standard (ETS) 300 211: "Network Aspects (NA); Metropolitan Area Network (MAN) Principles and Architecture", Dezember 1992.
- [43] ETSI European Telecommunication Standard (ETS) 300 217: "Network Aspects (NA); Connectionless Broadband Data Service (CBDS)", Dezember 1992.
- [44] ETSI European Telecommunication Standard (ETS) 300 275 (Draft): "Network Aspects (NA); Metropolitan Area Network (MAN) Interconnection of MANs", Juni 1993.
- [45] EWOS Technical Guide 002: "OSI Layers 1 to 4 Addressing", Januar 1990.
- [46] R. M. Falconer, J. L. Adams: "ORWELL: A Protocol for an Integrated Services Local Network", *British Telecom Technology Journal*, Vol. 3, No. 4, Oktober 1985, S. 27-35.
- [47] K. Fendick, V. Saksena: "Traffic Engineering Operations for the ACCUNET Packet Service Network", *Proceedings of the IEEE GLOBECOM '88*, USA, November/Dezember 1988, S. 342-346.
- [48] A. Fenyves, A. Lazzari: "Introducing Metropolitan Networks in Italian Urban Areas: Preliminary Results of a Field Experience", *Proceedings of the 11th International Conference on Computer Communication (ICCC)*, Genova, Italy, September 1992, S. 429-435.
- [49] P. C. Fetterolf, G. Anandalingam: "A Lagrangian Relaxation Technique for Optimizing Interconnection of Local Area Networks", *Operations Research*, Vol. 40, No. 4, Juli-August 1992, S. 678-688.
- [50] P. C. Fetterolf, G. Anandalingam: "Optimal Design of LAN-WAN Internetworks: An Approach Using Simulated Annealing", *Annals of Operations Research*, Vol. 36, No. 1-4, Mai 1992, S. 275-297.
- [51] W. Fischer, et. al.: "The Evolution from LAN/MAN to Broadband ISDN", *Proceedings of the ICC '91*, Denver, USA, Juni 1991, S. 1251-1257.
- [52] M. L. Fisher: "The Lagrangian Relaxation Method for Solving Integer Programming Problems", *Management Science*, Vol. 27, No. 1, Januar 1981, S. 1-18.
- [53] D. B. Fogel, "An Introduction to Simulated Evolutionary Optimization": *IEEE Transactions on Neural Networks*, Vol. 5, No. 1, Januar 1994, S. 3-14.
- [54] H. Frank, W. Chou, "Topological Optimization of Computer Networks": *Proceedings of the IEEE*, Vol. 60, No. 11, November 1972, S. 1285-1397.

- [55] R. Garcia: "Cluster-Analyse", *Der Fernmelde-Ingenieur*, 33. Jg., Heft 2, Februar 1979, S. 1-34.
- [56] M. W. Garrett, S. Li: "A Study of Slot Reuse in Dual Bus Multiple Access Networks", *Proceedings of the IEEE INFOCOM '90*, San Francisco, USA, Juni 1990, S. 617-629.
- [57] B. Gavish, I. Neumann: "A System for Routing and Capacity Assignment in Computer Communication Networks", *IEEE Transactions on Communications*, Vol. 37, No. 4, April 1989, S. 360-366.
- [58] A. Gerges, U. Reiff, H. Schenk: "A MAN System based on the DQDB Protocol", *Proceedings EFOC/LAN'91*, London, UK, Juni 1991, S. 367-373.
- [59] M. Gerla, L. Kleinrock: "On the Topological Design of Distributed Computer Networks", *IEEE Transactions on Communications*, Vol. 25, No. 1, Januar 1977, S. 48-60.
- [60] D. E. Goldberg: *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing Company Inc., 1989.
- [61] F. J. Gratzner, P. J. Sanchirico: "Switched Multi-Megabit Data Service (SMDS)", *Proceedings of the 11th International Conference on Computer Communication (ICCC)*, Genova, Italy, September 1992, S. 217-222.
