

Optimierung des Fahrzeugeinsatzes im öffentlichen Personennahverkehr

Modelle, Methoden und praktische Anwendungen

D I S S E R T A T I O N

zur Erlangung des akademischen Grades
doctor rerum politicarum
(dr. rer. pol.)
im Fach Wirtschaftsinformatik

eingereicht an der
Fakultät für Wirtschaftswissenschaften
Universität Paderborn

von
Frau Dipl.-Wirt.Inf. Natalia Kliewer

Dekan der Fakultät für Wirtschaftswissenschaften:

Prof. Dr. Peter F. E. Sloane

Gutachter:

1. Prof. Dr. Leena Suhl
2. Prof. Dr. Ludwig Nastansky

Paderborn, im März 2005

Danksagungen

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftliche Mitarbeiterin am Lehrstuhl für Wirtschaftsinformatik IV der Universität Paderborn. Ich möchte mich an dieser Stelle bei allen bedanken, die zum Gelingen dieser Arbeit beigetragen haben.

Zuerst gilt mein besonderer und herzlicher Dank Prof. Dr. Leena Suhl für die Betreuung der Arbeit und für die hervorragende Forschungsatmosphäre, die sie mit großen wissenschaftlichen und sozialen Kompetenz in ihrer Arbeitsgruppe schafft. Für die fachliche Unterstützung in der Themenfindungsphase und die fruchtbare Zusammenarbeit in der besonders wichtigen Anfangsphase der Promotion bedanke ich mich bei Prof. Dr. Taïeb Mellouli.

Ich möchte mich bei allen Kollegen am Lehrstuhl für die gute Zusammenarbeit und die vielen interessanten Diskussionen bedanken. Ausdrücklich nennen möchte ich die Kollegen, mit denen ich eng wissenschaftlich zusammengearbeitet habe: Claus Biederbick, Stefan Bunte, Vitali Gintner, Thomas Knechtel. Prof. Dr. Uwe Suhl und Achim Koberstein danke ich für die hilfreichen Ratschläge im Umgang mit LP- und MIP-Solvern.

Die Arbeit entstand im Rahmen einer Forschungskooperation mit der PTV Planung Transport Verkehr AG. Für die langjährige Zusammenarbeit im Bereich Busumlauf- und Fahrereinsatzplanung danke ich an dieser Stelle den Mitarbeitern der PTV AG in Karlsruhe und Braunschweig: Michael Beck, Alexander Hofmann und Dr. Peter Mott.

Ich möchte mich auch bei Prof. Dr. Ludwig Nastansky für die Begutachtung dieser Dissertation und bei Prof. Dr. Klaus Kleibohm und PD Dr. Stefan Bock für die interessanten Diskussionen und die wertvollen Hinweise zur schriftlichen Ausarbeitung der Dissertation bedanken. Bei Antonia Knübel, Thomas Knechtel und Michael Beck bedanke ich mich zusätzlich für das Korrekturlesen der Arbeit.

Schließlich möchte ich mich bei Georg, Jan, meinen Eltern und Freunden für ihren Rückhalt, ihre Unterstützung und Geduld in den letzten Jahren bedanken.

Vielen herzlichen Dank!

Paderborn, im Mai 2005

Natalia Kliewer

Inhaltsverzeichnis

1	Einleitung und Motivation	1
2	Der Planungsprozess im ÖPNV	7
2.1	Strategische und operative Planung	7
2.2	Planungsschritte der operativen Planung	9
2.2.1	Linienplanung	10
2.2.2	Fahrplanerstellung	11
2.2.3	Fahrzeugumlaufplanung	11
2.2.4	Dienstplanung	12
2.2.5	Reihenfolge der Planungsschritte	14
2.3	Umlaufbildung als Aufgabe der operativen Ressourceneinsatzplanung	14
2.3.1	Mehrdepot-Umlaufplanung	17
2.3.2	Kennzahlen für die Bewertung der Umlaufpläne	17
2.3.3	Praktische Anforderungen an die Umlaufplanung	19
2.4	Umlaufplanungswerkzeuge mit Einsatz von Optimierungsmethoden	24
3	Mehrdepot-Umlaufplanung - Stand der Forschung	27
3.1	Mathematische Optimierung	27
3.1.1	Branch-and-Bound Verfahren	29
3.1.2	Heuristische Verfahren	31
3.2	Mehrdepot-Umlaufplanungsproblem	31
3.3	Grundlegende Modellierungsansätze	33
3.3.1	Mehrgüter-Fluss-Modelle	34
3.3.2	Quasi-Assignment-Modelle	40
3.3.3	Set-Partitioning-Modelle	44
3.3.4	Spezielle Aspekte der Modellierung	45
3.3.5	Zielfunktionen und Kostenbeachtung	47
3.4	Lösungsansätze	50
3.4.1	Exakte Ansätze	51
3.4.2	Heuristische Ansätze	56
3.5	Motivation für eine alternative Modellierung	59

4	TSN-basierte Modellierung und Lösung	61
4.1	Modellierung	62
4.1.1	Das Netzwerk	63
4.1.2	Das Flussmodell	69
4.1.3	Das mathematische Modell	70
4.1.4	Einfaches Beispiel für den Netzwerkaufbau	74
4.2	Lösungsansatz	81
4.3	Numerische Ergebnisse	81
4.3.1	Probleminstanzen und Testergebnisse	82
4.3.2	Wahl des LP-Verfahrens: Laufzeitanalyse	85
4.3.3	Minimierung der Fahrzeuganzahl	86
5	Flussdekomposition für TSN	87
5.1	Dekomposition der Flüsse in Wege	87
5.2	Die Dekompositionsstrategien	88
5.2.1	<i>FIFO</i> : First In First Out	90
5.2.2	<i>LIFO</i> : Last In First Out	90
5.2.3	<i>MinAlternation</i> : Greedy-Verknüpfung	91
5.2.4	Zuordnungsproblem-Strategien	91
5.2.5	Vorsortierung der Haltestellen	95
5.2.6	Globale Dekompositionsstrategie: <i>SPP</i>	95
5.3	Numerische Ergebnisse	96
6	Einbeziehung praktischer Randbedingungen und Zielsetzungen	99
6.1	Multikriterielle Optimierung: Kosten und Linienreinheit	99
6.1.1	Linienkanten im Netzwerkmodell	100
6.1.2	Erweiterung des mathematischen Modells	101
6.1.3	Ergebnisse	102
6.2	Verschiebeintervalle für Fahrten	106
6.2.1	Erweiterung des Netzwerkmodells	107
6.2.2	Erweiterung des mathematischen Modells	113
6.2.3	Verschiebbare Fahrten	115
6.2.4	Testergebnisse	118
6.2.5	Überlegungen zu der Anschlussfahrten-Problematik	129
6.3	Das Depotgruppen-Konzept	129
6.3.1	Einsatzmöglichkeit	130
6.3.2	Modellierung	131
6.3.3	Ergebnisse	134
6.4	Weitere praktische Aspekte	138
6.4.1	Kapazitäten	138
6.4.2	Mindestaufenthaltszeit im Depot	139
6.4.3	Tageszeitabhängige Verbindungsdauer und -Entfernung	140
6.4.4	Vergabe von Fahrten	141

6.4.5	Vorbereitungs- und Wendezeiten	141
7	Maßnahmen zur Kostensenkung in der Umlaufplanung	143
7.1	Untersuchte Instanzen	143
7.2	Linienbündelung	144
7.2.1	Tests an ausgewählten Instanzen	146
7.3	Fahrzeugtypgruppen	149
7.3.1	Testergebnisse	151
7.3.2	Zusammenfassung	154
7.4	Mehrdepotplanung	154
7.4.1	Klassische Mehrdepotplanung	155
7.4.2	Depotgruppen für gleitende Fahrzeug-Depot-Zuordnung . . .	159
7.5	Zusammenfassung	163
8	Lösungsansätze für sehr große Probleminstanzen	165
8.1	Problemdekomposition im Preprocessing	165
8.2	Lambda-gesteuerte Reduktion der Verbindungskanten	167
8.3	Sequenzielles Scheduling und Clustering	168
8.4	Fahrtenketten-Heuristik: Fix-and-Optimize Methode	169
8.4.1	Phase 1	169
8.4.2	Phase 2	169
8.4.3	Numerische Ergebnisse	173
8.4.4	Automatische Erstellung vereinfachter Probleme	185
9	Zusammenfassung und Ausblick	187

Abbildungsverzeichnis

2.1	Planungsprozess im ÖPNV	8
2.2	Ausschnitt aus einem Liniennetz von PaderSprinter, Paderborn . . .	10
2.3	Typische Fahrtenmengen eines Betriebstages	12
3.1	Ein Netzwerk-Fluss-Modell für das MDVSP	34
3.2	Transformation eines SDVSP zu einem minimale-Kosten-Flussproblem in Anlehnung an [Löbel and Strubbe, 1996]	35
3.3	Mehrgüter-Fluss-Modell nach [Löbel and Strubbe, 1996]	36
3.4	Reduziertes Mehrgüter-Fluss-Modell nach [Grötschel et al., 1997] . .	37
3.5	Netzwerk nach [Banihashemi and Haghani, 2000]	38
3.6	Graph nach dem Modellansatz von [Dell’Amico et al., 1993]	41
3.7	Graph nach dem Modellansatz von [Fischetti et al., 1999]	42
3.8	Graph nach der Modellierung von [Freling et al., 2001]	44
3.9	Ablaufdiagramm für die Lösung des MDVSP nach [Löbel, 1997] . .	54
4.1	Timelines mit Kanten für potenzielle Service-, Depotausrück- und Depoteinrückfahrten	65
4.2	Anschlussknoten als aggregierte Gruppen von Ankünften und un- mittelbar darauffolgenden Abfahrten	66
4.3	Eliminierung redundanter Verbindungsfahrt-Kanten	67
4.4	Netztopologie	74
4.5	Erstellen der Timelines und Einfügen der Servicefahrt-Kanten . . .	76
4.6	Matches-Reduktion	77
4.7	Einfügen der Depot-Kanten	78
4.8	Knotenbildung, Einfügen der Warte- und Zirkulationsfluss-Kanten .	79
4.9	Ein optimaler Fluss von drei Flusseinheiten	80
5.1	Dekomposition eines S-T-Flusses in Wege	88
5.2	Dekomposition eines Flusses nach FIFO/LIFO	90
5.3	Ergebnisvergleich der verschiedenen Dekompositionsstrategien . . .	97
6.1	Line Arcs in Time-Space Netzwerk	100
6.2	Einfügen der Linienkanten	100
6.3	Einfügen der Linienkanten	101

6.4	Ergebnisse für drei Instanzen	105
6.5	Potential von Zeitfenstern	106
6.6	Verwendung von Zeitfenster-Kanten	107
6.7	Netzwerkaufbau mit Zeitfenstern (Schritt 1)	108
6.8	Netzwerk-Reduktion durch Löschen überflüssiger Zeitfenster-Kanten	110
6.9	Reduktion von Depot-Kanten für Zeitfenster	111
6.10	Netzwerk nach Reduktion überflüssiger Zeitfenster-Kanten	111
6.11	Optimaler Fluss im Netzwerk mit Zeitfenster-Kanten	112
6.12	Linien-Kanten und Zeitfenster-Kanten: Normalfall	113
6.13	Linien-Kanten und Zeitfenster-Kanten: Spezialfall	113
6.14	Ablauf der Verkürzungsheuristik	116
6.15	Fahrtenauslastung eines Fahrplans	117
6.16	Ermittlung kritischer Fahrten in der Schnitt-Heuristik	118
6.17	Glättung der Auslastungskurve durch die Schnitt-Heuristik	119
6.18	Fahrtensmengen der Instanzen über die Dauer eines Betriebstages . .	120
6.19	Einsparung und Laufzeit für globale Zeitfenster bis zu fünf Minuten	125
6.20	Einsparung und Laufzeit für große Zeitfenster für Monheim	125
6.21	Vergleich von Lösungsqualität zu Laufzeit für Halle	127
6.22	Vergleich von Lösungsqualität zur Gesamtverschiebung für Halle . .	127
6.23	Nur-Eigen-Depot A und Nur-Eigen-Depot B	131
6.24	Depot A mit Pausen in B und Rückkehr zu B	131
7.1	Time-Space Netzwerk für das Beispiel	145
7.2	Modellierung der Fahrzeugtypgruppen als Mengen	151
7.3	Gruppenbildung für BVO	153
7.4	BVO: Kostenentwicklung	157
7.5	BVO: Verteilung Depots	158
7.6	München: Variable Kosten bei unterschiedlicher Depotanzahl	159
7.7	München: Anzahl Fahrzeuge bei unterschiedlicher Depotanzahl . . .	160
8.1	Arbeitsweise der „fix-and-optimize“ Zwei-Phasen Heuristik	170
8.2	Basis-Algorithmus zur Fahrtenkettenerkennung	172
8.3	Laufzeiten versus Anzahl der Depots: city 1	177
8.4	Laufzeiten versus Anzahl der Depots: city 2	178
8.5	Laufzeiten versus Anzahl der Depots: city 3	179
8.6	Ergebnisse der Fahrtenketten-Erkennung für Fahrpläne city 1 bis 3 und mun 1	184

Tabellenverzeichnis

3.1	Aufgabe der Umlaufbildung	33
4.1	Fahrplan mit Servicefahrten	74
4.2	Eigenschaften der Probleminstanzen	83
4.3	Eigenschaften der Netzwerkmodelle	83
4.4	Exakte Optimierung: Ergebnisse	83
4.5	Eigenschaften der optimalen Lösungen	84
4.6	Laufzeiten-Vergleich: IPM (Barrier) versus dualer Simplex	85
4.7	Laufzeiten für die Minimierung der Fahrzeuganzahl	85
5.1	Laufzeitvergleich der verschiedenen Dekompositionsstrategien	96
6.1	Modellgröße und Laufzeit mit Linienkanten	103
6.2	Vergleich der Lösungen mit und ohne Linienkanten (LK)	103
6.3	Servicefahrt mit Zeitfenstern	107
6.4	Daten der verwendeten Instanzen	119
6.5	Ergebnisse bei globalen Zeitfenstern	122
6.6	Ergebnisse der Verkürzungs-Heuristik	123
6.7	Ergebnisse der Schnitt-Heuristik mit $\alpha = 10$ und $\beta = 10$	124
6.8	Fahrplandaten	135
6.9	2-Depots-Instanzen ohne und mit Depotgruppen	135
6.10	Drei Depots ohne und mit Depotgruppen	136
6.11	Verteilung der „depot-eigenen“ und „depot-gemischten“ Umläufe	137
6.12	Hinzufügen von Pausendepots (Fahrplan 2)	138
7.1	Eigenschaften untersuchter Testfälle	144
7.2	Beispielfahrplan	145
7.3	Beispiel - linienrein vs. liniengemischt	146
7.4	Halle linienrein vs. liniengemischt	146
7.5	Halle linienrein vs. liniengemischt, detailliert	146
7.6	München linienrein vs. liniengemischt	147
7.7	München linienrein vs. liniengemischt, detailliert	147
7.8	München linienrein vs. liniengemischt, Ein-/Ausrückfahrten	147
7.9	BVO linienrein vs. liniengemischt	149

7.10	BVO linienrein vs. liniengemischt, detailliert	149
7.11	BVO linienrein vs. liniengemischt, dicht	150
7.12	Halle: feste Vorgabe	152
7.13	Halle: alle Fahrzeugtypen in einer Gruppe	152
7.14	Beispiel: zwei Depots	155
7.15	Halle: Eindepot- versus Mehrdepot-Umlaufplanung	156
7.16	Halle: Eindepot- versus Mehrdepot-Umlaufplanung detailliert	156
7.17	BVO: Eindepot-Umlaufplanung	157
7.18	Halle: gleitende Depotzuordnung	161
7.19	Halle: gleitende Depotzuordnung, detailliert	161
7.20	BVO: gleitende Depotzuordnung	161
7.21	BVO: gleitende Depotzuordnung, detailliert	162
7.22	München: gleitende Depotzuordnung	162
7.23	München: gleitende Depotzuordnung, detailliert	162
8.1	Änderung von λ	168
8.2	Zuordnungsfunktion für Nachfolger-Servicefahrten	171
8.3	Eigenschaften untersuchter Testfälle	174
8.4	Ergebnisse der exakten Optimierung	175
8.5	Ergebnisse der Zwei-Phasen-Heuristik: city 1	178
8.6	Ergebnisse der Zwei-Phasen-Heuristik: city 2	179
8.7	Ergebnisse der Zwei-Phasen-Heuristik: city 3	180
8.8	Ergebnisse der Zwei-Phasen-Heuristik: mun 1	181
8.9	Anzahl der erkannten Fahrtenketten	183

Abkürzungsverzeichnis

B&B	Branch and Bound Methode
B&C	Branch and Cut Methode
DSS	Decision Support System
FIFO	First In – First Out
IPM	Interior Point Method
LIFO	Last In – First Out
LP	Lineares Programm
MDVSP	Multi-Depot Vehicle Scheduling Problem
MDVSPTW	Multi-Depot Vehicle Scheduling Problem with Time Windows
MinAlternation	Minimal Line Alternation
MIP	Mixed Integer Program
MVG	Münchener Verkehrsgesellschaft
ÖPNV	Öffentlicher Personennahverkehr
OR	Operations Research
SPP	Set Partitioning Problem
SDVSP	Single-Depot Vehicle Scheduling Problem
TSN	Time-Space Netzwerk
XMinAlternation	eXtended Minimal Line Alternation

Kapitel 1

Einleitung und Motivation

Das wachsende Bedürfnis nach Mobilität ist eine der ausschlaggebenden Determinanten moderner Gesellschaften. Mobil sein bedeutet, den Anforderungen der modernen Gesellschaft durch eine zeitliche und räumliche Flexibilität gerecht zu werden. Die Forderung nach höherer Mobilität trifft aufgrund der notwendigen Steigerung der Verkehrsdichte auf die begründete Sorge um die schädlichen Einflüsse des Verkehrsaufkommens auf die Umwelt. Dabei stellt sich die Frage, mit welchen Verkehrsformen der gestiegene Personenbeförderungsbedarf in der Zukunft befriedigt werden soll.

Der Rat für Nachhaltige Entwicklung beim Wissenschaftszentrum für Sozialforschung legt in einem Dialogpapier Ziele zur Nachhaltigen Entwicklung in Deutschland fest. Angesichts der steigenden Verkehrsdichte heißt es in diesem Papier: „Ein nachhaltiges Verkehrssystem muss die Wirtschaft unterstützen und negative ökologische Folgen vermeiden und soziale positive Effekte fördern. Es soll sozial gerecht, ökologisch verträglich und ökonomisch effizient organisiert werden“ ([Nachhaltigkeitsrat, 2004]).

Auch die fortschreitende Privatisierung des Marktes für die Personenbeförderung und die Strukturreform des öffentlichen Verkehrs in Deutschland bezwecken eine höhere ökonomische Effizienz der Personenbeförderung durch Etablierung eines stärkeren Wettbewerbs in diesem Bereich. Dafür werden die Leistungsbeziehungen für den Nahverkehr in Deutschland zwischen Aufgabenträger (Kreise und Städte) und Verkehrsunternehmen nach dem sogenannten „Besteller-Ersteller-Prinzip“ aufgeteilt. Eine klare Aufgabentrennung soll mehr Effizienz im öffentlichen Personennahverkehr (ÖPNV) fördern.

Den Aufgabenträgern für den ÖPNV obliegt es, die Weiterentwicklung und Bestellung des Nahverkehrs vor Ort sicherzustellen. Sie haben die Aufgabe, den straßengebundenen Nahverkehr im Wettbewerb zu organisieren und die jeweilige

Finanzierung des vor Ort für nötig befundenen gemeinwirtschaftlichen Verkehrsangebotes sicherzustellen. Dafür werden von den Aufgabenträgern einzelne Linien oder ganze Linienbündel ausgeschrieben. Die Verkehrsunternehmen können eine Konzession auf bestimmte Linien erwirken, indem sie sich auf die entsprechende Ausschreibung bewerben. Das günstigste Angebot erhält dabei den Zuschlag. Die Konzession wird aufgrund des Verkehrsvertrages zwischen Aufgabenträgern und Aufgabennehmern (Verkehrsunternehmen) geregelt. Auf diese Weise wird im Prozess der Deregulierung des Marktes für den öffentlichen Personennahverkehr eine Vermischung von unternehmerischer und öffentlicher Verantwortung abgeschafft. Durch die regelmäßigen Ausschreibungen der Verkehrsleistungen wird von den Verkehrsbetrieben eine höhere ökonomische Effizienz gefordert.

Die neue Struktur begrenzt die in der Vergangenheit möglichen Querfinanzierungen, Subventionierungen und Ausgleichszahlungen. Zum Teil werden zwar noch die alten „Schlupflöcher“ von den Unternehmen ausgenutzt, die aktuelle Entwicklung der Rechtssituation macht es aber auf Dauer unvermeidlich, Ausschreibungen zur Vergabe öffentlicher Aufträge für die Verkehrsleistungen durchzuführen. Die in der Vergangenheit übliche Quersubventionierung vom Energie- zum Verkehrsbereich sowie nachträgliche Verlustausgleichszahlungen von Kommunen an die Verkehrsunternehmen werden dann nicht mehr möglich sein. Oft wird diese neue Struktur des ÖPNV insbesondere von den kleineren Verkehrsunternehmen als Bedrohung empfunden. Im Bezug auf diese Entwicklung des ÖPNV meinte Dr. Axel Horstmann, Minister für Verkehr, Energie und Landesplanung des Landes Nordrhein-Westfalen [[Landkreistag-NRW, 2004](#)]: „Nicht die Großen fressen die Kleinen, sondern die Schnellen die Langsamen“. Man solle die aktuelle Entwicklung der Rechtssituation um den ÖPNV nicht als eine Bedrohung sehen, sondern als eine Chance für die Zukunft im europäischen Wettbewerb.

Für die Verkehrsunternehmen im ÖPNV bedeutet das in erster Linie, dass sie sich den neuen Anforderungen des Marktes stellen müssen, um im sich zuspitzenden Wettbewerb bestehen zu können. Schnelles und effizientes Handeln ist dabei unabdingbar. Um sich erfolgreich auf eine Ausschreibung zu bewerben, muss ein Verkehrsunternehmen das günstigste Angebot vorlegen. Der darin enthaltene Angebotspreis (vorgegeben in der Regel durch Kilometerkosten) muss aber für das Verkehrsunternehmen selbst kostendeckend sein. Einer der größten Kostenblöcke eines Unternehmens sind neben den Personalkosten die Kosten des Fahrzeugeinsatzes. Die Gesamtkosten des Fahrzeugeinsatzes bestehen aus den fixen Kosten für gekaufte oder gemietete Fahrzeuge, und aus den operativen Kosten, die durch die Ausführung von Fahrtleistungen anfallen. Die fixen Fahrzeugkosten ergeben sich aus Anschaffungs- und Wartungskosten der zur Bedienung der Linien benötigten Fahrzeuge. Die operativen Kosten werden in eine zeitabhängige und eine entfernungsabhängige Kostenkomponente aufgeschlüsselt. Zu beachten ist dabei,

dass die Anschaffung und der Betrieb von Fahrzeugen verschiedener Typen unterschiedlich viel kostet. Durch eine hohe Auslastung der Fahrzeuge und durch den Einsatz des jeweils günstigsten Fahrzeugtyps für jede ausgeschriebene Linienfahrt können die für das Unternehmen anfallenden Kosten stark reduziert werden. Dies würde einen günstigen Kilometerpreis in dem vom Unternehmen vorzulegenden Angebot ermöglichen und auf diese Weise im Wettbewerb mit der Konkurrenz um die Ausschreibung zu besseren Chancen führen. Um ein gutes Angebot bei einer Ausschreibung vorlegen zu können, muss die Kostenkalkulation somit auf einem effizienten Einsatz der Fahrzeuge und der Fahrer basieren.

Diese Arbeit legt den Fokus auf die effiziente Planung des Fahrzeugeinsatzes. Im Rahmen der Fahrzeugumlaufplanung, die auf einem fest vorgegebenen Fahrplan basiert, werden alle Fahrten eines Tages den Fahrzeugen zugewiesen. Es entsteht ein Umlaufplan, dessen Leerfahrten und Standzeiten die operativen Kosten verursachen. Diese operativen Kosten gilt es, zusammen mit den fixen Kosten für benötigte Fahrzeuge, zu minimieren.

Die Umlaufplanung wird im Rahmen eines komplexen, in mehrere Schritte unterteilten Produktionsplanungs- und Steuerungsprozesses im ÖPNV durchgeführt. Computergestützte Entscheidungsunterstützungssysteme spielen dabei eine entscheidende Rolle. Die Software-Produkte, die die Komplexität der Planungsaufgaben handhabbar halten, sind aus der Praxis der Ressourceneinsatzplanung in einem modernen ÖPNV-Betrieb nicht mehr wegzudenken. Darüber hinaus kann die Handlungsvorbereitung und die Entscheidungsfindung mit Hilfe von Methoden aus dem Bereich des Operations Research unterstützt werden. Solche Methoden, die zunehmend in die Planungswerkzeuge integriert werden, basieren auf Abbildungen der realen Entscheidungssituation in Form von mathematischen Modellen oder Netzwerkstrukturen.

Für die Optimierung des Fahrzeugeinsatzes wurden verschiedene Modelle und Verfahren in der Literatur vorgeschlagen. Die Verfahren aus der Literatur, die leistungsfähig genug sind, um Instanzen aus der Praxis kostenoptimal zu lösen, basieren auf unterschiedlichen Varianten der Column Generation Methode und weiteren fortgeschrittenen Techniken. Notwendig sind diese Techniken, weil die verwendeten Modelle eine sehr hohe Anzahl von Entscheidungsvariablen beinhalten. Der Einsatz spezialisierter Verfahren führt aber in der Regel zu einer Einschränkung der Aspekte, die im Modell berücksichtigt werden können.

Die Zielsetzung der vorliegenden Arbeit besteht in der Konzeption und Implementierung eines neuen Modells, das die Schwachstellen der existierenden Ansätze vermeidet und dadurch sehr große reale Problemstellungen mit zusätzlichen praktischen Anforderungen und Rahmenbedingungen in akzeptabler Zeit ohne Einsatz spezialisierter Algorithmen lösbar macht. Dies soll die Lücke zwischen den Anforde-

rungen aus der Praxis der Umlaufplanung und den Möglichkeiten der existierenden theoretischen Ansätze schließen.

Daraus lassen sich folgende Teilziele der vorliegenden Arbeit ableiten:

- Die Möglichkeit einer Modellierung und Lösung des Mehrdepot-Umlaufplanungsproblems mit Hilfe eines Time-Space-Netzwerkes soll untersucht werden.
- Die Anforderungen aus der Praxis sollen in dem zu erstellenden neuen Modell abbildbar sein, um die praktische Anwendbarkeit zu sichern.
- Es sollen Lösungsverfahren entwickelt werden, die eine Lösung auch für extrem große Probleminstanzen, insbesondere für Großstädte in annehmbarer Zeit ermöglichen.
- Die Modelle und die Lösungsverfahren sollen in ein Software-System implementiert werden, um die Möglichkeit des praktischen Einsatzes durch Testläufe belegen zu können.

Der beschriebenen Zielsetzung wird mit einem entsprechenden Aufbau der vorliegenden schriftlichen Ausarbeitung Rechnung getragen. Die Arbeit gliedert sich in neun Kapitel, wobei die Kapitel vier bis acht den inhaltlichen Kern der Arbeit ausmachen.

Kapitel 2 gibt einen Einblick in die Praxis der ÖPNV-Planung und beschreibt die einzelnen Planungsaufgaben mit einem Fokus auf die Fahrzeugumlaufplanung.

Ein Überblick über bereits veröffentlichte Modellierungs- und Lösungsansätze für das Mehrdepot-Umlaufplanungsproblem wird in Kapitel 3 gegeben. Dabei werden insbesondere die unterschiedlichen Modellierungsansätze kategorisiert und beschrieben.

Im Rahmen dieser Arbeit erfolgt die Modellierung mit Hilfe eines Time-Space-Netzwerkes. Diese Modellierungstechnik wird in Kapitel 4 vorgestellt und an mehreren Probleminstanzen aus der Praxis getestet.

In Kapitel 5 werden unterschiedliche Dekompositionsstrategien entwickelt. Diese werden dazu benötigt, aus einer optimalen Lösung des Netzwerkflussproblems aus Kapitel 4 die Fahrzeugumläufe zu extrahieren. Es wird untersucht, inwieweit sich durch unterschiedliche Dekompositionsstrategien praktische Anforderungen berücksichtigen lassen.

Einen weiteren Schwerpunkt der Arbeit bilden die in Kapitel 6 vorgestellten Möglichkeiten der Berücksichtigung verschiedener praktischer Aspekte im Aufbau des Netzwerkmodells und in Erweiterungen des Verfahrens.

In Kapitel 7 wird das Optimierungspotential untersucht, das sich durch eine Ausnutzung der Freiheitsgrade in der Ressourceneinsatzplanung, insbesondere in der Planung der Umläufe ergibt.

Kapitel 8 beinhaltet verschiedene Techniken, mit deren Hilfe Lösungszeiten reduziert werden, bzw. extrem große Probleminstanzen lösbar gemacht werden sollen, die bisher mit keinem existierenden exakten Verfahren lösbar waren. Dafür werden in diesem Kapitel approximative Modelle und Verfahren konzipiert, deren Leistungsfähigkeit durch Testergebnisse belegt wird.

Die Integration der vorgestellten Methoden in die kommerziellen Systeme zur Entscheidungsunterstützung in ÖPNV-Betrieben ist Gegenstand der Ausführungen in Kapitel 9, das auch eine Zusammenfassung der Ergebnisse und einen Ausblick auf weitere Forschungsaktivitäten enthält.

Kapitel 2

Der Planungsprozess im ÖPNV

Dieses Kapitel behandelt die Umlaufplanung als einen Teil der Ressourceneinsatzplanung im ÖPNV. Um diese Aufgabe in den Gesamtzusammenhang des Planungsprozesses einzuordnen, werden auch andere Planungsaufgaben beschrieben. Beleuchtet werden außerdem verschiedene Aspekte der Umlaufplanung, wie die Bewertung konkreter Umlaufpläne anhand unterschiedlicher Kennzahlen und die Modalitäten, die sich aus der Vielfalt der betrieblichen Randbedingungen ergeben. Daraus werden die Anforderungen abgeleitet, die an ein Modell zur Abbildung und Lösung der Umlaufplanungsaufgabe gestellt werden sollen.

Unter einem ÖPNV-Unternehmen werden im Rahmen dieser Arbeit Verkehrsunternehmen verstanden, die Fahrgastbeförderung auf bestimmten Linien anbieten. Weiterhin wird davon ausgegangen, dass die Linienfahrten in einem bestimmten Zeitraster angeboten werden. Eine Linie wird also nicht nur mit einer bestimmten Frequenz bedient, wie es in anderen Verkehrsformen oder anderen Ländern üblich ist, sondern auch zu den vorher festgelegten Zeitpunkten.

2.1 Strategische und operative Planung

Der Planungsprozess im öffentlichen Personennahverkehr lässt sich grundlegend in strategische und operative Planung unterteilen (siehe Abb. 2.1). Die Planung ist aufgrund zahlreicher Anforderungen, die es zu berücksichtigen gilt, ein komplexer Prozess. Das Ziel dieser Planung ist eine möglichst kostengünstige und trotzdem hochqualitative Bedienung der Reisewünsche von Personen im Nahverkehr.

Die zahlreichen Aufgaben der Planung, Organisation und Ausgestaltung des ÖPNV beinhalten insbesondere die strategische Bestellerfunktion - Rahmenregelungen, Vorgaben und Qualitätsziele. Nach einer Bedarfsermittlung werden dabei im Rahmen der strategischen ÖPNV-Planung die Passagierströme geschätzt und

	Planungsaufgabe		Ergebnis
strategisch	Bedarfsermittlung	↓	Quell/Ziel-Daten
	Angebotsplanung		Systeme, Tarife, Verfügbarkeit
	Netzplanung		Infrastruktur
operativ	Linienplanung		Linienverlauf, Fahrzeitprofile
	Fahrplanerstellung		Fahrplan mit Fahrgastfahrten
	Fahrzeugumlaufplanung		Fahrzeugumläufe
	Tagesdienstplanung		Tagesschichten
	Dienstreihenfolgeplanung		Dienstplan

Abbildung 2.1: Planungsprozess im ÖPNV

eine grobe Angebotsplanung vorgenommen. Die anschließende Netzplanung legt die Infrastruktur des Nahverkehrsnetzes fest. In der ersten Stufe der operativen Planung werden die zu bedienenden Linien festgelegt, die anschließend im Rahmen der Fahrplanerstellung mit genauen Zeitangaben für jede Linienfahrt versehen werden. Diese Aufgaben sowie die Steuerung des in Kapitel 1 erwähnten Vergabeverfahrens für den ÖPNV liegen in der kommunalen Hand.

Nach einer Ausschreibung können die Verkehrsunternehmen eine Konzession auf bestimmte Linien oder Linienbündel erwirken. Das Ziel der Linienbündelung in einer Ausschreibung ist es, dem Verkehrsunternehmen Freiraum für einen effizienten Ressourceneinsatz zu lassen. Daher enthält eine „gerechte“ Ausschreibung nicht eine einzelne Linie, sondern eine möglichst gute Mischung aus verschiedenen Linien.

Der Planer in dem Verkehrsunternehmen steht einer schwierigen Aufgabe gegenüber. Die von ihm erstellten Pläne müssen in sich schlüssig, also fahrbar sein, und dabei einen möglichst effizienten Ressourceneinsatz ermöglichen. Die Fahrzeugumläufe und Fahrdienste sind dabei so zu planen, dass alle Linienfahrten des Unternehmens mit möglichst geringen Fahrzeug- und Personalkosten bedient werden.

Verkehrsunternehmen müssen also im Rahmen des Planungsprozesses den genauen Streckenverlauf ihrer Fahrzeuge sowie deren Fahrpläne festlegen und die Arbeitspläne für die benötigten Fahrer erstellen. Bei der Planung dürfen jedoch nicht

nur die Kostenaspekte eine Rolle spielen. Es gibt viele zu beachtende Nebenbedingungen. Zu nennen sind unter anderem gesetzliche sowie tarifvertragliche Regelungen, die die Länge der Pausen oder der Arbeitszeiten der Fahrer festsetzen und die technischen Bedingungen (beispielsweise Wartungsintervalle oder Geschwindigkeit der Transportmittel).

Um eine gesamtkostenoptimale ÖPNV-Planung unter Beachtung aller Nebenbedingungen zu erstellen, wäre es nötig, alle Teilaufgaben des Planungsprozesses, wie Linien-, Fahrplan-, Umlauf- und Dienstplanung simultan zu lösen. Da dies aber auf Grund der komplexen Datenlage nicht möglich ist, wird der Planungsprozess in kleinere Teilprobleme zerlegt, die dann sequentiell gelöst werden (siehe Abb. 2.1). In diesem Prozess kann allerdings eine Wiederholung eines vorherigen Planungsschrittes nötig werden. Dies geschieht dann, wenn das Ergebnis von einem Teilprozess in Bezug auf die Gesamtplanung nicht den Wünschen entspricht, oder gar eine Lösung des nachfolgenden Teilproblems hierdurch unmöglich wird (vgl. [Hagemeyer, 2000]). Wird z.B. durch die Planung der zu bedienenden Linien eine zu grosse Anzahl an Fahrzeugen benötigt, so kann es sinnvoll sein, die Linienplanung zu revidieren, so dass weniger Fahrzeuge ausreichen würden. Zwar ist dann die Lösung des Linienplanungsproblems nicht mehr optimal im Sinne des für diese Aufgabe ursprünglich festgelegten Optimalitätskriteriums, aber die damit entstandenen Kosten können durch Einsparungen bei der Fahrzeugeinsatzplanung kompensiert werden.

Eine mögliche grobe Zerlegung des gesamten operativen Planung beinhaltet also die Teilaufgaben Linienplanung, Fahrplanerstellung sowie Fahrzeugumlauf- und Dienstplanung. (Letztere wird in der Regel in zwei Stufen durchgeführt - zunächst werden die Tagesschichten gebildet und erst in der nachfolgenden Dienstreihenfolgeplanung der eigentliche Dienstplan.) In der Praxis wird üblicherweise hierarchisch in der genannten Reihenfolge geplant (vgl. [Borndörfer et al., 2002]).

2.2 Planungsschritte der operativen Planung

Als Schwerpunkt der operativen Planung wird in dieser Arbeit die Umlaufbildungsaufgabe angesehen. Diese Planungsaufgabe ist natürlich immer im Kontext der gesamten operativen Planung zu sehen, da die einzelnen Planungsschritte aufeinander aufbauen und somit nicht losgelöst voneinander betrachtet werden können. Nachfolgend werden in diesem Zusammenhang die operativen Planungsschritte vorgestellt.

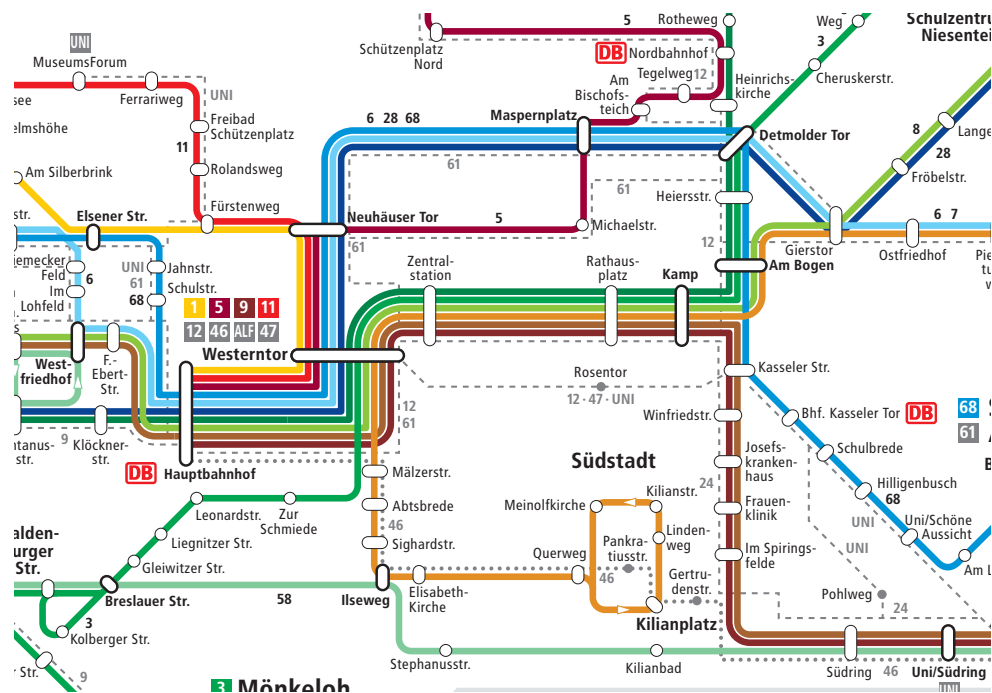


Abbildung 2.2: Ausschnitt aus einem Liniennetz von PaderSprinter, Paderborn

2.2.1 Linienplanung

In Rahmen der Linienplanung werden die Grundrouten und Fahrzeitprofile der zu bedienenden Linien bestimmt. Das Ziel dieses Planungsschrittes ist es, mit einer begrenzten Anzahl von Linien ein Maximum an Direktverbindungen (d.h. Verbindungen ohne Umsteigenotwendigkeit) anzubieten.

Unter dem Begriff *Linie* wird eine Folge von Haltestellen bzw. Wegpunkten verstanden, die nacheinander bedient werden sollen. Unter dem Begriff *Liniennetz* wird die Gesamtheit der erzeugten Linien mit ihren Haltestellen, Pausenräumen und Depots verstanden. In einem *Liniennetz* werden alle Linien erfasst, die zu bedienen sind. Die Abbildung 2.2 beinhaltet einen Ausschnitt des Liniennetzes der PaderSprinter in Paderborn.

Ausgangspunkt für diesen Planungsschritt sind die Ergebnisse der strategischen Planung, insbesondere die Schätzungen über die Nachfrage an benötigten Verkehrsverbindungen. Basierend auf dem geschätzten Beförderungsbedarf und auf einem vorgegebenen Streckennetz wird der Verlauf einzelner Linien festgelegt. Die Anzahl der Linien ist in der Regel begrenzt. Sie sollen so geplant werden, dass möglichst viele Direktverbindungen zustande kommen. Falls erforderlich, wird für jede Linie eine *Fahrzeugtypgruppe* bestimmt, die alle für die Fahrten dieser Linie geeignete Fahrzeugtypen enthält. Die Fahrzeitprofile beschreiben den Zeitbedarf für die ein-

zelenen Fahrabschnitte einer Linie.

Die Aufgabe der Linienplanung kann mit Optimierungsmethoden unterstützt werden. Einige Modellierungsansätze findet man z.B. in [Bordörfer et al., 2004] und speziell zu der Planung der Linienverläufe im schienengebundenen Verkehr in [Bussieck and Zimmermann, 1997].

2.2.2 Fahrplanerstellung

Durch die Festlegung des Linienverlaufs ist zwar der Zeitbedarf für die einzelnen Linien festgelegt, die Taktfrequenz der Linie und die Abfahrtszeitpunkte für jede Linienfahrt müssen jedoch noch bestimmt werden. Dies geschieht im Rahmen der Fahrplanerstellung. Die möglichen Wartezeiten der Fahrgäste beim Umsteigen werden dabei minimiert, so dass möglichst nur „verträgliche“ (also weder zu lange noch zu kurze) Wartezeiten entstehen. An dieser Stelle im Planungsprozess kommen wieder die zugrunde gelegten Daten der geschätzten Verkehrsverbindungen zum Tragen. In dieser Planungsaufgabe wird beispielsweise darauf geachtet, dass es bei Straßenbahnen nicht möglich ist, ein Gleis zur gleichen Zeit doppelt zu nutzen. Bei der Festlegung der Abfahrtszeiten müssen auch solche Nebenbedingungen zusätzlich beachtet werden.