- [62] U. Gremmelmaier, P. J. Kühn: "Breitband-Netzinfrastrukturen auf der Basis von Metropolitan Area Networks (MAN)", *Tagungsband der 17. Europäischen Congressmesse für Technische Kommunikation (ONLINE'94)*, Hamburg, Februar 1994, Beitrag C 224.
- [63] S. Gupta, K. W. Ross: "Performance Modeling and Optimization of Interconnected Ethernets", *Proceedings of the IEEE INFOCOM '91*, Bal Harbour, USA, April 1991, S. 1353-1359.
- [64] V. P. Gupta: "What is Network Planning?", *IEEE Communications Magazine*, Vol. 23, No. 10, Oktober 1985, S. 10-16.
- [65] R. Händel, M. N. Huber, S. Schröder: *ATM Networks*, Addison-Wesley Publishing Company, Wokingham, 1994.
- [66] E. L. Hahne, A. K. Choudhury, N. F. Maxemchuk: "Improving the Fairness of the Distributed Queue Dual Bus Networks", *Proceedings of the IEEE INFOCOM '90*, San Francisco, USA, Juni 1990, S. 175-184.
- [67] K. Heidtmann: "Methoden zur Zuverlässigkeitsanalyse von Rechnernetztopologien", *Informatik-Spektrum*, (1994) 17, S. 232-244.
- [68] F. S. Hillier, G. J. Lieberman: *Introduction to Operations Research*, 4th Edition, McGraw-Hill, Inc., 1986.
- [69] G. Horn: "A layered Concept for a Constant Bit Rate Service on a MAN", *Proceedings EFOC/LAN '92*, Paris, France, Juni 1991, S. 204-209.

- [70] IEEE Standard 802.1D: "Media Access Control (MAC) Bridges", 1990.
- [71] IEEE Draft Standard 802.1G: "Remote MAC Bridging", März 1993.
- [72] IEEE Standard 802.3: "Carrier Sense Multiple Access with Collision Detection (CSMA/CD): Access Method and Physical Layer Specifications", 1990.
- [73] IEEE Standard 802.4: "Token Passing Bus: Access Method and Physical Layer Specifications", 1990.
- [74] IEEE Standard 802.5: "Token Ring: Access Method and Physical Layer Specifications", 1990.
- [75] IEEE Standard 802.6: "Distributed Queue Dual Bus (DQDB) Subnetwork of a Metropolitan Area Network (MAN)", Dezember 1990.
- [76] K. Imai, T. Ito, H. Kasahara, N. Morita: "ATMR: Asynchronous Transfer Mode Ring Protocol", *Computer Networks and ISDN Systems*, 26 (1994), S. 785-798.
- [77] ISO IS 7498: "Information Processing Systems - Open Systems Interconnection - Basic Reference Model", 1984.
- [78] ISO IS 8348 Addendum 2: "Information Processing Systems - Data Communications - Network Service Definition: Addendum 2: Network Layer Addressing", 1988.
- [79] ISO/IEC TR 9575: "Information Technology - Telecommunications and Information Exchange between Systems - OSI Routing Framework", 1990.
- [80] ISO/IEC DIS 10589: "Intermediate system to Intermediate system Intra-Domain routing exchange protocol for use in Conjunction with the Protocol for providing the Connectionless-mode Network Service (ISO 8473)", 1990.
- [81] ITG-Empfehlungen: "ITG 1.6/0.1 Empfehlung 1987: ISDN-Begriffe", *Nachrichtentechnische Zeitschrift*, Bd. 40, Heft 11, S. 814-819, November 1987.
- [82] ITU-T Recommendation E.164: "Numbering Plan for the ISDN Era", 1988.
- [83] S. Jidarian, D. M. Shapiro: "Traffic Administration and Engineering of a Corporate Wide Area Network", *IEEE Network*, Vol. 5, No. 2, March 1991, S. 25-32.
- [84] D. S. Johnson et.al.: "Optimization by Simulated Annealing: An Experimental Evaluation; Part I, Graph Partitioning", *Operations Research*, Vol. 37, No. 6, November-Dezember 1989, S. 865-892.