Bei der Fahrplanerstellung wird also die räumliche Struktur eines Liniennetzes mit der zeitlichen Struktur des Fahrplans überlagert. Dabei werden Linien in Personenbeförderungsfahrten zwischen zwei Endhalttestellen zerlegt. Das Ziel der Fahrplanerstellung ist die Vermeidung bzw. Minimierung der gegenseitigen Behinderung von Fahrplanfahrten auf gemeinsam benutzten Straßen und eine optimale Ergänzung der Fahrten im Gesamtverkehrssystem. In dieser Planungsphase werden auch die Mengen der zulässigen Fahrzeugtypen für die einzelnen Fahrten festgelegt, falls sie nicht für alle Fahrten einer Linie gleich sein sollen.

Einige typische Verteilungen von Servicefahrten über einen Betriebstag sind in der Abbildung 2.3 durch die Tagesganglinien visualisiert. Sie zeigen die Anzahl der Fahrten, die zu einem bestimmten Zeitpunkt laut jeweiligem Fahrplan gefahren werden müssen. Oft haben solche Tagesganglinien zwei bis drei Spitzen, die einen erhöhten Beförderungsbedarf wegen Schülerfahrten oder Pendlerverkehr in den frühen Morgen- oder späten Nachmittagsstunden darstellen.

2.2.3 Fahrzeugumlaufplanung

In der Fahrzeugumlaufplanung (auch als Umlaufplanung bezeichnet) wird entschieden, welche Fahrzeugtypen bzw. Fahrzeuge auf den zuvor definierten Linien verkehren. In dieser Teilaufgabe der Planung werden die vom Fahrplan vorgegebenen

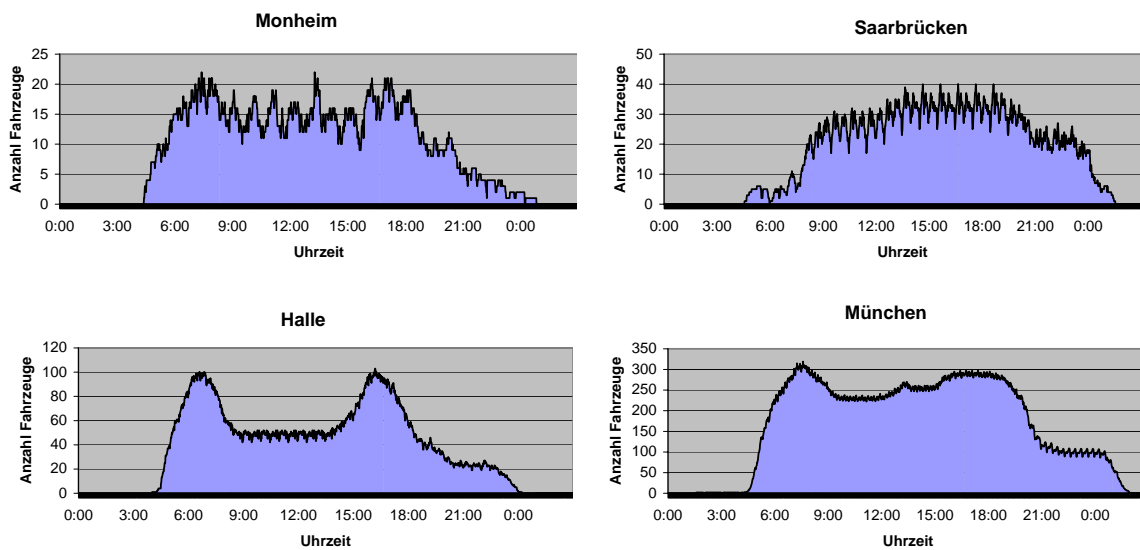


Abbildung 2.3: Typische Fahrtenmengen eines Betriebstages

Personenbeförderungsfahrten den vorhandenen Fahrzeugen zugeordnet. Bei der Zuordnung kann ein oder mehrere Ziele angestrebt werden. Es kann eine Minimierung der Anzahl einzusetzender Fahrzeuge oder die Minimierung der Leerfahrtzeiten und/oder die Einhaltung bestimmter Vorgaben, wie die Begrenzung der maximalen Anzahl der Linienwechsel pro Umlauf, sein.

Bei der Fahrzeugumlaufplanung unterscheidet man daher grundsätzlich zwischen dem regelbasierten und dem zielbasierten Vorgehen. Ziel der *regelbasierten* Planung der Umläufe ist es, bestimmte vordefinierten Regeln einzuhalten, um somit die Gültigkeit und die Qualität des resultierenden Umlaufplanes sicher zu stellen. Dieses Verfahren ist dank seiner Einfachheit noch heute sehr verbreitet. Bei dem *zielbasierten* Vorgehen wird anstatt beziehungsweise zusätzlich zu der Einhaltung von Regeln die bestmögliche Erfüllung eines oder mehrerer Ziele angestrebt. Letzgenannte Verfahren sind komplexer und ohne Computerunterstützung in der Regel nicht realisierbar.

Die zielbasierte Fahrzeugumlaufplanung mit dem Ziel der Minimierung der Kosten ist der Hauptuntersuchungsgegenstand dieser Arbeit und wird in Abschnitt 2.3 ausführlich beschrieben.

2.2.4 Dienstplanung

In diesem Planungsschritt geht es um die Planung des Personaleinsatzes zur Bedienung der im Rahmen der Umlaufplanung festgelegter Fahrzeugumläufe. Dabei werden mehrere Teilaufgaben, insbesondere Tagesdienstbildung (engl.: *crew sche-*

duling) und Dienstreihenfolgeplanung (engl.: *crew rostering*), sequentiell durchgeführt. Die Bildung der Tagesdienste liefert eine Menge anonymer Tagesdienste, die in der Dienstreihenfolgeplanung zu Wochendiensten zusammengefasst werden. Eine wichtige Nebenbedingung ist dabei die Einhaltung tarifvertraglicher Randbedingungen. So müssen beispielsweise die maximale Einsatzdauer ohne Unterbrechung und die vorgegebene Pausendauer eingehalten werden. Im Rahmen der Dienstplanung werden zunächst die aus der Fahrzeugeinsatzplanung erhaltenen Fahrzeugumläufe jedes Betriebstages in anonyme fahrerunabhängige Dienststücke zerlegt und anschließend so zusammengefügt, dass jedes Umlaufstück in einem Tagesdienst enthalten ist. Diese Dienste werden bei der Dienstreihenfolgeplanung verkettet, so dass sie später einzelnen Fahrern zugewiesen werden können. Nachfolgend werden diese beiden Teilaufgaben der Dienstplanung beschrieben.

Tagesdienstplanung

Die Dienstaktivitäten werden durch ein manuelles oder automatisches Schneiden der Fahrzeugumläufe in Teilumläufe gebildet. Aus diesen werden anschließend Einzeldienste und Dienstklassen grafisch-interaktiv oder automatisch generiert.

Beim Schneiden von Diensten oder Dienststücken aus der unverplanten Dienstplanmasse (*Gesamtheit der Aktivitäten des Umlaufplans*) müssen unterschiedliche Randbedingungen berücksichtigt werden. Dies sind beispielsweise die minimale bzw. maximale Länge eines Dienstes oder Dienststückes, die Zeitintervalle für Dienstbeginn und -ende oder die Länge der entstehenden Reststücke, um so unverplanbare Reststücke zu vermeiden.

Die meisten Entscheidungsunterstützungssysteme beinhalten ein grafisch-interaktives Bilden von Dienststücken und Diensten mit anschließendem automatischen Kombinieren der Dienststücke zu zulässigen Diensten (s. [PTV, 2004]).

Beim Bilden von Dienstklassen, wie Früh-, Mittel- oder Spätdienste, muss, soweit möglich, eine vorgegebene Anzahl unterschiedlicher Dienste einer Klasse generiert und bewertet werden, die den vorgegebenen Rahmenbedingungen entsprechen. Falls ein Optimierungsverfahren eingesetzt wird, wird aus der Gesamtmenge möglicher Dienste eine Teilmenge ausgewählt, die alle Dienstaktivitäten abdeckt und bestimmte Zielsetzungen, wie die Kostenminimierung, am besten erfüllt. Diese Tagesdienste werden an die Dienstreihenfolgeplanung weitergegeben.

Dienstreihenfolgeplanung

Anschließend werden die Dienstpläne der verschiedenen Wochentage unter Berücksichtigung von Mindestruhezeiten (auch für Feiertagsübergänge), durchschnittli-

chen Wochenarbeitszeiten und vorgebbaren Turnusfolgen von Arbeits- bzw. arbeitsfreien Tagen erstellt. Dabei werden mehrere Tagesdienste zu Wochen-Dienstplänen zusammengefasst, so dass sie hintereinander ausgeführt werden können, ohne dass Bedingungen wie die minimale Ruhezeit zwischen zwei Diensten verletzt werden. Danach erfolgt schließlich die Zuweisung der Dienstpläne zu den einzelnen Fahrern. Erst dann wird die Anonymität der Dienstpläne aufgegeben.

Wie in der Umlaufplanung gibt es auch für die Dienstplanung Verfahren zur automatischen Verteilung der Dienste, die sämtliche festgelegte Mindestruhezeiten je Wochentagsübergang berücksichtigen. Wird in der Dienstplanung ein zielbasiertes Vorgehen verfolgt, dann gehen die Parameter Ruhezeit und durchschnittliche Arbeitszeit beliebig gewichtet in die Zielfunktion der Optimierung ein.

2.2.5 Reihenfolge der Planungsschritte

Als Ziel der Planung im Rahmen des öffentlichen Personennahverkehrs kann eine möglichst günstige Bedienung des Beförderungsbedarfs der Bevölkerung angesehen werden. Allerdings darf der Kostenpunkt dabei nicht allein betrachtet werden, da es viele andere sowohl gesetzliche als auch tarifvertragliche Rahmenbedingungen gibt. Dazu zählen die Arbeitszeitrestriktionen für Fahrer, die technischen Bedingungen für Fahrzeuge oder auch eine Vorschrift zur Bedienung unrentabler Verbindungen.

Eine gesamtoptimale Lösung des Planungsprozesses unter Beachtung aller Nebenbedingungen würde eine simultane Betrachtung aller Teilprozesse erfordern. Allerdings erweist sich die Aufgabe schon wegen der Komplexität der einzelnen Probleme als sehr schwierig. Für realistische Problemgrößen ist sie mit dem heutigen Stand der mathematischen Optimierung und Computertechnik noch nicht exakt gesamt optimal lösbar. Daher wird es hier von einer sequentiellen Planung, wie in der Abbildung 2.2 bereits dargestellt wurde, ausgegangen. In der sequentiellen Planung ließe sich prinzipiell die Reihenfolge der Planungsschritte „Fahrzeugeinsatzplanung“ und „Dienstschichtplanung“ auch in umgekehrter Reihenfolge abarbeiten. Die hier vorgestellte Reihenfolge ist jedoch die in der Praxis eingesetzte.

2.3 Umlaufbildung als Aufgabe der operativen Ressourceneinsatzplanung

Nachdem in den vorangegangenen Abschnitten der Kontext der operativen Ressourceneinsatzplanung beschrieben wurde, wird in diesem Abschnitt die Umlaufplanungsaufgabe näher betrachtet. Im Folgenden werden zunächst zentrale Begriffe definiert, die für das weitere Verständnis der Arbeit von Bedeutung sind. Diese

Begriffe werden später auch in der mathematischen Formulierung des Umlaufplanungsproblems und in den vorgestellten Algorithmen verwendet.

Depot: Ein Depot ist ein Ort für das Abstellen der Fahrzeuge. Die Fahrzeuge beginnen und enden ihren Tagesumlauf an diesem Ort - von hier aus startet also das Fahrzeug zur ersten Fahrt des Tages, und hierhin kehrt es auch von der letzten Fahrt des Tages zurück. Alle Fahrzeuge werden über Nacht im Depot geparkt. In der Realität hat jedes Depot nur eine begrenzte Anzahl von Stellplätzen, die für die Übernachtung (Stationierung) oder für das vorübergehende Abstellen während des Tages benutzt werden können. Derartige Begrenzungen werden im Weiteren Depotkapazitäten genannt.

Ein Verkehrsunternehmen verfügt in der Regel über mehrere Depots. Die entsprechende Planungsaufgabe wird in diesem Fall als Mehrdepot-Umlaufplanung gekennzeichnet.

Fahrzeugtyp und Fahrzeugtypgruppe: Die eingesetzten Fahrzeuge können unterschiedliche Eigenschaften aufweisen. Sie können sich in ihrer Größe, Geschwindigkeit oder Ausstattung unterscheiden.

Als eine Fahrzeugtypgruppe wird eine Teilmenge der verfügbaren Fahrzeugtypen bezeichnet, die bestimmte Eigenschaften gemeinsam haben. Ein Beispiel für eine Fahrzeugtypgruppe wäre eine Zusammenfassung aller behindertengerechten Fahrzeuge, wie Niederflrbusse in Normal- oder Gelenkbusausführung sowie Fahrzeuge mit Rampe oder anderen Hilfen zum Ein- und Aussteigen. Beispielsweise haben bei Buslinien, die an Alten- und Behindertenheimen vorbeifahren, einige Fahrgäste Schwierigkeiten beim Ein- und Aussteigen. In diesem Fall wird allen Fahrten der betroffenen Linien schon bei der Planung der Fahrzeugeumläufe die Fahrzeugtypgruppe „behindertengerechte Fahrzeuge“ zugewiesen. Auf diese Weise wird erzwungen, dass ausschließlich die geeigneten Fahrzeuge eingesetzt werden. Durch Vorgabe bestimmter Fahrzeugtypgruppen können also Servicefahrten mit besonderen Anforderungen von geeigneten Fahrzeugen bedient werden.

Servicefahrten: Diese Fahrten werden auch Fahrplanfahrten oder Fahrgastfahrten genannt. Sie werden zur Personenbeförderung eingesetzt und sind durch den Fahrplan vorgegeben. Dazu gehören normale Linienfahrten sowie Sonderfahrten (wie Schulfahrten). Für jede Servicefahrt existiert ein Anfangs- und ein Endhaltepunkt sowie eine Start- und eine Endzeit. Eine Servicefahrt kann als eine fortlaufende Kette von Fahrtabschnitten einer Linie angesehen werden. Ein Fahrtabschnitt ist dabei eine planmäßige ununterbrochene Fahrt zwischen zwei Haltestellen einer Linie. Zwischen den Fahrtabschnitten steigen die Fahrgäste ein bzw. aus. Eine Servicefahrt kann insbesondere die

gleiche Anfangs- und Endstation besitzen, wenn der entsprechende Linienverlauf einen Zyklus darstellt.

Leerfahrten: Als Leerfahrten werden allgemein Fahrten ohne Beförderung der Fahrgäste bezeichnet. Sie werden eingesetzt, um den Standort des Fahrzeugs zu wechseln. Die Leerfahrten werden in Verbindungs-, Einrück- und Ausrückfahrten aufgeteilt.

Verbindungsfahrten: Zwei Servicefahrten heißen *kompatibel*, wenn sie von einem Fahrzeug nacheinander ausgeführt werden können. Dabei können die beiden Fahrplanfahrten entweder direkt verknüpft sein, oder durch eine Verbindungsfahrt verbunden werden, falls der Startpunkt einer Fahrplanfahrt nicht gleich der Endpunkt der vorhergehenden Fahrplanfahrt ist. Die entsprechende Fahrt ohne Passagiere von der Endhaltestelle der vorhergehenden Fahrplanfahrt zur Starthaltestelle der nächsten Fahrplanfahrt wird Verbindungsfahrt genannt.

Ein- und Ausrückfahrten: Neben den Verbindungsfahrten werden im Laufe eines Arbeitstages noch weitere Fahrten ohne Passagierbeförderung ausgeführt. Ausrückfahrten (oder Einsetzfahrten) führen vom Depot zum Anfangshaltepunkt einer Servicefahrt. Einrückfahrten (oder Aussetzfahrten) verbinden den Endhaltepunkt einer Servicefahrt mit dem Depot.

Umlauf: Allgemein wird mit einem Fahrzeugumlauf die Fahrtroute eines Fahrzeuges für einen Betriebstag beschrieben. Damit stellt ein Umlauf eine Sequenz von Fahrzeugaktivitäten, wie Servicefahrten, Leerfahrten und Wartezeiten dar, die im Laufe eines Arbeitstages nacheinander von einem Fahrzeug ausgeführt werden. Die Umläufe beginnen und enden immer in einem Depot. Die Gesamtheit der Fahrzeugumläufe wird in einem *Umlaufplan* festgehalten.

Das generelle Vorgehen der Umlaufplanung sieht vor, dass die Menge aller Servicefahrten eines Betriebstages in disjunkte Teilmengen zerlegt wird. Die in jeweils einer dieser Teilmengen enthaltenen Fahrten sollen von einem Fahrzeug nacheinander bewältigt werden können. Die Fahrzeugtypgruppen aller in einem Umlauf enthaltenen Servicefahrten müssen mindestens einen Fahrzeugtyp gemeinsam haben, also nicht disjunkt sein.

Zudem zeichnet sich die Umlaufplanung im ÖPNV dadurch aus, dass Leerfahrten zum Standortwechsel zwischen der Ausführung von Fahrplanfahrten erlaubt sind. Das heißt, dass ein Bus nach der Ausführung einer planmäßigen Linienfahrt seinen Standort ändern darf, indem er zu einer anderen Haltestelle leer fährt, um dort die nächste Servicefahrt bedienen zu können. Solche Leerfahrten verursachen zusätzliche operative Kosten, können aber den Fahrzeugeinsatz wesentlich reduzieren. Die Gesamtkosten des Fahrzeugeinsatzes, die aus fixen und variablen Kosten-

komponenten bestehen, sind bei der Planung der Umläufe zu minimieren. Diese Planungsaufgabe wird in der Literatur als das Umlaufplanungsproblem (engl.: *vehicle scheduling problem, VSP*) behandelt. Ist dabei für jede Fahrt die Zuordnung zu einem Depot und einem Fahrzeugtyp vorgegeben, so handelt es sich um das Eindepot-Umlaufplanungsproblem, das mit geringem Aufwand gelöst werden kann.

2.3.1 Mehrdepot-Umlaufplanung

Schwieriger gestaltet sich die Lösung des Umlaufplanungsproblems wenn für eine Fahrt mehrere Depots oder Fahrzeugtypen zum Einsatz kommen können. Dabei ist einerseits zu beachten, dass die Schnittmenge der Fahrzeugtypgruppen aller in einem Umlauf enthaltenen Fahrten nicht leer ist und andererseits, dass ein Fahrzeug am Ende eines Betriebstages zum selben Depot zurückkehrt, aus dem es morgens zum ersten Einsatzort ausgerückt war.

Es gilt also, einen Umlaufplan zu bestimmen, der

- alle Fahrten eines gegebenen Fahrplans abdeckt,
- die Anforderungen an Fahrzeugtypen und Depot-Rückkehr erfüllt, und
- minimale Gesamtkosten für das Verkehrsunternehmen verursacht.

Diese Aufgabe wird in der Literatur als das Mehrdepot-Umlaufplanungsproblem (engl.: *multiple depot vehicle scheduling problem, MDVSP*) bezeichnet. Weil jede Fahrplanfahrt von mehreren Depots und mit einigen bestimmten Fahrzeugtypen bedient werden kann, besitzt die Problemstellung eine hohe kombinatorische Komplexität. Wie [Bertossi et al., 1987] zeigen, ist das Mehrdepot-Umlaufplanungsproblem NP-hart.

Ferner können Kapazitätsanforderungen aufgestellt werden. Ein bestimmtes Depot kann eine begrenzte Anzahl von Stellplätzen für die Übernachtung (Stationierung) oder für das vorübergehende Abstellen der Busse während des Tages haben. Diese Kapazitäten können eventuell auch nach Fahrzeugtyp differenziert sein. Das Mehrdepot-Umlaufplanungsproblem mit Kapazitäten ist ebenfalls NP-hart ([Bertossi et al., 1987]).

2.3.2 Kennzahlen für die Bewertung der Umlaufpläne

Im Allgemeinen lassen sich als Kennzahlen oder betriebliche Kennziffern solche Zahlen betrachten, die quantitativ erfassbare Sachverhalte in konzentrierter Form

erfassen. Nach [Reichmann, 1997] sind Kennzahlen vielseitige Instrumente zur Durchführung aussagekräftiger und wirksamer Kontrollen in der betrieblichen Praxis. In der Umlaufplanung werden unterschiedliche Kennzahlen verwendet, die zum einen die Kostenaspekte und zum anderen die Auslastungen und die Betriebsleistung charakterisieren.

Kosten

Eine für die Entscheidungssituation der Umlaufplanung relevante Kennzahl stellt die Höhe der Betriebskosten dar. Damit sind alternative Planungsmöglichkeiten vergleichbar. Dies ermöglicht eine Bewertung der Umlaufplanung im Bezug auf die Wirtschaftlichkeit. Die Kosten hängen im Wesentlichen davon ab, wie viele Fahrzeuge zum Bedienen des Fahrplans gekauft und betrieben werden sollen und wie hoch der Gesamtumfang von den im Umlaufplan enthaltenen Leerfahrten ist. Ein weiterer Kostenfaktor ist die eigentliche Einsatzzeit der Fahrzeuge - ein Fahrzeugaufenthalt (Fahren oder Stehen) außerhalb von vordefinierten Abstellorten verursacht ebenfalls operative Kosten.

Die (anteiligen) Anschaffungskosten für eingesetzte Fahrzeuge bilden den Block der *fixen Kosten*. Sie sind unabhängig von der eigentlichen Fahrleistung. Die von der Einsatzzeit und Kilometerleistung abhängigen Kostenkomponenten werden als *variable Kosten* bezeichnet.

Die Vielfalt der benutzten Fahrzeugtypen spielt bei der Bewertung eines Fahrzeugumlaufplans eine wichtige Rolle. Die Eigenschaften, die die Fahrzeuge für eine Fahrt qualifizieren, können bei vielen Fahrzeugtypen gleich ausgeprägt sein; die fixen und variablen Kosten dieser Fahrzeuge sind aber in der Regel unterschiedlich.

Zur Berechnung der Gesamtkosten eines Umlaufplans ist entsprechend die Angabe von drei Kostensätzen pro Fahrzeugtyp erforderlich: Fahrzeugkosten, Kilometerkosten und Stundenkosten. Die Kostensätze für Fahrzeuge hängen insbesondere davon ab, ob ein Fahrzeug unternehmenseigen oder angemietet ist.

Betriebsleistung und Auslastungen

Nutzeinsatzzeit, Leerfahrtszeit sowie die Längen der produktiven und unproduktiven gefahrenen Strecken werden als Kennzahlen für die Betriebsleistung verwendet.

Als Auslastungskennzahlen werden Fahrplanwirkungsgrade nach unterschiedlichen Berechnungsvorschriften ermittelt. So gibt der *Zeitwirkungsgrad* eines Fahrplans den Anteil der produktiven Dauer an der Gesamteinsatzzeit ohne Betriebshofaufenthalte wieder. Der *Wegwirkungsgrad* wird analog als Anteil der produktiven Streckenlänge an der Gesamtstrecke berechnet.

Diese Berechnungen können aber innerbetrieblich variieren, insbesondere dadurch, dass die Depotfahrten bei manchen Unternehmen als produktive Fahrten in die Berechnung eingehen und bei anderen nicht.

Manchmal werden die eigentlich optimalen Pläne von den ÖPNV-Planern manuell „verbessert“. Es muss also neben den benannten Kennzahlen bestimmte nicht-quantitative Kriterien geben, die nicht formal definierbar sind und ausschließlich dem Expertenwissen erfahrener Planer entspringen. Diese qualitativen Aspekte sind nicht objektiv erfassbar, da der Planer in der Praxis oft nach seinem subjektiven Empfinden handelt.

2.3.3 Praktische Anforderungen an die Umlaufplanung

In der Praxis der Umlaufplanung trifft man auf zahlreiche praktische Randbedingungen und Zielsetzungen, die von Betrieb zu Betrieb unterschiedlich ausgeprägt sind und Einfluss auf die Bildung der Umläufe nehmen können. Diese Randbedingungen und Zielsetzungen geben die Anforderungen an die Planung und Bewertung der Umläufe vor. Sie müssen bei der Konzeption und Implementierung der Modelle und Planungswerkzeuge für die Umlaufplanung beachtet werden. Ein Entscheidungsunterstützungswerkzeug für die Umlaufplanung sollte möglichst viele solcher Randbedingungen aus der Praxis berücksichtigen können. Nachfolgend werden insbesondere folgende Randbedingungen und deren Modalitäten erläutert:

- Zuordnung von Fahrzeugtypen bzw. Fahrzeugtypgruppen zu Fahrten,
- Zuordnung von Fahrzeugumläufen zu den Depots,
- tageszeitabhängige Verbindungszeiten und Entfernungen zwischen den Haltestellen,
- linienreine vs. liniengemischte Planung.

Fahrzeugtypen und Gruppen

Der Fuhrpark eines Unternehmens im öffentlichen Personennahverkehr besteht, wie eingangs des Abschnitts 2.3 erläutert, nahezu niemals aus Fahrzeugen ein und desselben Modells. Vielmehr kommen Fahrzeuge verschiedener Hersteller zum Einsatz, die sich in vielerlei Hinsicht unterscheiden. Diese Unterschiede der einzelnen Fahrzeugmodelle müssen bei der Konstruktion des Umlaufplans von den Planern mitberücksichtigt werden. So müssen unter anderem die Beförderungskapazitäten bei der Zuordnung von Fahrzeugen zu Fahrten genau beachtet werden. Im Berufsverkehr - z.B. morgens zwischen 6:30 und 9:00 Uhr und nachmittags zwischen 16:00 und

17:00 Uhr - müssen viele Personen befördert werden. Dementsprechend können auf manchen Linien nur Fahrzeuge mit hoher Transportkapazität eingesetzt werden.

Weitere Aspekte, die bei der Planung berücksichtigt werden müssen, beziehen sich auf spezielle technische Merkmale der Fahrzeuge. So werden auf Strecken, die Haltestellen an Krankenhäusern und Altersheimen haben, bevorzugt Niederflrbusse eingesetzt.

Als weiteres Beispiel kann der Einsatz von Gelenkbussen dienen. Diese relativ großen Fahrzeuge können nicht auf jeder Strecke eingesetzt werden, da sie durch ihre Ausmaße manche Kurven oder auch Wendeschleifen gar nicht oder nur unter großen Problemen durchfahren können. Doppeldecker-Busse können an einigen Linien nicht eingesetzt werden, falls Brücken mit niedriger Höhe auf der Fahrtstrecke vorhanden sind. Die entsprechenden Fahrten müssen von Fahrzeugen „ohne Überlänge“ bzw. „mit Normalhöhe“ bedient werden.

Wird nur ein Fahrzeugtyp für jede Fahrt erlaubt, so führt es zu einer Erhöhung des Fahrzeugbedarfs. Aus diesen Überlegungen werden die Fahrzeugtypen mit gemeinsamen Eigenschaften zu Fahrzeugtypgruppen zusammengefasst. Solche Fahrzeugtypgruppen können z.B. Niederflrbusse, Gelenkbusse, Kleinbusse oder auch Klimabusse sein. Ein Fahrzeugtyp kann dabei in mehreren Gruppen enthalten sein. So würde ein Gelenkbus mit Niederflur sowohl der Gruppe „behindertengerechte Fahrzeuge“ als auch der Gruppe „Großraumfahrzeuge“ angehören.

Jeder Fahrt wird somit eine Fahrzeugtypgruppe zugeordnet, die die für diese Fahrt geeigneten Fahrzeugtypen beinhaltet. Inwieweit die geforderten Fahrzeugeigenschaften nun den Bedarf an Fahrzeugen erhöhen, hängt stark davon ab, wie die Fahrzeugtypgruppen definiert sind. Sind diese sehr klein, im Extremfall mit genau einem festgelegten Fahrzeugtyp für jede Fahrt, so ist von einem erhöhten Fahrzeugbedarf auszugehen. Daher lohnt es sich schon in der Fahrplanerstellungphase, in der auch die zulässigen Fahrzeugtypen für jede Fahrt festgelegt werden, die Menge der zulässigen Fahrzeugtypen für jede einzelne Fahrt so gross wie möglich und so klein wie nötig zu halten. Dies belegen auch die „what-if“ Untersuchungen in Kapitel 7.

Zuordnung von Fahrzeugumläufen zu Depots

In der in Abschnitt 2.3.1 vorgestellten Aufgabe der Mehrdepot-Umlaufplanung wird jeder Umlauf eindeutig einem Depot zugeordnet. Das entsprechende Fahrzeug darf dann nur in diesem Depot starten, enden und parken. In der Regel verursachen sämtliche Pausen im eigenen Depot, im Gegensatz zu Pausen an Haltestellen, keine oder geringere Personalkosten und sind deswegen vorzuziehen. Allerdings ist eine Fahrt zum zugeordnetem Depot und zurück zwecks einer Pause in vielen Fäl-

len wegen der großen Entfernung zum Depot nicht rentabel bzw. überhaupt nicht möglich.

Ist es für ein Fahrzeug aber möglich, am Ende seiner Dienstfahrten zu einem anderen Depot zurückzukehren, so können dadurch operative Kosten eingespart werden. Die Bedingung, dass ein Fahrzeug am Ende eines Betriebstages wieder in dem Depot stehen muss, von dem es ausgerückt ist, muss dafür aufgehoben werden.

Die Regelung, dass das Start- und Enddepot eines Umlaufs bzw. Umlaufstücks zwischen zwei Depotaufenthalten verschieden sein darf, führt zu einer Einsparung der Leerfahrten und der „teuren“ Standzeiten außerhalb des Depots. Das Fahrzeug ist schneller wieder in einem Depot, verbraucht also weniger Kraftstoff, und der Fahrer kann seinen Arbeitstag früher beenden, was zu geringeren Lohnkosten durch weniger Arbeitsstunden führt.

Durch diese Regelung können also nicht nur lange Aus- bzw. Einrückfahrten eingespart werden, sondern auch viele Leerfahrten. Die letzten können dabei durch eine Fahrt über eins der vordefinierten Depots ersetzt werden. Außerdem wird in vielen Fällen anstatt des „teuren“ Wartens an einer Anschlusshaltestelle eine günstigere Alternative mit einer Pause in einem der Depots möglich.

Damit aber am nächsten Tag alle Fahrten wieder so bedient werden können wie am Vortag, muss eine weitere Restriktion bei der Modellierung beachtet werden: Nachdem alle Fahrzeuge in ein Depot zurückgekehrt sind, müssen sich in jedem Depot so viele Fahrzeuge eines Typs befinden, wie am Morgen desselben Tages. Nur so kann gewährleistet werden, dass der Fahrplan am nächsten Tag entsprechend der Modelllösung abgearbeitet werden kann. Nach einer endlichen Anzahl von Tagen wird jedes Fahrzeug in das ursprüngliche Startdepot zurückkehren.

Eine weitere Kostensenkung entsteht, falls einem Fahrzeug die Möglichkeit gegeben wird, seinen Umlauf nicht nur in den Depots für Pausen unterbrechen zu können, sondern auch an bestimmten Haltestellen, die wegen ihrer speziellen Eigenschaften auch für die durch Arbeitsverträge vereinbarte Ruhepausen geeignet sind. Dies können große Haltestellen sein, wo es sanitäre und gastronomische Einrichtungen gibt. Die Fahrer können ihre Pausen so bequemer gestalten, als an einer gewöhnlicher Haltestelle. Für das Unternehmen bringt das ebenfalls Vorteile, da solche Pausen keine Personalkosten verursachen und während eines längeren Wartens an einer Anschlusshaltestelle zwischen zwei Fahrten stattfinden können.

Es sind also mehrere Varianten der Zuordnungen von Fahrzeugumläufen zu Abstellorten möglich:

- jeder Fahrzeugumlauf startet und endet an einem Ort,
- Start- und Enddepot eines Umlaufs dürfen unterschiedlich sein, und zwar

beliebige Depots des Unternehmens oder

- es werden Depotgruppen analog zu den Fahrzeugtypgruppen definiert, so dass Start- und Enddepot eines Umlaufs derselben Depotgruppe angehören müssen.

Tageszeitabhängige Verbindungszeiten und Entfernungen zwischen den Haltestellen

Nach Ausführung einer Servicefahrt kann ein Fahrzeug seinen Standort wechseln, indem es zu einer anderen Haltestelle leer fährt. Wie lange eine solche Leerfahrt dauert, hängt nicht nur von dem Start- und Endhaltestellen-Paar ab, sondern auch von der Tageszeit, zu der die Fahrzeugumsetzung stattfinden soll. Es ist nahelegend, dass im städtischen Verkehr die Fahrtzeiten zu den Hauptverkehrszeiten stark verlängert sein können. Zusätzlich können auch bestimmte Straßen, Straßenabschnitte oder Brücken zu bestimmten Tageszeiten gesperrt sein. In diesem Fall hängt nicht nur die Fahrtdauer, sondern auch die Länge der zu fahrenden Strecke zwischen zwei Haltestellen von der Tageszeit ab.

So stellt sich an ein adäquates Entscheidungsunterstützungswerkzeug die Anforderung, gegebenenfalls die Fahrtdauer und Entfernung für jede mögliche Relation „von Haltestelle - nach Haltestelle“ tageszeitabhängig abbilden und berücksichtigen zu können.

Regional- vs. Stadtverkehr

Die Unterschiede zwischen städtischen und regionalen Verkehrsformen ergeben sich aus der Topologie des Liniennetzes, der Verkehrsdichte und der Lage von Abstellorten für die Fahrzeuge. Diese Unterschiede haben insbesondere Auswirkungen darauf, wie stark die aufeinanderfolgende Planungsphasen Umlaufplanung und Dienstplanung miteinander gekoppelt werden sollen. Beim schienengebundenen Verkehr stellt die Belegung der Endgleise eine weitere Anforderung an die Planungswerkzeuge – stehen mehrere Fahrzeuge an einem Endgleis, so können sie nur nach dem Last In-First Out (LiFo)-Prinzip zur Bedienung der nächsten Fahrten ausrücken.

Linienreine oder liniengemischte Planung

Es gibt grundsätzlich zwei Möglichkeiten, die Umläufe zu verplanen:

linienrein - ein Fahrzeug bedient im Laufe des Tages nur Fahrten einer Linie,

liniengemischt - ein Fahrzeug kann im Laufe des Tages von einer Linie zu anderen wechseln.

Die Anforderungen bezüglich der Linienreinheit variieren vom Betrieb zu Betrieb. Während die Fahrer eines Unternehmen beliebig viele Linien im Laufe eines Arbeitstages fahren können, wird bei einem anderen Betrieb eine Anzahl von vier Linien pro Tag schon als eine echte Herausforderung angesehen und darf auf keinen Fall überschritten werden.

Die Anforderung, möglichst linienrein zu planen, wird dabei auf unterschiedliche Weise begründet. Die „technische“ Argumentation eines PTV interplan-Anwenders basiert auf der Tatsache, dass sich an Bord der Fahrzeuge sich sogenannte Fahrerkarten befinden, die beschreiben, welche Haltestellenfolge das Fahrzeug abfahren muss. Werden mehrere Linien befahren, wird entweder das entsprechende Blatt Papier sehr groß sein oder die Schrift sehr klein gewählt werden. Eine zu große Vorlage ist unhandlich und bei kleiner Schrift wird der Fahrer das Schriftstück während der Fahrt nicht lesen können. Aus betrieblicher Sicht ist es für einen reibungslosen, fehlerfreien Betrieb sicherlich nicht wünschenswert, dass ein Busfahrer beim Lesen der Fahrerkarte, beim Fahren oder beim Fahrzeugwechseln durcheinanderkommt. So hat, beispielsweise die Linie 34 in München 49 Haltestellen, die angefahren werden müssen. Schon bei vier vergleichbaren Linien in einem Busumlauf wären 196 unterschiedliche Haltestellen anzufahren und dementsprechend auch vom Fahrer zu merken. Hinzu kommt, dass die Fahrer auf den jeweiligen Linien geschult werden müssen.

Die Zulässigkeit von liniengemischten Umläufen hängt insbesondere davon ab, wie stark die Fahrzeugumläufe bei der Bildung der Fahrerdienste zerstückelt werden, also wie oft ein Fahrerwechsel an einem Fahrzeug im Laufe des Tages stattfindet. Stimmen die Fahrerdienste zum größten Teil mit den Fahrzeugumläufen überein, so muss bei der Dienstbildung vermieden werden, dass einem Fahrer ein Fahrzeug zugewiesen wird, dass eine Linie befährt, für die dieser Fahrer keine Ausbildung hat.

Als ein weiteres Argument für die Begrenzung der Anzahl der Linienwechsel pro Umlauf wird erwähnt, dass durch eine einzige Verspätung bei einem Bus, alle Linien, die er noch befährt, und gegebenenfalls auch alle betroffenen Anschlüsse dieser Linien verspätet werden. Sind im Fahrplan viele Anschlüsse definiert, so kann eine lokal aufgetretene Störung zu Anschlusskonflikten und Verspätungen in großen Teilen des Busnetzes führen. Je weniger Linien ein Umlauf enthält, desto weniger verheerend wären die Auswirkungen einer einzigen Verspätung.

Allgemein lässt sich sagen, dass, falls es diese Möglichkeit gibt, die Umläufe aus Wirtschaftlichkeitsgründen liniengemischt geplant werden sollten. Eine ausführlichere Begründung wird in Kapitel 7 dargelegt. Wird allerdings eine Linienberück-

sichtigung in der Umlaufplanung ausdrücklich gewünscht, so sollte unterschieden werden, ob die Anzahl der Linienwechsel, die Anzahl unterschiedlicher Linien pro Umlauf oder beides begrenzt oder minimiert werden sollte. Alle diese Modalitäten sollten von einem Entscheidungsunterstützungswerkzeug unterstützt werden können.

Vergabe an externe Unternehmen

Verkehrsbetriebe sind in der Regel bestrebt, die für den eigenen Betrieb unwirtschaftlichen Fahrten durch private Unternehmen bzw. durch Subunternehmen durchführen zu lassen. Die Auswahl der Fahrten, die ausgelagert werden sollen, geschieht in der heutigen Praxis eher unsystematisch.

Die Subunternehmen mit ihren Depots können aber in die Umlaufplanung des Verkehrsbetriebs miteinbezogen werden, so dass von vornherein die Verteilung der Umläufe am günstigsten für alle Beteiligten ist. Alternativ kann eine volumenbezogene Vergabe von Fahrtleistungen realisiert werden. Dabei wird nicht die gesamte Menge der Fahrten eines Fahrplans durch den Umlaufplan abgedeckt, sondern nur ein durch die Anzahl der Fahrtkilometer oder Stunden vorgegebenes Volumen. Die restlichen, unverplanten Fahrten werden an die privaten Unternehmen vergeben. Besonders bei den größeren Betrieben ist es erstrebenswert, auch diese Planungsaktivität durch geeignete Modelle und Methoden zu unterstützen. Dabei ist es sinnvoll zu überlegen, ob das Outsourcing auf der Grundlage von ganzen Linien, von fertigen Fahrzeugumläufen oder von einzelnen Fahrten geschieht.

2.4 Umlaufplanungswerkzeuge mit Einsatz von Optimierungsmethoden

Wo früher noch Karteikarten und Registerordner gereicht haben, ist die Umlaufplanung im modernen ÖPNV nicht mehr ohne elektronische Hilfsmittel denkbar. Die vormals manuellen Vorgänge sind durch den Einsatz der interaktiven Planungskomponenten weitgehend ersetzt worden. In den letzten Jahren setzte ein zweiter Übergang ein - von den regelbasierten interaktiven Planungswerkzeugen zu Softwareprodukten, die die Methoden der mathematischen Optimierung für eine kostenoptimale und somit zielbasierte Planung der Fahrzeugumläufe einsetzen. Immer mehr Betriebe nutzen spezielle Entscheidungsunterstützungssoftware für die Planung und Optimierung der Betriebsabläufe. Das Einsparungspotential kann dabei mehrere Millionen Euro betragen.

Produkte dieser Art werden von unterschiedlichen Anbietern auf dem Markt für

die Planungssoftware angeboten (vgl. auch [Huschke, 2004]). Dazu zählen unter anderem das System *HASTUS*, das von der Firma GIRO Inc. Montréal ([GIRO, 2004]) vertrieben wird, *microbus* von der IVU Traffic Technologies AG, Berlin ([IVU, 2004]) und *ptv interplan* von der PTV AG ([PTV, 2004]), Karlsruhe.

Die im Rahmen dieser Arbeit konzipierten und implementierten Ansätze wurden von der PTV AG in *ptv interplan* integriert und werden für die Erstellung optimaler Fahrzeugumläufe in der aktuellen Softwareversion eingesetzt. Transparent für den Benutzer fügt sich das Optimierungsmodul in eine an den Windows-Standards gehaltene grafische Oberfläche ein. *ptv interplan* deckt die wichtigsten Planungsbereiche eines Verkehrsbetriebs mit Linienverkehren ab: Fahrplanerstellung, Umlaufplanung, Dienstplanung und Dienstreihenfolgebildung. Als ein Betriebsplanungssystem zur Optimierung von Fahrplänen, Umlaufplänen und Dienstplänen fügt sich *ptv interplan* in das Software-Paket für Verkehrsplanung und Verkehrstechnik *ptv vision* ein, das Werkzeuge für die strategische und operative Planung im öffentlichen Verkehr beinhaltet.

In diesem Kapitel wurden die Ziele und Rahmenbedingungen für die Umlaufplanung beschrieben. Diese werden auch als Anforderungen an die Konzeption eines Modellierungs- und Lösungsansatzes vorgegeben. Im nächsten Kapitel erfolgt eine Überblick der bestehenden Modellierungs- und Lösungsansätze für die optimale Planung der Fahrzeugumläufe.

Kapitel 3

Mehrdepot-Umlaufplanung - Stand der Forschung

Die Aufgabe der Umlaufbildung wird in der Operations Research-Literatur als das Umlaufplanungsproblem abstrahiert. Mehrere Autoren schlagen für dieses Problem Lösungsansätze vor, die auf unterschiedlichen Modellierungsarten basieren. In diesem Kapitel erfolgt zunächst ein Exkurs in die mathematische Optimierung, um weitere zentrale Begriffe, die später bei der Beschreibung der Modelle und der Lösungsansätze benutzt werden, zu erklären. Danach werden in einem Überblick über vorhandene Arbeiten insbesondere die unterschiedlichen Modellierungsansätze kategorisiert und beschrieben. Ein weiterer Aspekt dieses Kapitels ist die Beleuchtung der unterschiedlichen Lösungsansätze. Dabei wird auch beachtet, Probleminstanzen welcher Größe mit dem jeweiligen Ansatz lösbar sind. Es werden die jeweiligen Zielfunktionen vor allem in Hinsicht auf die Beachtung der operativen Kosten untersucht. Zusätzlich werden spezielle Aspekte einzelner Arbeiten erwähnt, deren Berücksichtigung häufig in den anderen Arbeiten nicht gegeben ist.