- [85] D. S. Johnson et.al.: "Optimization by Simulated Annealing: An Experimental Evaluation; Part II, Graph Coloring an Number Partitioning", *Operations Research*, Vol. 39, No. 3, Mai-Juni 1991, S. 378-406.
- [86] J. Jordan: *Untersuchung der Leistungsfähigkeit von Genetischen Algorithmen*, Semesterarbeit Nr. 1332, Universität Stuttgart, Institut für Nachrichtenvermittlung und Datenverarbeitung, Juni 1994.

- [87] H. Jüchter, M. Hammer: "Rechnergestützte Netzplanung für PDH- und SDH-Netze", *Philips Telecommunication Review*, Vol. 51, No. 1, März 1993, S. 43-50.
- [88] D. Jungnickel: *Graphen, Algorithmen, Datenstrukturen*, Wissenschaftsverlag Mannheim, 1990.
- [89] J. Kadirire: "Minimising Packet Copies in Multicast Routing by Exploiting Geographic Spread", *ACM SIGCOMM Computer Communications Review*, Vol. 24, No. 3, Juli 1994, S. 47-62.
- [90] B. W. Kernighan, S. Lin: "An Efficient Heuristic Procedure for Partitioning Graphs", *Bell System Technical Journal*, Vol. 49, Februar 1970, S. 291-307.
- [91] A. Kershenbaum, et. al.: "MENTOR: An Algorithm for Mesh Network Topological Optimization and Routing", *IEEE Transactions on Communications*, Vol. 39, No. 4, April 1991, S. 503-513.
- [92] A. Kershenbaum: *Telecommunications Network Design Algorithms*, McGraw-Hill, Inc., New York, USA, 1993.
- [93] G. C. Kessler, D. A. Train: *Metropolitan Area Networks: Concepts, Standards and Services*, McGraw Hill, Inc., New York, USA, 1992.
- [94] K. M. Khalil, P. A. Spencer: "A Systematic Approach for Planning, Tuning and Upgrading Local Area Networks", *Proceedings of the IEEE GLOBECOM '91*, Phoenix, USA, Dezember 1991, S. 658-653.
- [95] T. King, G. Whalley, A. Smith, S. Evans: "Customer Internetworking Solutions", *Proceedings of the 15th International Switching Symposium (ISS)*, Berlin, April 1995, Paper A3.1.
- [96] S. Kirkpatrick et.al.: "Optimization by Simulated Annealing", *Science*, Mai 1983, S. 671-680.
- [97] L. Kleinrock: *Queueing Systems, Volume II: Computer Applications*, J. Wiley, USA, 1976.
- [98] P. J. Kühn: "Approximate Analysis of General Queueing Networks by Decomposition", *IEEE Transactions on Communications*, Vol. 27, No. 1, Januar 1979, S. 113-126.
- [99] P. J. Kühn: *Nachrichtenvermittlung I*, Vorlesungsmanuskript, IND, Universität Stuttgart, Stuttgart, 1993.
- [100] P. J. Kühn: *Wartezeitprobleme der Daten- und Nachrichtenverkehrstheorie*, Vorlesungsmanuskript, IND, Universität Stuttgart, Stuttgart, 1993.
- [101] L. J. Leblanc, S. Narasimhan: "Topological Expansion of Metropolitan Area Networks", *Computer Networks and ISDN Systems*, 26 (1994), S. 1235-1248.
- [102] J. Le Boudec: "The Asynchronous Transfer Mode: A Tutorial". *Computer Networks and ISDN Systems*, 24 (1992), S. 279-309.

- [103] B. M. Leiner: "Internet Technology", *Communications of the ACM*, Vol. 37, No. 8, August 1994, S. 32.
- [104] V. C. M. Leung, N. Quian, A. S. Malyan, R. W. Donaldson: "Call Control and Traffic Transport for Connection-Oriented High-Speed Wireless Personal Communications Over Metropolitan Area Networks", *IEEE Journal on Selected Areas in Communications*, Vol. 12, No. 8, Oktober 1994, S. 1376-1388.
- [105] S. Liang, J. R. Lee: "Locating Internet Gateways to Minimize Nonlinear Congestion Costs", *IEEE Transactions on Communications*, Vol. 42, No. 9, September 1994, S. 2740-2750.
- [106] F. Y. S. Lin: "Link Set Sizing for Networks Supporting SMDS", *IEEE/ACM Transactions on Networking*, Vol. 1, No. 6, Dezember 1993, S. 729-739.
- [107] F. Y. S. Lin: "Link Set Capacity Augmentation Algorithms for Networks Supporting SMDS", *Proceedings of the ICC '94*, Denver, USA, Juni 1994, S. 624-629.
- [108] S. Lin: "Heuristic Programming as an Aid to Network Design", *Networks*, Vol. 5, 1975, S. 33-43.
- [109] S. Lin: "Computer Solutions of the Traveling Salesman Problem", *Bell System Technical Journal*, Vol. 44, 1965, S. 2245-2269.
- [110] A. D. Malyan et.al.: "Network Architecture and Signaling for Wireless Personal Communications", *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 6, August 1993, S. 830-841.
- [111] P. Marsden: "Interworking IEEE 802/FDDI LAN's Via the ISDN Frame Relay Bearer Service", *Proceedings of the IEEE*, Vol. 79, No. 2, Februar 1991, S. 223-229.
- [112] J. McGibney, D. D. Botvich, T. Curran: "Modern Global Optimisation Heuristics in the long term Planning of Networks", *Proceedings of the Australian Telecommunication Networks and Applications Conference 1994*, Melbourne, Australien, Dezember 1994, S. 317-322.
- [113] P. V. McGregor, D. Shen: "Network Design: An Algorithm for the Access Facility Location Problem", *IEEE Transactions on Communications*, Vol. 25, No. 1, Januar 1977, S. 61-73.
- [114] I. Meyer: *Analyseverfahren für das MAN DQDB*, Diplomarbeit Nr. 1142, Universität Stuttgart, Institut für Nachrichtenvermittlung und Datenverarbeitung, Mai 1992.
- [115] Z. Michalewicz: *Genetic Algorithms + Data Structures = Evolutionary Programs*, Springer-Verlag, Berlin, 1992.
- [116] N. Modiri: "The ISO Reference Model Entities", *IEEE Network*, Vol. 5, No. 4, Juli 1991, S. 24-33.

- [117] J. F. Mollenauer: "Standards for Metropolitan Area Networks", *IEEE Communications Magazine*, Vol. 26, No. 4, April 1988, S. 15-19.
- [118] C. L. Monma, D. F. Shallcross: "Methods for Designing Communication Networks with Certain Two-Connected Survivability Constraints", *Operations Research*, Vol. 37, No. 4, Juli-August 1989, S. 531-541.
- [119] H. Müller-Merbach: *Optimale Reihenfolgen*, Springer-Verlag, Berlin, 1970.
- [120] B. Mukherjee, C. Bisdikian: "A Journey through the DQDB Network Literature", *Performance Evaluation*, Vol. 16, No. 1-3, November 1992, S. 129-158.
- [121] K. H. Muralidhar, M. K. Sundareshan: "On the Decomposition of Large Communication Networks for Hierarchical Control Implementation", *IEEE Transactions on Communications*, Vol. 34, No. 10, Oktober 1986, S. 985-987.
- [122] H. Nagano, S. Tasaka: "The Effect of the Gateway Location on the Performance of DQDB Networks", *Proceedings of the ICC '94*, Denver, USA, Juni 1994, S. 962-967.
- [123] B. Neef: "CBDS/SMDS Broadband Service Introduction", *Proceedings of the 11th International Conference on Computer Communication (ICCC)*, Genova, Italy, September 1992, S. 229-234.
- [124] R. M. Newman, Z. L. Budrikis, J. L. Hullett: "The QPSX MAN", *IEEE Communications Magazine*, Vol. 26, No. 4, April 1988, S. 20-28.
- [125] H. van Norman: *LAN/WAN Optimization Techniques*, Artech House, Inc., Boston, London, 1992.