3.1 Mathematische Optimierung

Der Begriff der *mathematischen Optimierung* oder auch *mathematischen Programmierung*, beschreibt die Maxi- oder Minimierung einer Zielfunktion von Variablen, wobei die Variablen bestimmten Nebenbedingungen (Restriktionen) unterliegen. Dabei wird davon ausgegangen, dass eine Entscheidungssituation durch eine Menge von Variablen, die die Entscheidungen kodieren, definiert werden kann. Das System schreibt durch die *Restriktionen* vor, welche Werte die Variablen annehmen dürfen. Die Restriktionen sind eine Menge von Gleichungen und Ungleichungen, die die Werte der oben genannten Variablen begrenzen. Einen bestimmten Systemzu-

stand erhält man, indem den Variablen konkrete Werte zugewiesen werden. Die Gesamtheit aller Systemzustände (Lösungen), die die Restriktionen erfüllen, stellt die Menge der zulässigen Lösungen dar. Die Bewertung eines zulässigen Systemzustandes kann einen Funktionswert der den Zustand beschreibenden Variablen liefern. Diese Funktion, die es zu minimieren oder zu maximieren gilt, wird *Zielfunktion* (engl.: *objective function*) oder Kostenfunktion genannt. Das Optimierungsproblem ist somit die Aufgabe der Bestimmung einer zulässigen Lösung, die den optimalen Zielfunktionswert verursacht.

Bestehen die Restriktionen und die Kostenfunktion aus linearen Verknüpfungen von Variablen, so spricht man von einem *linearen Programm (LP)*. Die lineare Programmierung spielt eine besondere Rolle in der mathematischen Optimierung, da zum einen eine Vielzahl verschiedener Probleme als lineare Programme modelliert werden kann und zum anderen bereits eine Reihe schneller Algorithmen, die riesige lineare Programme lösen können, entwickelt wurde (s.[Nemhauser and Wolsey, 1988]). Die Methoden der linearen Programmierung spielen ferner eine wichtige Rolle bei der Analyse und Lösung nichtlinearer mathematischer Optimierungsprobleme.

Man spricht von einem *gemischt-ganzzahligen Optimierungsproblem* (engl.: *mixed integer problem, MIP*), wenn eine Teilmenge der Variablen nur ganzzahlige Werte annehmen darf. Beispielsweise kann die Entscheidung, ob zwei Fahrplanfahrten mit einem Fahrzeug nacheinander gefahren werden sollen durch binäre Variable beschrieben werden.

Lineare Optimierungsprobleme, die nur kontinuierliche Variablen beinhalten, sind in polynomieller Zeit, z.B. mit der Interior-Point-Methode lösbar. Der Zeitaufwand für die Lösung derartiger Probleme lässt sich als Polynom bezüglich der Größe der Problem Instanz darstellen. Probleme, für die keine polynomiellen Algorithmen bekannt sind, werden als NP-vollständig (nichtdeterministisch polynomiell) klassifiziert. Die folgenden Abschnitte sollen eine Übersicht über vorhandene Methoden geben, auf die man zur Lösung der gemischt-ganzzahligen Optimierungsprobleme zurückgreifen kann.

Unter *exakten Verfahren* versteht man Verfahren, die mit Sicherheit eine optimale Lösung des vorliegenden Problems finden oder zeigen, dass keine optimale Lösung dafür existiert. Angesichts der NP-Vollständigkeit mancher Probleme wird vermutet, dass die Laufzeit dieser Algorithmen im schlechtesten Fall (*worst-case*) nicht polynomiell ist. Dies bedeutet aber nicht, dass die tatsächliche Laufzeit auf bestimmten Problem Instanzen immer dem worst-case entspricht. Bei der Lösung vieler Probleme zeigt sich, dass sich manche exakte Verfahren „gutartig“ verhalten und in der Praxis sogar lineares Laufzeitverhalten aufweisen. Bei größeren Problem Instanzen werden jedoch die Grenzen des Machbaren relativ schnell erreicht.

Probleme, die im Bereich der Transportplanung auftreten, werden meistens als

gemischt-ganzzahlige Optimierungsprobleme formuliert. Es wurden in diesem Bereich sehr weit fortgeschrittene Lösungstechniken entworfen, die langjährige Erfahrungen widerspiegeln. Die Grundlage fast aller dieser Techniken bildet das im folgenden beschriebene *Branch-and-Bound Verfahren*.

3.1.1 Branch-and-Bound Verfahren

Dieses Verfahren ist ein enumeratives Verfahren, dessen Hauptkomponenten Branching und Bounding sind. Ein Entscheidungsbaum, der am Anfang nur einen Wurzelknoten enthält, wird sukzessive aufgebaut und abgearbeitet. Die Blätter des Baumes entsprechen den zulässigen Lösungen des gemischt-ganzzahligen Problems.

Branching Neue Knoten werden durch die Bildung von Unterproblemen generiert, indem einer *Branching-Variable* Integer-Werte zugewiesen werden. Als Branching-Variable wird eine Variable gewählt, die ganzzahlig sein sollte, aber in der LP-Lösung einen fraktionalen Wert hat. Die Auswahl der Variable kann nach verschiedenen Kriterien erfolgen. Zum Beispiel kann die Variable gewählt werden, deren Wert am nächsten an einem ganzzahligen Wert liegt. Die Lösungsmengen der Unterprobleme sollen möglichst disjunkt sein und deren Vereinigung den MIP-Lösungsraum des übergeordneten Problems darstellen.

Bounding Für die neu generierten Knoten werden zunächst die unteren oder oberen Schranken (je nachdem, ob eine Minimierung oder Maximierung der Zielfunktion vorliegt) berechnet. Sie können durch die Lösung einer *LP-Relaxation* des MIP-Problems geliefert werden. Eine LP-Relaxation erhält man, indem die im MIP-Modell enthaltenen Ganzzahligkeitsbedingungen aufgehoben werden. Nach der Lösung der LP-Relaxation liefert der optimale Wert der Zielfunktion eine untere Schranke (obere Schranke im Falle der Maximierung) für das MIP. Falls für eine LP-Relaxation eines Teilproblems eine optimale Lösung gefunden wird, die für das Ausgangsproblem (MIP) zulässig (erfüllt alle Ganzzahligkeitsbedingungen) und besser als alle bisher gefundenen MIP-Lösungen ist, werden der Zielfunktionswert und die Lösung als die bisher besten gespeichert. Für den Knoten ist kein weiteres Branching mehr nötig. In diesem Fall wird das eigentliche Bounding durchgeführt: es werden alle Knoten und ihre Unterbäume gestrichen, deren untere Schranken (obere Schranke im Falle der Maximierung) schlechter als der Zielfunktionswert der eben gefundenen MIP-Lösung ist. Erfüllt die Lösung der LP-Relaxation die Ganzzahligkeitsbedingungen nicht, wird geprüft, ob die berechnete Schranke besser ist, als die beste der bisher gefundenen MIP-Lösungen. Ist dies nicht der Fall, wird der Knoten als „bounded“ erkannt und nicht weiter bearbeitet.

Zwei wesentliche Punkte bei Branch-and-Bound Verfahren, die unterschiedlich gehandhabt werden können, sind: die Branching-Strategie, nach der die Auswahl der Variablen erfolgt, die für den Aufbau der neuen Zweige im Branch&Bound-Baum benutzt werden, und die Art, wie man die Unterprobleme löst, die in jedem Knoten des Branch&Bound-Baumes entstehen, um eine *Bounding*-Funktion zu erhalten. Fortschritte in diesen zwei Punkten erweitern die Grenzen des Einsatzbereiches von Branch-and-Bound-Verfahren. Im Folgenden werden Verfahren beschrieben, die innerhalb des Branch-and-Bound-Verfahrens zur Ermittlung der oberen oder der unteren Schranken eingesetzt werden können:

Simplex-Methoden - untersuchen die Ecken des konvexen Polyeders, das den Lösungsraum eines Linearen Problems darstellt (Simplex), indem sie von einer Basislösung (Ecke des Polyeders) zur nächsten, besseren durch einen Basistausch gelangen, bis keine Verbesserung mehr möglich ist. Es existieren mehrere Varianten der Simplex-Methode, wie die primale, die duale, die primal-duale und die Netzwerk-Simplex-Methode (s. [Chvatal, 1983]).

Interior-Point-Methode - Diese Methode besteht darin, ausgehend von einem inneren Punkt des LP-Polyeders, eine optimale Lösung der Relaxation zu finden, die auf einer Fläche des LP-Polyeders liegt. Um dann weiter mit Branch-and-Bound-Verfahren arbeiten zu können, sucht man ausgehend von dieser optimalen Lösung der Relaxierung eine optimale Basislösung. Ausführliche Beschreibungen der Interior-Point-Methode findet man in [Andersen, 1998] und in [Wright, 1997].

Column Generation - Bei diesem Verfahren handelt es sich um eine Optimierungstechnik für lineare Programme mit sehr vielen Variablen. Zunächst wird ein partielles Modell mit einer beliebigen Auswahl von Variablen gebildet. Die Lösung des partiellen Modells liefert auch die reduzierten Kosten der Variablen und die Schattenpreise der Restriktionen. Unter Benutzung dieser Informationen wird ein Hilfsmodell formuliert, dessen Lösung eine neue Spalte (Variable) generiert, unter Hinzunahme derer eine bessere Lösung des partiellen Modells gewährleistet ist. Wenn keine solche Variable generiert werden kann, ist die Lösung des partiellen Modells bereits für das Ausgangsproblem optimal. [Barnhart et al., 1996] enthält eine ausführliche Beschreibung dieses Verfahrens.

Lagrange-Relaxation - Bei diesem Verfahren werden die „schwierigen“ Restriktionen des linearen Optimierungsproblems in eine neue Zielfunktion übertragen, so dass die verbleibenden Restriktionen ein Problem bilden, das einfacher gelöst werden kann. Eine wichtige Aufgabe dabei besteht darin, die Lagrange-Multiplikatoren geeignet zu definieren. Dieses Unterproblem wird

meistens mit einem Subgradient-Verfahren gelöst. Eine gute Beschreibung der Lagrange-Relaxation-Methode ist in [Reeves, 1995] nachzulesen.

Branch-and-Cut

Die modernen MIP-Solver (vgl. z.B. [ILOG, 2004] und [Suhl, 2000]) verwenden erweiterte Versionen der Branch-&-Bound Methode - die Branch-and-Cut Methode. Dabei werden in jedem Knoten die sogenannten Schnittebenen (engl.: *cuts*) hinzugefügt, die die optimale LP-Lösung dieses Knotens abschneiden, um dadurch bessere untere Schranken zu berechnen. Weitere Informationen zu Branch-and-Cut-Ansatz sind [Hoffmann and Padberg, 1985] und [Nemhauser and Wolsey, 1988] zu entnehmen.

3.1.2 Heuristische Verfahren

Zu heuristischen Verfahren zählen einerseits allgemeine, universelle Metaheuristiken, wie lokale Suche, Simulated Annealing, Genetische Algorithmen, Tabu Search und andererseits auch viele Algorithmen, die auf ein spezielles Problem zugeschnitten wurden und die die Problemstruktur in irgendeiner Weise ausnutzen. Bei diesen Verfahren gelingt es nicht immer, die Optimalität der gefundenen Lösungen zu beweisen. Trotzdem liefern sie für viele Problemstellungen sehr gute Lösungen in akzeptabler Zeit. In der Regel lässt ein heuristisches Verfahren eine Steuerung der Laufzeit und der Lösungsqualität durch Parametereinstellungen zu. Einen guten Einstieg ins Thema heuristische Verfahren bietet [Reeves, 1995].

3.2 Mehrdepot-Umlaufplanungsproblem

Die Aufgabe der Umlaufplanung, die in Kapitel 2 beschrieben wurde, beinhaltet mehrere Anforderungen und Rahmenbedingungen. Beim Bilden der Fahrzeugumläufe ist zu beachten, dass

- jede Fahrplanfahrt in genau einem Fahrzeugumlauf enthalten ist,
- die Fahrplanfahrten eines Fahrzeugumlaufs von einem Fahrzeug ausführbar sind, d.h. sie beinhalten keine zeitlichen Überschneidungen und sind alle von mindestens einem gemeinsamen Fahrzeugtyp fahrbar,
- jedes Fahrzeug am Ende des Betriebstages zu demselben Depot zurückkehrt, von dem aus es morgens zum ersten Einsatz gestartet ist und

- die Gesamtkosten der Ausführung des Umlaufplans und/oder die Anzahl der benötigter Fahrzeuge minimiert werden.

Es wird grundsätzlich unterschieden, ob eine Umlaufplanungsaufgabe ein Depot (Single Depot Vehicle Scheduling Problem, *SDVSP*) oder mindestens zwei Depots (Multiple Depot Vehicle Scheduling Problem, *MDVSP*) enthält. [Bertossi et al., 1987] haben bewiesen, dass ein Umlaufplanungsproblem NP-hart ist, wenn es mindestens zwei Depots besitzt. Das *SDVSP* ist dagegen in polynomieller Zeit als minimale-Kosten-Flussproblem (engl.: *min cost flow*) lösbar.

Die Komplexität dieser Aufgabe hängt im Wesentlichen von der Anzahl der Depots, der Anzahl der Fahrplanfahrten und der Anzahl möglicher Verbindungsfahrten ab. Zusätzlich können weitere Aspekte wie Beschränkungen der Fahrzeuganzahl zu beachten sein, was zu einer weiteren Erhöhung der Problemkomplexität führt.

Bei den meisten Verkehrsunternehmen werden mehrere unterschiedliche Fahrzeugtypen eingesetzt. In diesem Fall muss darauf geachtet werden, dass eine Fahrplanfahrt gegebenenfalls nicht mehr von Fahrzeugen aus allen Depots bedient werden kann. Es kann auch eine maximale Umlauflänge zu berücksichtigen sein. Dann muss ein Fahrzeug nach einer bestimmten Zeit zum Depot zurückkehren, um zum Beispiel Kraftstoff zu tanken.

Die Veröffentlichungen zu dem *MDVSP* beinhalten unterschiedliche Modelle und Lösungsmethoden. Es werden teilweise auch noch weitere zusätzliche Aspekte beachtet, die nicht Bestandteil des allgemeinen Modells sind. Neben unterschiedlichen Modellierungsansätzen werden in der Literatur auch unterschiedliche heuristische und exakte Lösungsansätze für das *MDVSP* vorgeschlagen. Aufgrund der in den letzten Jahren rapide gestiegenen Rechnerleistungen können immer schnellere und komplexere Lösungsmethoden implementiert werden, mit denen dementsprechend immer größere Probleminstanzen gelöst werden können.

Die Zielfunktionen sollen zwar eine kostenminimale Lösung garantieren, jedoch variieren sie aufgrund unterschiedlicher Annahmen oder auch Kostenbeachtungen. So kann es sein, dass lediglich die Anzahl der Fahrzeuge, die zur Abdeckung aller Fahrplanfahrten benötigt werden, minimiert wird. In den meisten Modellen werden die Gesamtkosten minimiert. Allerdings ergeben sich auch Unterschiede in der Ermittlung der Bestandteile der Gesamtkosten - der operativen Kosten und der Fixkosten. Zusammenfassend kann die Aufgabe der Umlaufbildung wie in der Tabelle 3.1 beschrieben werden.

Das Ziel der Umlaufplanung im ÖPNV kann folgendermaßen formuliert werden: Es soll eine vorgegebene Menge von Servicefahrten eines bereits festgelegten Fahrplans zu Umläufen verkettet werden, die den Depots und den Fahrzeugtypen

Eingabe	Fahrplan (Fahrten mit erlaubten Fahrzeugtypen) Netz (Haltestellen und Depots)
Parameter	Regeln für den Aufbau der Umläufe Dauer jeder möglichen Verbindungsfahrt Definition der Kostensätze weitere Vorgaben und Rahmenbedingungen
Ausgabe	Umlaufplan: Umläufe als Ketten von Fahrzeugaktivitäten mit Angabe des Depots und Fahrzeugtyps optimale Kosten des Umlaufplans

Tabelle 3.1: Aufgabe der Umlaufbildung

zugeordnet werden, so dass diese Umläufe mit einer Minimalanzahl an Fahrzeugen durchgeführt werden können. Neben der Minimierung der Fixkosten für die Fahrzeugflotte wird dabei konkurrierend auch eine Verminderung der operativen Kosten durch eine Minimierung der Leerfahrten und der Pausenzeiten der Fahrzeuge angestrebt.

3.3 Grundlegende Modellierungsansätze

Alle in der Literatur erwähnten Modellierungsarten basieren auf bestimmten Netzwerkstrukturen. Die Depots und die Fahrten werden dabei durch die Knoten oder Kanten im Netzwerk dargestellt. Eine Fahrplanfahrt kann z.B. durch einen einzigen Knoten, durch einen Knoten für den Startpunkt und einen Knoten für den Endpunkt oder durch eine Kante repräsentiert werden. Ein Depot wird in den meisten Netzwerkmodellen ebenfalls durch zwei Knoten repräsentiert. Im ersten starten alle Ausrückfahrten und im zweiten enden die Einrückfahrten. In manchen Fällen führt zusätzlich eine Zirkulationsflusskante vom zweiten Depotknoten zum ersten Depotknoten, um zum Beispiel die Depotkapazitäten berücksichtigen zu können. Alle zulässigen Verbindungsfahrten sowie die Ausrück- und Einrückfahrten werden in dem Netzwerk durch gerichtete Kanten dargestellt, die die entsprechenden Knoten verbinden.

Die durch das Netzwerk fließenden Flusseinheiten entsprechen jeweils einem Fahrzeug. Die Fahrplanfahrten eines Umlaufs für ein Fahrzeug ergeben sich dann aus den besuchten Knoten dieser Flusseinheit. Dabei muss sichergestellt werden, dass ein Fahrzeugumlauf in einem Depot beginnt, Fahrplanfahrten durchführt, die jeweils paarweise kompatibel und gegebenenfalls durch Verbindungsfahrten verknüpft sind, und wieder in demselben Depot endet.

Ein MDVSP mit zwei Depots d_1 und d_2 sowie drei Fahrplanfahrten mit den

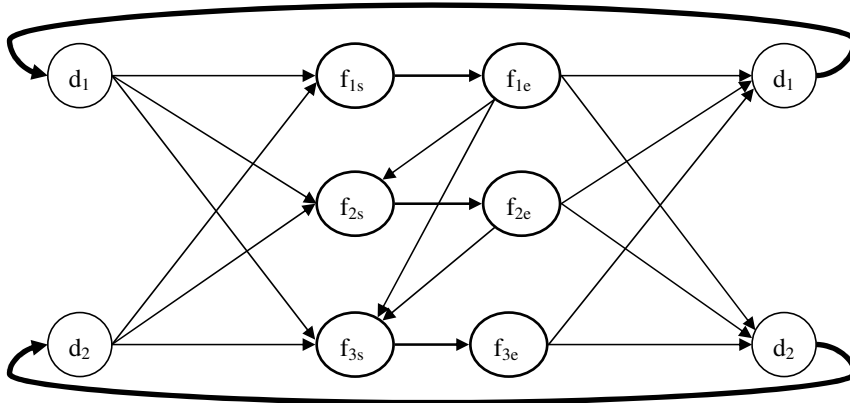


Abbildung 3.1: Ein Netzwerk-Fluss-Modell für das MDVSP

jeweiligen Startpunkten f_{1s} , f_{2s} , f_{3s} und den Endpunkten f_{1e} , f_{2e} , f_{3e} kann wie in der Abbildung 3.1 als ein Netzwerk-Fluss-Modell dargestellt werden. Eine Fahrplanfahrt wird in diesem Fall jeweils durch zwei Knoten, die durch eine gerichtete Kante verbunden sind, modelliert.

Ein Charakteristikum dieses Netzwerkes ist, dass es azyklisch ist, wenn die Zirkulationsflusskanten außer Acht gelassen werden. Das liegt daran, dass zwei Knoten, die unterschiedliche Fahrplanfahrten repräsentieren, nur dann mit einer Kante verbunden werden, wenn die beiden Fahrplanfahrten kompatibel sind. Mit diesem Netzwerk-Fluss-Modell kann das SDVSP zum Beispiel leicht als minimale-Kosten-Flussproblem gelöst werden. Dazu ist eine Transformation nötig, die in der Abbildung 3.2 dargestellt ist. Die Endpunkte und die Startpunkte der Fahrplanfahrten werden zu Angebotsknoten bzw. Bedarfsknoten (mit einem Bedarf bzw. Angebot von Eins), die Depotknoten stellen reine Umladeknoten dar. Dieses Modell kann mit Standard-Netzwerkalgorithmen gelöst werden, die beispielsweise in [Ahuja et al., 1983] beschrieben sind.