- [126] R. Oechsle, H. L. Truong: "Shortest Path Bridging for Broadband LAN Interconnection", *Proceedings EFOC/LAN '92*, Paris, France, Juni 1992, S. 184-189.
- [127] R. Oppliger, S. Weber, D. Hogrefe: "Entwurf von virtuell privaten Netzen", *Tagungsband Kommunikation in verteilten Systemen*, ITG/GI-Fachtagung, München, März 1993, S. 428-441.
- [128] A. R. Pach, S. Palazzo, D. Panno: "Internetworking of DQDB Subnetworks: A Maximum Throughput Analysis", *Proceedings of the International Conference on Local and Metropolitan Communication Systems*, Kyoto, Japan, Dezember 1994, S. 335-348.
- [129] A. Patel, V. Ryan: "Introduction to Names, Addresses and Routes in an OSI Environment", *Computer Communications*, Vol. 13, No. 1, Januar/Februar 1990, S. 27-36.
- [130] R. Perlman: *Interconnections: Bridges and Routers*, Addison-Wesley, Reading, MA, USA, 1992.
- [131] H. Pirkul, A. Amiri: "Routing in Packet-Switched Communication Networks", *Computer Communications*, Vol. 17, No. 5, Mai 1994, S. 307-316.

- [132] A. Proietti et.al.: "High-Speed Network Field Trials: Early Solutions for Broadband Service Provision", *Proceedings of the 11th International Conference on Computer Communication (ICCC)*, Genova, Italy, Sep. 1992, S. 423-428.
- [133] A. Proietti et.al.: "The Role of High-Speed Network Field Trials for the Assessment of Early Available Products", *Proceedings of the 14th International Switching Symposium (ISS)*, Yokohama, Japan, Oktober 1992, Paper A1.4.
- [134] I. Rechenberg: *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, Friedrich Frommann Verlag, Stuttgart, 1973.
- [135] K. Richter, P. Bachmann, S. Dempe: *Diskrete Optimierungsmodelle: Effektive Algorithmen und näherungsweise Lösung*, VEB Technik Verlag, Berlin, 1988.
- [136] F. E. Ross: "An Overview of FDDI: The Fiber Distributed Data Interface", *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 7, September 1989, S. 1043-1051.
- [137] M. N. O. Sadiku, A. S. Arvind: "Annotated Bibliography on Distributed Queue Dual Bus", *Computer Communication Review*, Vol. 24, No. 1, Januar 1994, S. 21-36.
- [138] V. R. Saksena: "Topological Analysis of Packet Networks", *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 8, Oktober 1989, S. 1243-1252.
- [139] P. von Schau: "Entwicklung der Telekommunikationsnetze: Markt und Technik im Wandel", *Nachrichtentechnische Zeitschrift*, Bd. 47 (1994), Heft 3, S. 146-152.
- [140] H. J. J. H. Schepers et. al.: "Naming, Addressing, Routing and Relaying in the OSI Environment", *Computer Networks and ISDN Systems*, Vol. 23, (1992), S. 241-251.
- [141] H. J. J. H. Schepers et. al.: "LAN/WAN Interworking in the OSI Environment", *Computer Networks and ISDN Systems*, Vol. 23, (1992), S. 253-266.
- [142] W. Schödl: *Kopplung von DQDB-Regionalnetzen mit ATM-Weitverkehrsnetzen: Architektur, Steuerstrategien und Leistungsverhalten*, 58. Bericht über verkehrstheoretische Arbeiten (Dissertationsschrift), Universität Stuttgart, Institut für Nachrichtenvermittlung und Datenverarbeitung, 1994.
- [143] M. Schwartz, T. E. Stern: "Routing Techniques Used in Computer Communication Networks", *IEEE Transactions on Communications*, Vol. 28, No. 4, April 1980, S. 539-552.
- [144] P. Semal, K. Wirl: "Optimal Clustering and Ring Creation in the Network Planning System PHANET", *Proceedings of the Sixth International Planning Symposium (Networks '94)*, Budapest, Hungary, September 1994, S. 303-308.