Viele Modellierungsansätze für das MDVSP stellen Mehrgüter-Fluss-Modelle dar, die in Abschnitt 3.3.1 vorgestellt werden. Der Abschnitt 3.3.2 beschreibt die Quasi-Assignment-Modelle und der Abschnitt 3.3.3 enthält ein Set-Partitioning-Modell. In Abschnitt 3.3.4 werden die speziellen Aspekte der Modelle aus den einzelnen Arbeiten hervorgehoben.

3.3.1 Mehrgüter-Fluss-Modelle

Mit dem Graphen aus der Abbildung 3.1 kann nicht sichergestellt werden, dass ein Umlauf in dem Depot endet, von dem aus er auch gestartet ist, oder dass bestimmte Fahrplanfahrten nur von bestimmten Depots aus oder nur mit bestimmten Fahr-

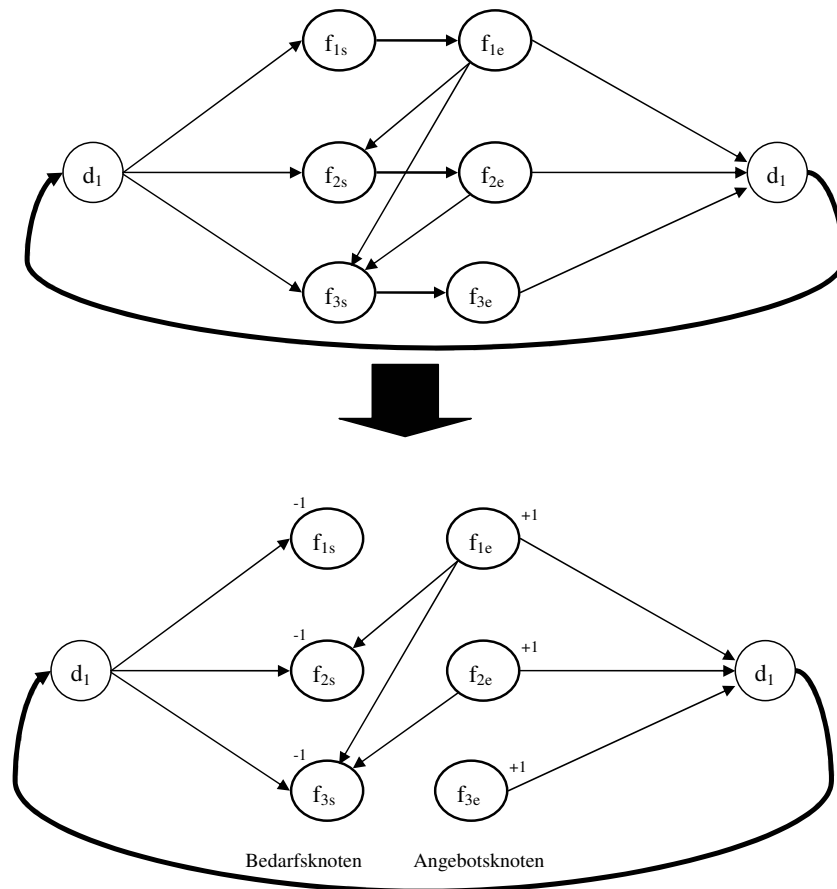


Abbildung 3.2: Transformation eines SDVSP zu einem minimale-Kosten-Flussproblem in Anlehnung an [Löbel and Strubbe, 1996]

zeugtypen bedient werden können. Deswegen wird in Mehrgüter-Fluss-Modellen meistens für jedes Depot eine eigene Netzwerkschicht angelegt¹. Das erhöht erheblich die Komplexität des Modells durch die zusätzlichen Restriktionen, die sicherstellen müssen, dass eine Fahrplanfahrt nur von genau einem Fahrzeug bedient wird. Ein Mehrgüter-Fluss-Modell mit zwei Netzwerk-Schichten wird in der Abbildung 3.3 dargestellt. Zur Lösung des entsprechenden Problems wird ein lineares Programm aufgestellt.

[Forbes et al., 1994] präsentieren eine Mehrgüter-Fluss-Formulierung, die allerdings mehr in den Ausdrücken der Quasi-Assignment-Formulierung angegeben wird. Im Netzwerkflussmodell Modell wird ein Gut mit jeweils einem Bus aus ei-

¹Diese Bedingung muss nicht zwangsläufig durch unterschiedliche Netzwerk-Schichten sichergestellt werden. Es können auch andere Bedingungen im linearen Programm dafür sorgen, dass eine Fahrplanfahrt nur von bestimmten Depots aus bedient werden kann.

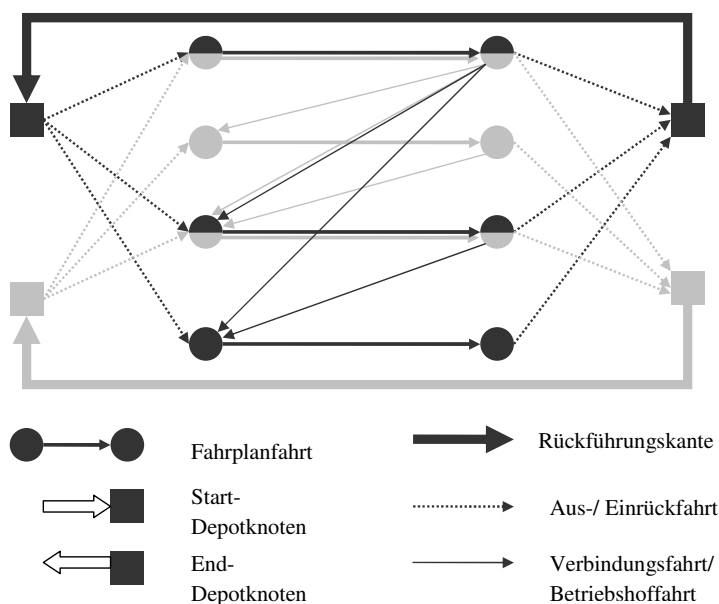


Abbildung 3.3: Mehrgüter-Fluss-Modell nach [Löbel and Strubbe, 1996]

nem Depot assoziiert. Das entsprechende Netzwerk sieht dann wie das Netzwerk in der Abbildung 3.3 aus. Das MDVSP wird auch als Quasi-Assignment-Problem mit einem entsprechenden linearen Programm formuliert, dessen Lösung als Ausgangspunkt zur Lösung der LP-Relaxation der Mehrgüter-Fluss-Formulierung dient.

[Löbel and Strubbe, 1996] formulieren das MDVSP als ein Mehrgüter-Fluss-Modell. Ein entsprechender Graph mit zwei Depots und vier Fahrplanfahrten ist in der Abbildung 3.3 dargestellt. Damit die Anzahl der Umläufe minimiert wird, werden zwei zusätzliche Konzepte angewandt. Zunächst wird zu allen Gewichtungen der Ausrückfahrten ein hinreichend großer Wert M addiert, der größer als jede mögliche Gewichtung einer zulässigen Umlaufplanung ist. Da aber in der Realität ein Fahrzeug während eines Umlaufs ein Depot zwecks einer Pause besuchen kann, ist durch diese sogenannte Big-M-Methode nicht sichergestellt, dass eine minimale Anzahl an Umläufen generiert wird. Deshalb wird ein Depotbesuch durch eine Betriebshoffahrt modelliert. Sie gleicht einer Verbindungsfahrt, da sie auch zwei Fahrplanfahrten verbindet. Die Gewichtung der Betriebshoffahrten wird durch Parameter gesteuert.

In [Grötschel et al., 1997] wird das gleiche Modell wie in [Löbel and Strubbe, 1996] verwendet. Es werden auch untere und obere Schranken für jede Kante eingeführt. Bei einer Zirkulationsflusskante ist die untere Schranke gleich der unteren Kapazitätsschranke des Depots und die obere Schranke gleich der maximalen Kapazität des Depots. Für die übrigen Kanten beträgt der Wert der unteren Schranke null und

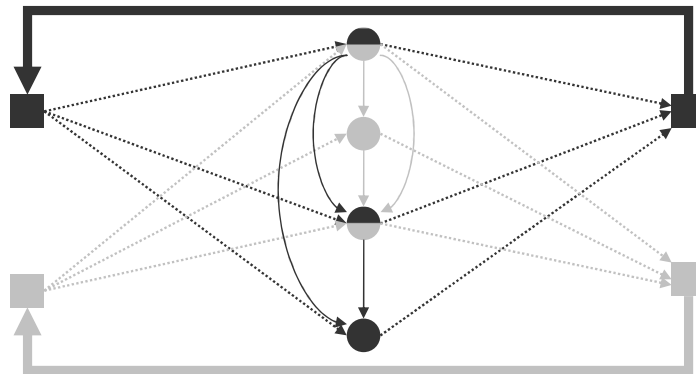


Abbildung 3.4: Reduziertes Mehrgüter-Fluss-Modell nach [Grötschel et al., 1997]

der Wert der oberen Schranke eins. Der Graph aus der Abbildung 3.3 wird in der Abbildung 3.4 in einen „reduzierten“ Graphen überführt, in dem die Fahrplanfahrten nur durch einen Knoten dargestellt werden. Dieses Mehrgüter-Fluss-Problem wird als ganzzahliges lineares Programm gelöst. Jeder Knoten, der eine Fahrplanfahrt repräsentiert, wird von genau einer Flusseinheit besucht.

Das Netzwerkmodell in [Löbel, 1997] ist dasselbe wie in [Grötschel et al., 1997]. Die Formulierung des ganzzahligen linearen Programms, das aus der Mehrgüter-Fluss-Formulierung abgeleitet ist, ist dieselbe wie die von [Forbes et al., 1994]. Der einzige Unterschied besteht darin, dass in [Löbel, 1997] untere Depotkapazitäten berücksichtigt werden. Wenn der Zeitabstand zwischen zwei kompatiblen Fahrplanfahrten eine bestimmte Grenze (zum Beispiel 60 Minuten) überschreitet, soll keine Verbindungsfahrt zwischen diesen beiden Fahrplanfahrten möglich sein, da dann zum Beispiel die Wartezeit zu lang wird. Um aber trotzdem die Ausführung dieser beiden Fahrplanfahrten von einem Fahrzeug nicht zu verhindern, da das ein Bestandteil der optimalen Lösung sein könnte, werden an dieser Stelle eine Einrückfahrt zum Depot und eine Ausrückfahrt vom Depot zu der zweiten Fahrplanfahrt eingeschoben. Nach der ersten dieser beiden Fahrplanfahrten wird das Fahrzeug also im Depot geparkt, bevor es dann die zweite Fahrplanfahrt durchführt. Bei den Berechnungen wurden allerdings die Depotkapazitäten nicht berücksichtigt. Der Author erwähnt, dass diese als weiche Nebenbedingungen berücksichtigt werden können und in der Praxis nicht so relevant wären.

[Banihashemi and Haghani, 2000] erweitern das Mehrgüter-Fluss-Modell. Die kompatiblen Fahrplanfahrten werden in *depot-kompatible* und *straßen-kompatible* Fahrplanfahrten aufgeteilt. Zwei Fahrplanfahrten sind depot-kompatibel, wenn es zulässig und weniger kostspielig für das Fahrzeug ist, zwischen dem Bedienen der beiden Fahrplanfahrten zum Depot zurückzukehren. Zwei Fahrplanfahrten heißen straßen-kompatibel, wenn es unzulässig oder kostspieliger für das Fahrzeug ist,

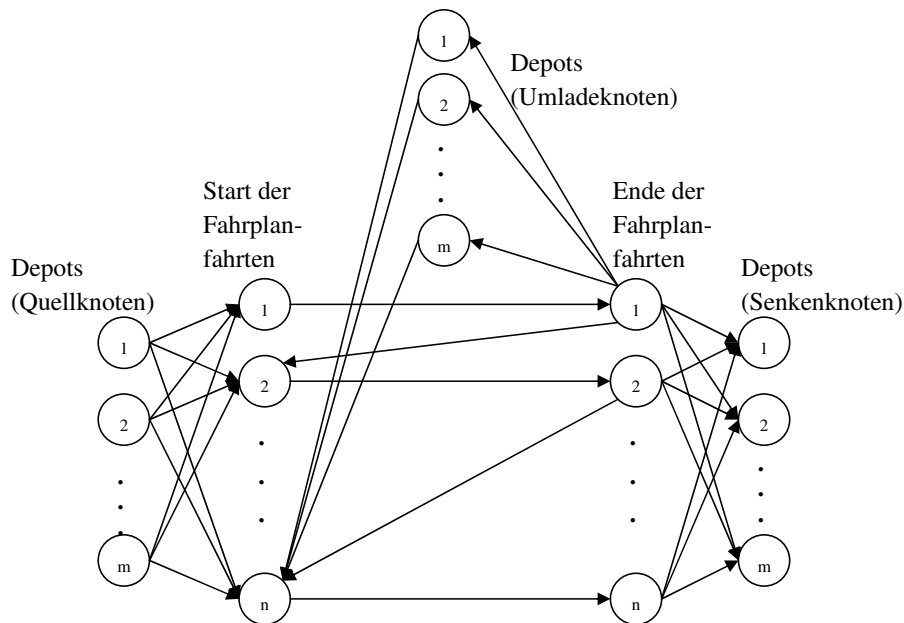


Abbildung 3.5: Netzwerk nach [Banihashemi and Haghani, 2000]

zwischen dem Bedienen der beiden Fahrplanfahrten zum Depot zurückzukehren. Alle Fahrplanfahrten werden in die drei Mengen - Morgenfahrten, Mittagsfahrten und Nachmittagsfahrten eingeteilt. Dabei ist das Hauptkriterium, dass alle Fahrplanfahrten aus der Menge der Morgenfahrten depot-kompatibel zu allen Fahrplanfahrten aus der Menge der Nachmittagsfahrten sind. Die restlichen Fahrplanfahrten gehören dann der Menge der Mittagsfahrten an.

Das entsprechende Netzwerkmodell ist in der Abbildung 3.5 dargestellt. Die Anzahl der Kanten wird reduziert, indem die Morgenfahrten keine ausgehenden Kanten zu den Nachmittagsfahrten, dafür aber welche zu den Umladeknoten (Depots), haben. Dementsprechend haben die Nachmittagsfahrten zwar zusätzlich eingehende Kanten von den Depots, dafür aber keine von den Morgenfahrten. Dadurch wird die Anzahl der Variablen enorm verringert, da es viel weniger Depots als Morgen- oder Nachmittagsfahrten gibt. In der Abbildung 3.5 stellt die Fahrplanfahrt 1 eine Morgenfahrt dar, Fahrplanfahrt 2 eine Mittagsfahrt und Fahrplanfahrt m eine Nachmittagsfahrt. Es wird gezeigt, dass diese Formulierung äquivalent zu der von [Forbes et al., 1994] ist.

Zusätzlich werden in [Banihashemi and Haghani, 2000] Zeitbeschränkungen für Blöcke im MDVSP berücksichtigt. Ein Block ist ein Teil eines Umlaufs und beginnt mit einer Ausrückfahrt und endet mit der nächsten Einrückfahrt. Ein Grund für diese Beschränkung ist zum Beispiel die Notwendigkeit, nach einer bestimmten zurückgelegten Strecke Kraftstoff zu tanken. Deshalb werden der mathematischen

Formulierung zusätzliche Bedingungen zur Reduktion eines zu langen Blocks um mindestens eine Fahrplanfahrt hinzugefügt. Diese Bedingungen werden erst nach einem ersten Lösen des MDVSP ohne die Zeitbeschränkungen hinzugefügt. Große Probleminstanzen mit einer hohen Anzahl an Fahrplanfahrten und dementsprechend vielen Variablen in der mathematischen Formulierung sind als MDVSP mit Zeitbeschränkung (engl.: *MDVSP with route time constraints; MDVSPRTC*) praktisch nicht mehr handhabbar. Deshalb werden zwei heuristische Techniken vorgestellt, um die Problemgröße zu verringern. Die erste Technik fasst zwei oder mehr Fahrplanfahrten als eine Fahrplanfahrt zusammen, wenn

- i die entsprechenden Fahrplanfahrten paarweise kompatibel sind,
- ii in allen kompatiblen Paaren von Fahrplanfahrten der Endpunkt der ersten Fahrplanfahrt dem Startpunkt der zweiten Fahrplanfahrt entspricht,
- iii die Gesamtzeit vom Beginn der ersten Fahrplanfahrt bis zum Ende der zweiten Fahrplanfahrt inklusive der benötigten Zeit für Leerfahrten von und zu den Depots die maximal erlaubte Zeit nicht überschreitet und
- iv die Zeit zwischen der Ankunft der ersten Fahrplanfahrt und dem Start der zweiten Fahrplanfahrt in allen kompatiblen Paaren eine vorgegebene maximale Wartezeit nicht überschreitet.

Mit der zweiten Technik wird die Anzahl der Variablen im Modell weiter verringert. Durchschnittlich betrug die Reduktionsrate 78%.

[Huisman et al., 2001] präsentieren ein Modell zur Lösung des dynamischen VSP. Hier wird zusätzlich zu allgemeinen MDVSP-Anforderungen versucht, die Anzahl der auftretenden Verspätungen möglichst gering zu halten. Normalerweise werden in einem ÖPNV-Unternehmen die operativen Planungsaufgaben in bestimmten Zeitabständen gelöst. Dabei werden ein Fahr-, ein Fahrzeugumlauf- und ein Dienstplan berechnet. [Huisman et al., 2001] schlagen vor, auf diese Weise, also in der Regel halbjährlich, nur noch den Fahrplan festzulegen und die Umlauf- und Dienstpläne dynamisch zu berechnen. Begründet wird dieses Vorgehen durch die variablen Fahrtzeiten, die nicht exakt vorhergesagt werden können. Eine Zielsetzung dabei ist, die Fahrzeug- und Fahrereinsatzpläne bei den auftretenden Verspätungen zu „reparieren“. Dafür wird das VSP mehrere Male pro Tag gelöst. Die Fahrtzeiten für die nächste Zeitscheibe werden als bekannt angenommen, für den Rest der Planungsperiode werden stochastische Daten mit Hilfe von Szenarien berücksichtigt. Im mathematischen Modell werden neben den Kosten für Fahrzeuge auch die induzierten Verzögerungszeiten, gewichtet mit Wahrscheinlichkeiten für Szenarien, minimiert. Dadurch können bei auftretenden Verspätungen Fahrzeuge

zu anderen Fahrplanfahrten zugeordnet werden, um möglichst Kettenreaktionen durch eine verspätete Fahrplanfahrt zu vermeiden.

Es wird zuerst eine Problemformulierung für das statische VSP angegeben. Die Grundidee dieser Formulierung ist ebenfalls eine Mehrgüter-Fluss-Formulierung mit den unterschiedlichen Netzwerken für die jeweiligen Depots. Um aber die Anzahl der Variablen zu reduzieren, wird noch die Kantenmenge verringert. Unter der Annahme, dass keine Kosten zu beachten sind, wenn ein Fahrzeug im Depot parkt, werden lange Kanten definiert. Zwei kompatible Fahrplanfahrten werden durch eine lange Kante miteinander verbunden, wenn es möglich und gleichzeitig nicht teurer ist, dass ein Fahrzeug zwischen den beiden Fahrplanfahrten zum Depot zurückkehrt. Dann können die langen Kanten aus dem Modell gelöscht werden. Um die Anzahl der Verspätungen zu reduzieren, werden zusätzliche Variablen eingeführt. Darauf wird in dieser Arbeit aber nicht näher eingegangen.

3.3.2 Quasi-Assignment-Modelle

In den Arbeiten, die ein Quasi-Assignment-Modell benutzen, gibt es keine Standard-Netzwerk-Formulierung, da für unterschiedliche Aspekte unterschiedliche Netzwerke benutzt werden. Die Quasi-Assignment-Formulierung stellt eine Erweiterung des klassischen Zuordnungsproblems (engl.: *assignment problem*), bei dem n Objekte den n Positionen kostenminimal zugeordnet werden sollen. Das lineare Programm eines Quasi-Assignment-Modells entspricht in etwa einem Transportproblem, bei dem nicht alle Verbindungen zwischen Angebots- und Nachfrageknoten erlaubt sind. In den folgenden Arbeiten ist das Quasi-Assignment-Problem zumindest die Grundlage für die Modellansätze, auch wenn sie sonst erweitert oder leicht abgeändert wurden.

[Dell'Amico et al., 1993] und [Fischetti et al., 1999] benutzen dasselbe oder zumindest ein sehr ähnliches Modell, in dem vorausgesetzt wird, dass alle Fahrzeuge identisch sind. Die Folge davon ist, dass keine Fahrzeugtypgruppen für Fahrplanfahrten beachtet werden. Das bedeutet, dass jede Fahrplanfahrt von jedem Depot aus bedient werden kann und dementsprechend auch keine depotspezifischen Kosten berücksichtigt werden müssen. Das Netzwerk enthält für jede Fahrplanfahrt und jedes Depot einen Knoten, wobei die Menge der Knoten in die Teilmenge W und V zerlegt wird, die die Depot- bzw. die Fahrplanfahrten-Knoten beinhalten. Zum Lösen des MDVSP muss eine Menge von Zyklen gefunden werden, wobei jeder Zyklus nur einen Knoten aus W enthält. Dieser Knoten aus W entspricht dem Depot, in dem der durch den Zyklus repräsentierte Fahrzeugumlauf startet und endet. Außerdem muss sichergestellt werden, dass jeder Knoten aus V in genau einem Zyklus enthalten ist, jeder Knoten aus W in höchstens so vielen Zyklen enthalten ist,

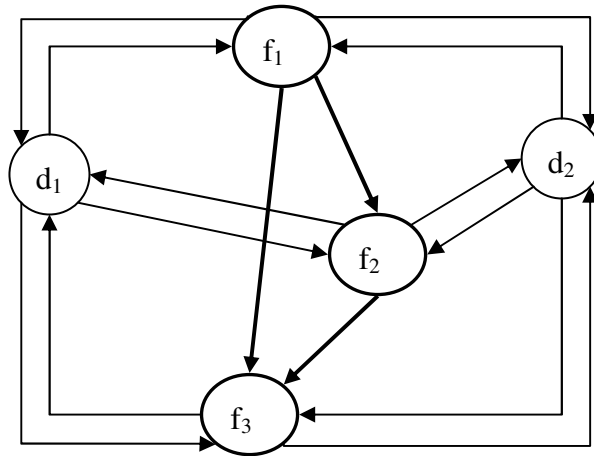


Abbildung 3.6: Graph nach dem Modellansatz von [Dell'Amico et al., 1993]

wie Fahrzeuge in dem entsprechenden Depot stationiert sind. Dabei sollen die Anzahl der Zyklen und deren Kosten minimiert werden. Abbildung 3.6 zeigt, wie das Netzwerk mit zwei Depots und drei Fahrplanfahrten nach [Dell'Amico et al., 1993] aussieht.

In [Fischetti et al., 1999] werden die Knoten des Netzwerks $V = 1, \dots, m + n$ aufgeteilt: die erste Teilmenge $W = \{1, \dots, m\}$ enthält für jedes Depot D_k einen Knoten k , und die zweite Teilmenge $N = \{m+1, \dots, m+n\}$ enthält für jede Fahrplanfahrt T_j einen Knoten $m+j$. Die Kantenmenge A enthält alle Kanten zwischen allen Paaren von Knoten. Allerdings werden die Kanten, die zwei nicht kompatible Fahrplanfahrten verbinden, mit unendlichen Kosten belegt. Diese Kanten werden nicht aus dem Modell gelöscht, um die Notation zu vereinfachen. Ohne die Kanten mit unendlichen Kosten ist das Netzwerk dem in [Dell'Amico et al., 1993] vorgestellten ähnlich und ist in der Abbildung 3.7 mit zwei Depots und drei Fahrplanfahrten dargestellt. Jeder Zyklus mit endlichen Kosten, der genau einen Knoten aus der Teilmenge W und mindestens einen Knoten aus der Menge N enthält, entspricht einem gültigen Umlauf.

Das MDVSP wird dann als die Bestimmung einer kostenminimalen Menge von Zyklen mit jeweils genau einem Knoten aus der Menge W formuliert. Alle Fahrplanfahrten-Knoten aus N müssen genau einmal besucht werden, wobei jeder Depotknoten $k \in W$ höchstens r_k mal besucht wird. Die Kapazität des Depots d_k beträgt r_k . Das lineare Programm macht deutlich, dass das MDVSP als Quasi-Assignment-Modell mit zusätzlichen pfadorientierten Bedingungen formuliert ist. Die Entscheidungsvariable x_{ij} hat den Wert eins, falls die Kante (i, j) in der optimalen Lösung benutzt wird (die Fahrplanfahrt j nach der Fahrplanfahrt i von demselben Fahrzeug durchgeführt wird). Das lineare Programm sieht dann wie folgt aus:

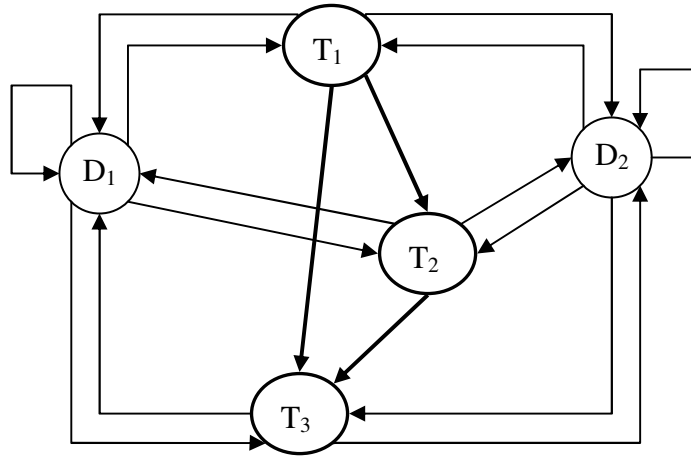


Abbildung 3.7: Graph nach dem Modellansatz von [Fischetti et al., 1999]

$$\min \sum_{i \in V} \sum_{j \in V} c_{ij} x_{ij} \quad (3.1)$$

$$\sum_{i \in V} x_{ij} = r_j \quad \forall j \in V \quad (3.2)$$

$$\sum_{j \in V} x_{ij} = r_i \quad \forall i \in V \quad (3.3)$$

$$\sum_{(i,j) \in P} x_{ij} = |P| - 1 \quad \forall P \in \Pi \quad (3.4)$$

$$x_{ij} \geq 0 \quad \forall (i,j) \in V \quad (3.5)$$

$$x_{ij} \text{ integer} \quad \forall (i,j) \in V \quad (3.6)$$

Für jedes $i \in N$ ist $r_i := 1$ definiert. Π ist die Menge der Pfade mit endlichen Kosten, die zwei verschiedene Depot-Knoten aus W verbinden. Die Bedingungen 3.2 und 3.3 stellen sicher, dass jeder Knoten genau r_k mal besucht wird. Werden aber nicht alle r_k Fahrzeuge eines Depots benötigt, so wird diese Bedingung trotzdem aufgrund der Kante (k, k) mit den Kosten null erfüllt. Diese Kante gibt also die Anzahl der nicht benötigten Fahrzeuge des Depots d_k an. Die Bedingung 3.4 ist pfadorientiert. Sie sorgt dafür, dass es keinen Pfad gibt, der zwei verschiedene Depots miteinander verbindet, indem sie fordert, dass in einem solchen ungültigen Pfad die Entscheidungsvariable mindestens einer Kante den Wert null hat. Falls das Modell nur ein Depot beinhaltet, ist die Menge Π leer, und das lineare Programm reduziert sich zu einem Transportproblem.

[Silva et al., 2000] präsentieren ein Modell für das SDVSP. Es wird ein Netzwerk gebildet, in dem jede Fahrplanfahrt durch einen Knoten für den Startpunkt und

einen Knoten für den Endpunkt dargestellt wird. Wenn zwei Fahrplanfahrten kompatibel sind, dann verbindet eine Kante den Endpunkt der ersten Fahrplanfahrt mit dem Startpunkt der zweiten Fahrplanfahrt. [Silva et al., 2000] definieren zwei Fahrplanfahrten als kompatibel, wenn sie von einem Fahrzeug hintereinander ausgeführt werden können und zusätzlich eine vordefinierte Pufferzeit zwischen dem Ankunftszeitpunkt der ersten Fahrplanfahrt und dem Abfahrtszeitpunkt der zweiten Fahrplanfahrt vorhanden ist. Dadurch sollen mögliche Verspätungen beachtet werden. Das Depot wird durch den Abfahrtsknoten g und den Ankunfts-knoten p repräsentiert. Vom Abfahrtsknoten des Depots führt jeweils eine Kante zu jedem Startpunkt einer Fahrplanfahrt, und von jedem Endpunkt einer Fahrplanfahrt führt jeweils eine Kante zum Ankunfts-knoten des Depots. Da die Anzahl der Variablen aber sehr hoch ist, werden in dem Netzwerk zunächst nicht alle Kanten eingetragen. Sogenannte lange Kanten werden weggelassen. Eine Kante zwischen zwei Fahrplanfahrten ist eine lange Kante, wenn ein Fahrzeug zwischen dem Bedienen der beiden Fahrplanfahrten zum Depot zurückkehren und dort eine vordefinierte minimale Zeit parken kann. Diese Kanten werden in späteren Schritten der Lösungsprozedur Schritt für Schritt identifiziert und einbezogen, falls sie die Lösung verbessern können.

[Freling et al., 2001] definieren ebenfalls ein Modell für das SDVSP, da sie einen Algorithmus entwickeln, um das SDVSP als Teilproblem von größeren Problemen, wie zum Beispiel der kombinierten Fahrzeug- und Fahrereinsatzplanung, möglichst schnell lösen zu können. Das Modell wird ein asymmetrisches Assignment-Modell genannt, weil das Depot durch n Knoten repräsentiert wird, während es beim linearen Assignment-Modell durch $2n$ Knoten repräsentiert wird, wobei n die Anzahl der Fahrplanfahrten ist. Das Netzwerk für das Modell enthält für jede Fahrplanfahrt $\in \{1, 2, \dots, n\}$ einen Knoten i und für jeden Knoten i noch einen Knoten $n + i$. Die Knotenmenge $\{n + 1, \dots, 2n\}$ repräsentiert das Depot. Zwei Fahrplanfahrtnoten sind durch gerichtete Kanten miteinander verbunden, wenn die beiden Fahrplanfahrten kompatibel sind. Von jeder Fahrplanfahrt i führt eine gerichtete Kante zu dem Knoten $n + i$. Die Abbildung 3.8 zeigt ein Beispiel mit drei Fahrplanfahrten.

Anhand des Netzwerks wird allerdings die Assignment-Formulierung nicht deutlich. Deshalb wird an dieser Stelle zusätzlich das entsprechende lineare Programm angegeben. Dafür sind einige Definitionen notwendig. Sei $N = \{1, 2, \dots, n\}$ die Menge der Fahrplanfahrten, die aufsteigend nach der Startzeit nummeriert sind. Die Knotenmenge ist dann definiert als $M = N \cup \{n + 1, \dots, 2n\}$. Die Menge der Kanten sei dann $E^T = E \cup \{(i, n + i), i = 1, \dots, n\}$, wobei $E = \{(i, j) | i < j, i \text{ und } j \text{ sind kompatibel}, i \in N, j \in N\}$. Seien c_{ij} die Kosten, die anfallen, wenn ein Fahrzeug die Fahrplanfahrt j nach der Fahrplanfahrt i bedient, c_{it} die Kosten, die anfallen, wenn ein Fahrzeug nach der Fahrplanfahrt i in das Depot zurückkehrt, und c_{sj} die Kosten, die anfallen, wenn ein Fahrzeug nach dem Ausrücken

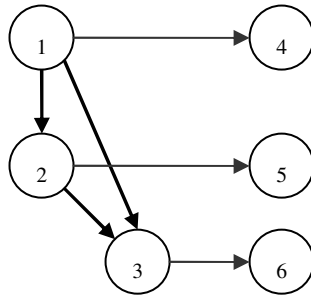


Abbildung 3.8: Graph nach der Modellierung von [Freling et al., 2001]

aus dem Depot die Fahrplanfahrt j bedient. Da in dem Graph die Ausrückfahrten nicht einbezogen werden, werden die Kosten für alle möglichen Ausrückfahrten in die Zielfunktion übernommen und dafür die Kosten für eine Ausrückfahrt von den Kosten für eine Verbindungsfahrt subtrahiert, wenn die Ausrückfahrt nicht durchgeführt wird. Deshalb sei $b_{ij} = \begin{cases} c_{ij} - c_{sj}, & \forall (i, j) \in E \\ c_{ij}, & \forall (i, j) \in (E^T - E). \end{cases}$

Mit x_{ij} als Entscheidungsvariable, die den Wert eins hat, falls ein Fahrzeug die Fahrplanfahrt j direkt nach der Fahrplanfahrt i bedient, und sonst den Wert null hat, kann dann das lineare Programm wie folgt formuliert werden:

$$\min \sum_{(i,j) \in E^T} b_{ij} x_{ij} + \sum_{j \in N} c_{sj} \quad (3.7)$$

$$\sum_{j: (i,j) \in E^T} x_{ij} = 1 \quad \forall i \in N \quad (3.8)$$

$$\sum_{i: (i,j) \in E^T} x_{ij} \leq 1 \quad \forall j \in M \quad (3.9)$$

$$x_{ij} \in \{0, 1\} \quad \forall (i, j) \in E^T \quad (3.10)$$

3.3.3 Set-Partitioning-Modelle

Das MDVSP kann als ein Set-Partitioning-Problem modelliert werden. Grundsätzlich geht es beim Set-Partitioning-Problem darum, dass aus einer bestimmten Menge von Spalten eine kostenminimale Teilmenge ausgewählt werden muss, die jede Zeile genau einmal überdeckt. Beim MDVSP bestehen die Zeilen aus allen Fahrplanfahrten und die Spalten sind die möglichen Umläufe. Ein Umlaufplan soll die Menge aller Fahrplanfahrten partitionieren, d.h. in Teilmengen (Umläufe) aufteilen, die sich nicht überschneiden und deren Vereinigung der Gesamtmenge aller Fahrplanfahrten entspricht. Set-Partitioning-Modelle werden, im Gegensatz zu den bisher beschriebenen kantenorientierten Modellen, allgemein pfadorientierte Modelle

genannt, da die Entscheidungsvariablen der Set-Partitioning-Modelle den gesamten Umläufen, also Pfaden, entsprechen.

In [Ribeiro and Soumis, 1994] wird das MDVSP als ein Set-Partitioning-Problem modelliert. Es wird für jedes Depot ein Graph betrachtet, dessen Knotenmenge aus der Menge $N = 1, \dots, n$ mit jeweils einem Knoten für jede Fahrplanfahrt und einem Knoten k für das entsprechende Depot aus der Menge $K = 1, \dots, m$ besteht. Für die Set-Partitioning-Formulierung wird zusätzlich für jedes Depot k die Menge von Pfaden betrachtet, die im Depot k starten, danach einige Knoten im Graphen (also Fahrplanfahrten) besuchen und wieder zum Depot k zurückkehren. In dem entsprechenden linearen Programm wird sichergestellt, dass jeder Knoten, der eine Fahrplanfahrt repräsentiert, von genau einem Pfad bzw. Zyklus besucht wird. Eine weitere Bedingung stellt sicher, dass nicht mehr Pfade von einem Depot ausgehen als Fahrzeuge in diesem Depot vorhanden sind. Dabei bezeichnet r_k die Fahrzeugkapazität des Depots k . Jeder Zyklus repräsentiert einen Umlauf für ein Fahrzeug. Für jeden Pfad $p \in \Omega = \bigcup_{k=1}^m \Omega_k$ sei c_p die Summe der Kosten von dessen Kanten. Die Variable a_{jp} habe den Wert eins, wenn der Pfad p den Knoten $j \in N$ besucht, sonst ist $a_{jp} = 0$. Für jeden Pfad $p \in \Omega$ sei y_p die entsprechende 0/1-Variable. Das lineare Programm zu dieser Set-Partitioning-Modellierung sieht dann wie folgt aus:

$$\min \sum_{k=1}^m \sum_{p \in \Omega_k} c_p y_p \quad (3.11)$$

$$\sum_{k=1}^m \sum_{p \in \Omega_k} a_{jp} y_p = 1 \quad \forall j \in N \quad (3.12)$$

$$\sum_{p \in \Omega_k} y_p \leq r_k \quad \forall k \in K \quad (3.13)$$

$$y_p \in \{0, 1\} \quad \forall p \in \Omega \quad (3.14)$$

Da die LP-Relaxation dieses linearen Programms mit Column Generation gelöst wird, wird auch ein lineares Programm für das Lösen des Unterproblems, der Generierung der Spalte mit den minimalen Grenzkosten, angegeben. Dieses Problem besteht darin, einen kürzesten Pfad zu finden, der von einem Depot aus über mindestens einen Knoten, der eine Fahrplanfahrt repräsentiert, wieder zurück zum Depot führt.

3.3.4 Spezielle Aspekte der Modellierung

Es wurden zum Teil schon bei der Beschreibung der einzelnen Modellansätze spezielle Aspekte des jeweiligen Modells erwähnt. In diesem Abschnitt werden sie ex-

plizit zusammengefasst, um eine gute Übersicht darüber zu geben, wo bestimmte Bedingungen in die Modelle einbezogen oder auch weggelassen werden.

In den Modellen von [Dell'Amico et al., 1993] und [Fischetti et al., 1999] wird angenommen, dass alle Fahrzeuge identisch sind. Es werden also keine Fahrzeugtypen- und Fahrzeugtypgruppen beachtet.

[Ribeiro and Soumis, 1994] geben an, dass die Software GENCOL, mit der das Column Generation implementiert wurde, einige zusätzliche Bedingungen beachten kann. Dazu gehören Zeitfenster, priorisierte Verbindungsfahrten, Ressourcenverbrauch und Längenbeschränkungen.

[Löbel and Strubbe, 1996] führen als zu berücksichtigende Nebenbedingungen eine maximale Umlauflänge, eine maximale Anzahl der Linienwechsel pro Fahrzeug, eine Mindest- und Höchstwendezeit unter anderem aufgrund von Pausenregelungen und eine Einteilung des Netzes in Netzsegmente, um Kettenreaktionen bei Verspätungen zu vermeiden. Allerdings wird nicht beschrieben, wie diese Nebenbedingungen in der Lösungsmethode berücksichtigt werden.

In dem Modell von [Grötschel et al., 1997] können zusätzlich folgende Sonderfälle beachtet werden: Vermeidung von Linienwechsel (für Fahrzeuge, bei denen ein Linienwechsel aufwendige Umrüstungen erfordert); Vermeidung von Fahrtartenwechsel (Stamm- und Ergänzungsfahrten wie Schulbusfahrten sollen nicht gemischt werden); Berücksichtigung einer maximalen Wendezeit; Festlegung von Erst- und Letztfahrten; Berücksichtigung von Bereitschaften und vordefinierten (Teil-)Umläufen; geplante Depotwechsel. Diese Sonderfälle können im Modell berücksichtigt werden, wenn bestimmte Kanten in dem Graphen gestrichen oder neu eingeführt werden oder wenn die Kosten entsprechend modifiziert werden. Es können aber nicht eine maximale Umlaufdauer bzw. Umlauflänge für bestimmte Fahrzeugtypen und eine maximale Anzahl von Linienwechseln pro Umlauf berücksichtigt werden. Diese Anforderungen könnten nur durch eine heuristische oder manuelle Nachbearbeitung berücksichtigt werden.

Im Modell von [Löbel, 1997] werden Schranken für die Anzahl der Fahrzeuge berücksichtigt. Bei der Lösung der Testinstanzen werden aber keine Depotkapazitäten berücksichtigt. Um zu lange Wartezeiten zwischen der Durchführung zweier aufeinanderfolgender Fahrplanfahrten zu vermeiden, werden ab einer bestimmten Zeitgrenze statt einer Verbindungsfahrt eine Ein- und eine Ausrückfahrt in das Netzwerk eingefügt.

[Desaulniers et al., 1998] behandeln das MDVSP mit Zeitfenstern für Fahrten unter Berücksichtigung der Kosten für Wartezeit. Die Modellierung erfolgt in einem Mehrgüterfluss-Netzwerk und zur Lösung wird ein in einer Branch-and-Bound-Methode eingebetteter Column Generation Ansatz benutzt. Die Berechnung der

unteren Schranken erfolgt durch Dantzig-Wolfe-Dekomposition. Die Zielfunktion beschreibt die Gesamtkosten des Umlaufplans inklusive Kosten für die Wartezeit und ist wegen der Zeitfenster für Fahrten nicht linear.

Im Gegensatz zu den meisten anderen Arbeiten werden in [Silva et al., 2000] zwei Fahrplanfahrten als kompatibel definiert, wenn sie von einem Fahrzeug hintereinander ausgeführt werden können und zusätzlich eine vordefinierte Pufferzeit zwischen dem Ankunftszeitpunkt der ersten Fahrplanfahrt und dem Abfahrtszeitpunkt der zweiten Fahrplanfahrt vorhanden ist. Dadurch sollen mögliche Verspätungen beachtet werden. Diese Definition kann aber auch leicht in die Definition aus den anderen Arbeiten übernommen werden. Parkt ein Fahrzeug während seines Umlaufs im Depot, so muss es eine definierte minimale Zeit dort parken. Die operativen Kosten während des Parkens sind ein fester Wert. Sie sind also unabhängig von der Länge der Aufenthaltsdauer des Fahrzeugs im Depot. Da einem Planer manchmal schon im Voraus bekannt sein kann, dass manche Fahrplanfahrten idealerweise einen bestimmten Nachfolger haben sollten, ist es möglich durch Extra-Kosten diese Verknüpfung der beiden Fahrplanfahrten im Netzwerk zu bevorzugen.

[Banihashemi and Haghani, 2000] lösen das MDVSP mit Zeitbeschränkungen für Blöcke. Dabei ist ein Block ein Teil eines Umlaufs und beginnt mit einer Ausrückfahrt und endet mit der nächsten Einrückfahrt. Ein Grund für diese Beschränkung ist zum Beispiel die Notwendigkeit zu tanken.

[Huisman et al., 2001] präsentieren ein Modell zur Lösung des dynamischen VSP, um die Anzahl der Verspätungen möglichst gering zu halten. Dabei wird das VSP mehrere Male pro Tag gelöst. Depotkapazitäten werden nicht beachtet. Es sei aber einfach, diese einzubeziehen.