- [145] M. E. Sosa, M. A. Figueroa: "An OSPF-Based Routing Algorithm for SMDS Networks", *Proceedings of the GLOBECOM '92*, Orlando, USA, Dezember 1992, S. 1057-1061.

- [146] J. Spaeth, U. Gremmelmaier, U. Briem, M. N. Huber: "Design and Planning Aspects of Future Photonic Transport Networks", *Proceedings EFOC/LAN '95*, Brighton, UK, Juni 1995, S. 60-63
- [147] M. Srinivas, L. M. Patnaik: "Genetic Algorithms: A Survey", *IEEE Computer*, Vol. 27, No. 6, Juni 1994, S. 17-26.
- [148] S. L. Sutherland, J. Burgin: "B-ISDN Interworking", *IEEE Communications Magazine*, Vol. 31, No. 8, August 1993, S. 60-63.
- [149] J. Stieb, L. Griebel: "Datex-M: Schon ATM-kompatibel", *telekom praxis*, 12/93, S. 19-22.
- [150] T. Tai, M. Gerla: "LAN Interconnection: A Transparent, Shortest-Path Approach", *Proceedings of the ICC '91*, Denver, USA, Juni 23-26 1991, S. 1666-1670.
- [151] A. S. Tanenbaum: *Computer Networks*, Second Edition, Prentice-Hall, Eaglewood Cliffs, N.J., USA, 1989.
- [152] M. Tangemann: *Modellierung und Analyse von lokalen Hochgeschwindigkeitsnetzen mit zeitgesteuerten Token-Passing-Medienzugriffsprotokollen*, 57. Bericht über verkehrstheoretische Arbeiten (Dissertationsschrift), Universität Stuttgart, Institut für Nachrichtenvermittlung und Datenverarbeitung, 1994.
- [153] A. N. Tantawy, M. Zitterbart: "A Scheme for High-Performance LAN Interconnection Across Public MAN's", *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 8, Oktober 1993, S. 1133-1144.
- [154] P. Tran-Gia, T. Stock: "Approximate Performance Analysis of the DQDB Access Protocol", *Computer Networks and ISDN Systems*, Vol. 20 (1990), S. 231-240.
- [155] G. Walther: *Entwurf und Implementierung eines Netzplanungstools für Metropolitan Area Networks*, Diplomarbeit Nr. 1268, Universität Stuttgart, Institut für Nachrichtenvermittlung und Datenverarbeitung, August 1993.
- [156] J. L. Wang, E. A. White: "Integrated Methodology for Supporting Packet Network Performance Management and Planning", *Computer Communications*, Vol. 13, No. 9, November 1990, S. 558-570.
- [157] J. L. Wang: "An Integrated Methodology for Supporting Network Planning and Traffic Engineering with Considerations to SMDS Service", *Proceedings of the IEEE INFOCOM '91*, Bal Harbour, USA, April 1991, S. 333-342.
- [158] I. Wegener: *Theoretische Informatik*, Teubner Verlag, Stuttgart, 1993.
- [159] G. Wöhlbier (Hrsg.): *Planung von Telekommunikationsnetzen - Band I*, R. v. Decker's Verlag, G. Schenck, Heidelberg, 1990.
- [160] J. R. Yee, S. Liang: "A Second-Order Greedy Algorithm for Interconnecting Ethernet", *Proceedings of the ICC '91*, Denver, USA, Juni 1991, S. 1305-1310.

- [161] B. Zell, P. C. Fetterolf: "An Approach to Generating Network Requirements", *Proceedings of the IEEE INFOCOM '91*, Bal Harbour, USA, April 1991, S. 754-763.
- [162] T. G. Zerbic: "Considering the Past and Anticipating the Future for Private Data Networks", *IEEE Communications Magazine*, Vol. 30, No. 3, März 1992, S. 36-46.
- [163] T. Zillgitt: *Erweiterung eines MAN-Netzplanungstools um ein neues Routing-Verfahren*, Semesterarbeit Nr. 1331, Universität Stuttgart, Institut für Nachrichtenvermittlung und Datenverarbeitung, Juni 1994.