3.3.5 Zielfunktionen und Kostenbeachtung

In der Regel werden in den Zielfunktionen der linearen Programme zur Lösung eines MDVSP die Kosten minimiert. Durch die Einbeziehung von Fixkosten oder ähnliche Methoden wird meist auch eine minimale Fahrzeuganzahl erreicht. Nachfolgend wird angegeben, was in den Zielfunktionen einzelner Modelle minimiert wird. Zusätzlich wird dargestellt, in welcher Weise die Kosten berechnet worden sind.

In [Dell'Amico et al., 1993] wird die Anzahl der eingesetzten Fahrzeuge (die in diesem Modell identisch sind) minimiert, indem die Anzahl der Zyklen im Netzwerk minimiert wird. Es wird behauptet, dass der verwendete heuristische Algorithmus immer die minimale Anzahl an Fahrzeugen garantiert, was jedoch nicht bewiesen wird. Als untergeordnetes Ziel werden auch die Kosten minimiert. Die

zufällig erzeugten Probleminstanzen haben eine sehr ähnliche Struktur wie die in [Ribeiro and Soumis, 1994] generierten Instanzen, und die Kosten der einzelnen Kanten werden ebenfalls ähnlich wie in [Ribeiro and Soumis, 1994] mit Hilfe der euklidischen Distanz berechnet. Für die Lösung der Probleminstanzen wird außerdem als untergeordnetes Ziel die Minimierung einer Mischung aus Fahrt- und Wartezeiten der Fahrzeuge angegeben.

Im Modell von [Ribeiro and Soumis, 1994] wird die Summe der Kosten der Umläufe, die in der optimalen Lösung benutzt werden, minimiert. Die Kosten eines Umlaufs entsprechen der Summe der Kosten für die Verbindungsfahrten sowie der Kosten für die Ausrückfahrt und der Kosten für die Einrückfahrt. Für die gelösten Zufallsinstanzen wurden die Kosten für die Leerfahrten mit Hilfe der euklidischen Distanz des Startpunktes und des Endpunktes der Leerfahrt ausgedrückt. Die Kosten einer Verbindungsfahrt werden durch die Summe des zehnfachen der euklidischen Distanz und des doppelten der Wartezeit berechnet. Zu den Ausrück- und Einrückfahrten wurde jeweils zusätzlich ein so großer Wert addiert, dass eine optimale Lösung mit der minimalen Anzahl an Fahrzeugen gefunden wird. Dementsprechend werden diese Kosten durch die Summe des zehnfachen der euklidischen Distanz und des Wertes 5.000 berechnet.

[Forbes et al., 1994] minimieren die Kosten, die aus den variablen Kosten und den Fixkosten für den Einsatz der Busse bestehen. Die variablen Kosten können depotspezifisch beachtet werden. Die Fixkosten werden jeweils zur Hälfte zu den Kosten für die Aus- und Einrückfahrten hinzuaddiert. Es wurden Zufallsinstanzen wie in [Ribeiro and Soumis, 1994] gelöst, wobei die Kosten nach demselben Schema ermittelt wurden.

[Grötschel et al., 1997] minimieren die Summe der Kosten für die in der optimalen Lösung benutzten Kanten. Die Kosten für die Kanten können dabei depotspezifisch angegeben werden. In die Kosten für Verbindungsfahrten können zum Beispiel auch Kosten für die Wartezeit des Busses einbezogen werden, wenn der Bus nach der Verbindungsfahrt nicht sofort die nächste Fahrplanfahrt bedient. Genaue Kostendefinitionen müssen betriebsspezifisch erfolgen, wobei die Zielsetzungen des Betriebes quantitativ umzusetzen sind. Es werden dieselben Methoden benutzt wie in [Löbel and Strubbe, 1996], um eine minimale Anzahl an Fahrzeugen zu garantieren.

[Löbel, 1997] beschreibt als Zielsetzung die Minimierung der Anzahl der eingesetzten Fahrzeuge mit dem untergeordneten Ziel der Minimierung der operativen Kosten. In der Zielfunktion werden allerdings die Kosten für alle Leerfahrten minimiert. Eine minimale Fahrzeuganzahl wird aber dadurch erreicht, dass zu jeder Ausrückfahrt ein hinreichend großes M addiert wird, das größer ist als die operativen Kosten jeder zulässigen Lösung. Sollen depotspezifische Kosten für eine

Fahrplanfahrt berücksichtigt werden, so können dementsprechend die Kosten der vorausgehenden Verbindungsfahrt geändert werden.

Die Zielfunktion in [Desaulniers et al., 1998] beschreibt die Gesamtkosten des Umlaufplans inklusive Kosten für die Wartezeit und ist wegen der Berücksichtigung von Zeitfenstern für Fahrten nicht linear.

[Fischetti et al., 1999] minimieren die Gesamtkosten der Leerfahrten. Alle Kanten zwischen nicht kompatiblen Fahrplanfahrten sind mit unendlichen Kosten belegt. Es werden drei Möglichkeiten für die Bewertung der Kanten mit Kosten vorgeschlagen. Wenn die Anzahl der Fahrzeuge minimiert werden soll, sollten die Kosten der Kanten für die Ausrückfahrten den Wert eins erhalten. Alle anderen Kanten sollten den Wert null erhalten. Sollen die Gesamtkosten minimiert werden, müssen die Kanten mit Kostenwerten, die die Betriebskosten der Fahrzeuge repräsentieren, und mit Strafkosten für Wartezeiten etc. belegt werden. Drittens können die beiden erwähnten Möglichkeiten irgendwie kombiniert werden. Zur Lösung der zufällig generierten Testinstanzen wurden die Kosten genauso wie in [Ribeiro and Soumis, 1994] ermittelt.

[Silva et al., 2000] minimieren die Gesamtkosten der Umläufe. Um zu erreichen, dass eine minimale Anzahl an Fahrzeugen ermittelt wird, werden jeweils die Hälfte der Fixkosten für ein Fahrzeug zu den Kosten jeder Kante, die eine Ausrückfahrt repräsentiert, und zu den Kosten jeder Kante, die eine Einrückfahrt repräsentiert, hinzuaddiert. Die Kosten werden ähnlich wie in [Ribeiro and Soumis, 1994] ermittelt. Allerdings wird hier mehr Spielraum für die Kostenberechnung gelassen, indem bestimmte Konstanten nicht fest vorgegeben werden.

[Banihashemi and Haghani, 2000] minimieren in ihrer Zielfunktion die Summe aller Kosten. Dabei werden beim Beginn und Ende jeden Umlaufs jeweils die Hälfte der Fixkosten für die Unterhaltung eines Fahrzeugs zu den Kosten für die Aus- und Einrückfahrten addiert. Wird ein Fahrzeug während seines Umlaufs zwischenzeitlich im Depot geparkt, so werden die Kosten der entsprechenden Ein- und Ausrückfahrten nicht noch zusätzlich mit den halben Fixkosten belastet. Damit wird sichergestellt, dass ein Fahrzeug zwischenzeitlich zum Depot zurückkehrt, falls das günstiger ist. In der Kostendefinition sind nicht nur die Kosten für Leerfahrten, sondern auch die Kosten für die Fahrplanfahrten enthalten. Durch diese Relation zwischen den Fixkosten und den operativen Kosten wird eine minimale Fahrzeuganzahl erreicht. Dabei können die Kostensätze für die Durchführung von Fahrplanfahrten, für die Durchführung von Leerfahrten, für Wartezeiten und für das Parken im Depot unterschiedlich sein. Allerdings setzten [Banihashemi and Haghani, 2000] voraus, dass die Kosten pro Minute für die Durchführung von Fahrplanfahrten, für die Durchführung von Leerfahrten und für Wartezeiten identisch seien, während die Kosten für das Parken im Depot gleich null sind.

[Freling et al., 2001] minimieren die Kosten, die den Kanten im Netzwerk zugeordnet sind. Die entsprechenden Kosten zur Lösung der Testinstanzen sind ähnlich wie in [Ribeiro and Soumis, 1994] ermittelt worden. Ähnlich wie bei [Silva et al., 2000] wird aber mehr Spielraum für die Festlegung einiger Variablen gelassen. Durch diese Kostenberechnungen wird eine minimale Anzahl an Fahrzeugen garantiert.

[Huisman et al., 2001] minimieren die Summe der Fixkosten und der operativen Kosten. Die operativen Kosten stellen für jede Netzwerkkante eine Funktion der Fahrt- und der Wartezeit, oder eine Funktion der Verspätung (bzw. eine hinreichend große Zahl, wenn Verspätungen nicht vorkommen sollen). So sollen die Gesamtkosten und die Verspätungen minimiert werden. In den Testinstanzen betragen die Fixkosten pro Fahrzeug 10.000 Kosteneinheiten. Die variablen Kosten wurden mit einer Kosteneinheit pro Minute für die Zeit, in der ein Fahrzeug ohne Passagiere ist, ermittelt. Es wird somit keine Unterscheidung zwischen Fahren und Warten gemacht, was zu einer ungenauen Abbildung der realen Kostensituation führt. Für verspätete Fahrplanfahrten wurden zusätzlich Kosten in Höhe von 10.000 Kosteneinheiten berücksichtigt. Die Ergebnisse machen deutlich, dass die Verspätungen auf Kosten einer nur gering höheren Fahrzeuganzahl reduziert werden können.

3.4 Lösungsansätze

Ein grundlegendes Unterscheidungsmerkmal für die existierenden Lösungsansätze ist, ob mit einem Lösungsansatz das MDVSP exakt optimal oder annähernd optimal (heuristisch) gelöst wird. Für das SDVSP gibt es effiziente Algorithmen, die das Problem exakt lösen. Der Grund für einen heuristischen Lösungsansatz beim MDVSP ist, dass damit oft Probleme mit mehr Fahrplanfahrten und Depots in kürzeren Zeit gelöst werden können und die Abweichung zur optimalen Lösung oft nur minimal ist. Da in den letzten Jahren die Rechnerleistung stark gestiegen ist, können immer größere Probleme exakt gelöst werden. Die Frage, ob ein Lösungsansatz exakt oder heuristisch sein soll, geht also oft mit der Frage einher, ob ein großes Problem relativ schnell gelöst werden soll, oder ob die Lösung auf jeden Fall exakt sein soll und somit nicht unbedingt sehr große Probleme gelöst werden können. In Abschnitt 3.4.1 werden zunächst die exakten in Abschnitt 3.4.2 die heuristischen Lösungsansätze aus der Literatur betrachtet. Die jeweiligen Arbeiten werden in etwa historisch angeordnet. Es wird also gleichzeitig die chronologische Entwicklung der Möglichkeiten zur Lösung des MDVSP deutlich. Zusätzlich werden nach Möglichkeit die Größe der gelösten Instanzen, also die Anzahl der Fahrplanfahrten und die Anzahl der Depots der größten mit dem jeweiligen Lösungsansatz gelösten Instanz, sowie die dafür benötigte Rechenzeit und Rechnerleistung, mit der das Problem gelöst wurde, angegeben.

3.4.1 Exakte Ansätze

Die meisten hier betrachteten Lösungsansätze ermöglichen eine exakte Lösung des VSP. In einigen Arbeiten werden allerdings zusätzlich zu einem exakten Lösungsansatz noch heuristische Lösungsansätze angegeben, mit denen größere Probleminstanzen gelöst werden können.

[Forbes et al., 1994] lösen das MDVSP in drei Stufen. In der ersten Stufe wird das Quasi-Assignment-Problem gelöst, wobei das Original-Problem in ein SDVSP relaxiert wird, das als ein reines Netzwerkfluss-Problem schnell lösbar ist. In der zweiten Stufe wird die optimale Lösung des relaxierten Problems dazu genutzt, eine dual zulässige Basis für die LP-Relaxation zu schaffen, die dann mit dem dualen Simplex-Algorithmus optimal gelöst wird. Stufe drei benutzt einen Branch-and-Bound-Ansatz, um eine ganzzahlige optimale Lösung zu erhalten. Die Tests haben ergeben, dass über die Hälfte der Probleme schon nach der zweiten Stufe eine ganzzahlige Lösung hatten.

Durchgeführt wurden Tests auf einem Intel 80486/25 MHz-Prozessor für zwei Klassen von Problemen mit unterschiedlichen Strukturen. Die größte gelöste Instanz beinhaltet 200 Fahrplanfahrten und drei Depots. Für das Lösen der Probleme aus der ersten Klasse wurden für Stufe 2 durchschnittlich 1.771 Sekunden und für Stufe 3 insgesamt 522 Sekunden benötigt, während zur Lösung der Probleme aus der zweiten Klasse für Stufe 2 durchschnittlich 2.558 Sekunden und für Stufe 3 insgesamt 17.165 Sekunden benötigt wurden. Die Zeiten für Stufe 1 werden nicht angegeben, da die Zeit zur Lösung immer nur einen kleinen Prozentsatz von der Zeit, die für Stufe 2 benötigt wird, ausmacht. Die beiden Probleme haben 37.258 bzw. 38.786 Variablen durchschnittlich.

Zusätzlich wird eine Instanz mit 600 Fahrplanfahrten und 91.556 Variablen gelöst, wobei allerdings ungefähr 72% der Fahrplanfahrten nur einem Depot zugeordnet werden, 25% von zwei Depots aus bedient werden können und lediglich 3% Fahrzeugen aus allen drei Depots zugeordnet werden können. Außerdem werden hier keine Depotkapazitäten berücksichtigt. Für die Lösung dieser Instanz werden in Stufe 2 durchschnittlich 4.017 Sekunden und in Stufe 3 insgesamt 1.396 Sekunden benötigt.

[Ribeiro and Soumis, 1994] lösen die LP-Relaxation ihres linearen Programms mit Column Generation, wobei zur Generierung neuer Spalten mit minimalen marginalen Grenzkosten ein Kürzeste-Wege-Problem gelöst wird. Es wird ein kürzester Pfad gesucht, der, ausgehend vom entsprechenden Depot-Knoten, mindestens einen Fahrplanfahrt-Knoten besucht und dann wieder zum Depot zurückkehrt. Zusätzlich wird ein Branch-and-Bound Algorithmus benutzt.

Es wird gezeigt, dass die durch das Column Generation erreichte untere Schran-

ke mindestens so gut ist wie die durch zusätzliche Techniken erreichbare untere Schranke. Der Column Generation-Ansatz wird mit der Software GENCOL implementiert, wobei als LP-Solver ILOG CPLEX benutzt wird. Die Mehr-Depot-Fahrzeugumlaufprobleme wurden auf einer Sun Sparc 2 Workstation gelöst. Die größte gelöste Instanz, die zufällig erzeugt wurde, besitzt sechs Depots und 300 Fahrplanfahrten. Die Lösung wird in durchschnittlich 3.346,8 Sekunden erreicht, wobei die maximal benötigte Zeit bei 5.253 Sekunden liegt. Die gelösten Instanzen sind vier bis fünf Mal größer als die, die bis dahin auf einem schnellen Rechner gelöst werden konnten.

In [Löbel and Strubbe, 1996] wird das mathematische Modell für das MDVSP durch ein Schnittebenenverfahren in Verbindung mit einem Branch-and-Bound-Algorithmus gelöst. Dafür werden Pricing-Strategien und Column Generation benutzt. Mit dieser Lösungsmethode werden flottenminimale Lösungen innerhalb weniger Stunden auf einer leistungsfähigen SUN Workstation berechnet. Leider fehlen genaue Angaben zur Größe der gelösten Testinstanzen. In [Grötschel et al., 1997] werden ebenfalls Column Generation und ein Branch-and-Bound-Algorithmus auf das MDVSP angewendet. Es handelt sich anscheinend um eine Verbesserung des Lösungsverfahrens aus [Löbel and Strubbe, 1996]. [Grötschel et al., 1997] bezeichnen ihre Lösungsmethode als ein Branch-and-Cut-Verfahren. Es wird eine Optimallösung für eine reale Problem Instanz mit 21.003 Fahrplanfahrten angegeben. Diese wurde auf einer SUN SPARC Station 20-71 in 41 Stunden berechnet. Allerdings konnten die Fahrplanfahrten nicht allen Depots zugeordnet werden. Es fehlen Angaben darüber, von wie vielen Depots eine Fahrplanfahrt durchschnittlich bedient werden durfte.

In [Löbel, 1997] werden unterschiedliche Lösungsmethoden für die Lösung von MDVSP beschrieben. Diese werden so miteinander kombiniert, dass große Problem Instanzen gelöst werden können. Zuerst wird eine untere Grenze für die minimale Flottengröße und die minimalen betrieblichen Kosten so nah wie möglich am ganzzahligen optimalen Wert durch Lagrange Relaxation ermittelt. Danach wird eine obere Grenze durch die Eröffnungsheuristik Schedule-Cluster-Reschedule oder Cluster first-Schedule second berechnet. Nachdem die Grenzen ermittelt worden sind, wird die LP-Relaxation durch Column Generation und Column Elimination gelöst, wobei ein neues Lagrange Pricing benutzt wird. Optional kann an dieser Stelle versucht werden, durch die Heuristik LP-Plunging die aktuelle ganzzahlige obere Grenze zu verbessern. Wenn die obere Grenze aus einer praktischen Sichtweise klein genug ist, wird der Algorithmus beendet. In vielen Testfällen wurde schon in dieser Stufe eine optimale Lösung generiert, so dass der Algorithmus gestoppt werden kann. Sonst wird Column Generation angewandt. Die resultierende LP-Relaxation wird fixiert und mit Branch-and-Cut zu einer bewiesenen Optimalität gelöst. Branch-and-Cut ist nur eine Heuristik, wenn nicht alle Spalten mit

bestimmten reduzierten Kosten generiert worden sind. Ein Ablaufdiagramm für die Lösung des MDVSP wird in der Abbildung 3.9 dargestellt. Die alternativ betrachtete Dantzig-Wolfe-Dekompositionsmethode hat sich als eine für das MDVSP ungeeignete Methode herausgestellt.

Die Tests sind ohne Beachtung von Depotkapazitäten auf einem SUN 170 UltraSPARC mit 512 MB Arbeitsspeicher und 1,7 GB virtuellem Speicher durchgeführt worden. Alle LP-Relaxationen sind mit CPLEX 4.0.9 gelöst worden.

Die größten optimal gelösten Instanzen sind „Berlin-Spandau 1“ mit durchschnittlich 4,94 Depots, 2.424 Fahrplanfahrten und 164.000 Leerfahrten, „Berlin-Spandau 5“ mit durchschnittlich 2,25 Depots, 3.331 Fahrplanfahrten und 75.000 Leerfahrten, „Hamburg 1“ mit durchschnittlich 2,23 Depots, 8.563 Fahrplanfahrten und 1.322.000 Leerfahrten sowie „Hamburg-Holstein 2“ mit durchschnittlich 3,65 Depots, 5.447 Fahrplanfahrten und 1.054.000 Leerfahrten.

Die Instanzen „Berlin 2“ mit durchschnittlich 1,56 Depots, 24.906 Fahrplanfahrten und 304.000 Leerfahrten, „Berlin-Spandau 2“ mit durchschnittlich 5,49 Depots, 3.308 Fahrplanfahrten und 327.000 Leerfahrten und „Berlin-Spandau 8“ mit durchschnittlich 5,02 Depots, 3.308 Fahrplanfahrten und 283.000 Leerfahrten sind mit einer minimalen Flotte, aber nicht mit bewiesenen minimalen Kosten gelöst worden. Leider werden nur Laufzeiten für einzelne Methoden, nicht aber eine Gesamtlaufzeit zur Lösung der Instanzen, angegeben.

[Desaulniers et al., 1998] modellieren das MDVSP mit Zeitfenstern für Fahrten unter Berücksichtigung der Kosten für Wartezeit durch ein Multi-Commodity-Flow Modell. Zur Lösung wird ein in einer Branch-and-Bound-Methode eingebetteter Column Generation Ansatz benutzt. Die Berechnung der unteren Schranken erfolgt durch Dantzig-Wolfe-Dekomposition. Instanzen mit bis zu 250 Fahrten und 5 Depots wurden auf einer HP 9000/735 Workstation in der Zeit unter einer Stunde exakt optimal gelöst. Mit dem heuristischen Verfahren wurden in vergleichbarer Zeit größere Instanzen mit bis zur 500 Fahrten gelöst.

[Fischetti et al., 1999] stellen für die optimale Lösung von MDVSP einen Branch-and-Cut-Ansatz vor. Dafür wird zunächst eine neue Klasse von Bedingungen, die Pfad-Eliminierungs-Bedingungen genannt werden, eingeführt, die die pfadorientierten Bedingungen aus dem linearen Programm ersetzen sollen, da die Relaxation des ursprünglichen Modells normalerweise schlechte untere Grenzen liefert. Eine solche neue Bedingung dominiert die schwache alte Bedingung. Es werden unterschiedliche Methoden vorgestellt, um möglichst viele verletzte Pfad-Eliminierungs-Bedingungen zu finden. Für Branch-and-Cut-Algorithmus werden die Cuts zu dem linearen Programm hinzugefügt, die zu den neuen Klassen von Bedingungen gehören. Es wird auch ein neues Pricing-Schema benutzt, das reduzierte Kosten berechnet, um Variablen zu dem linearen Programm hinzuzufügen. Da bei großen In-

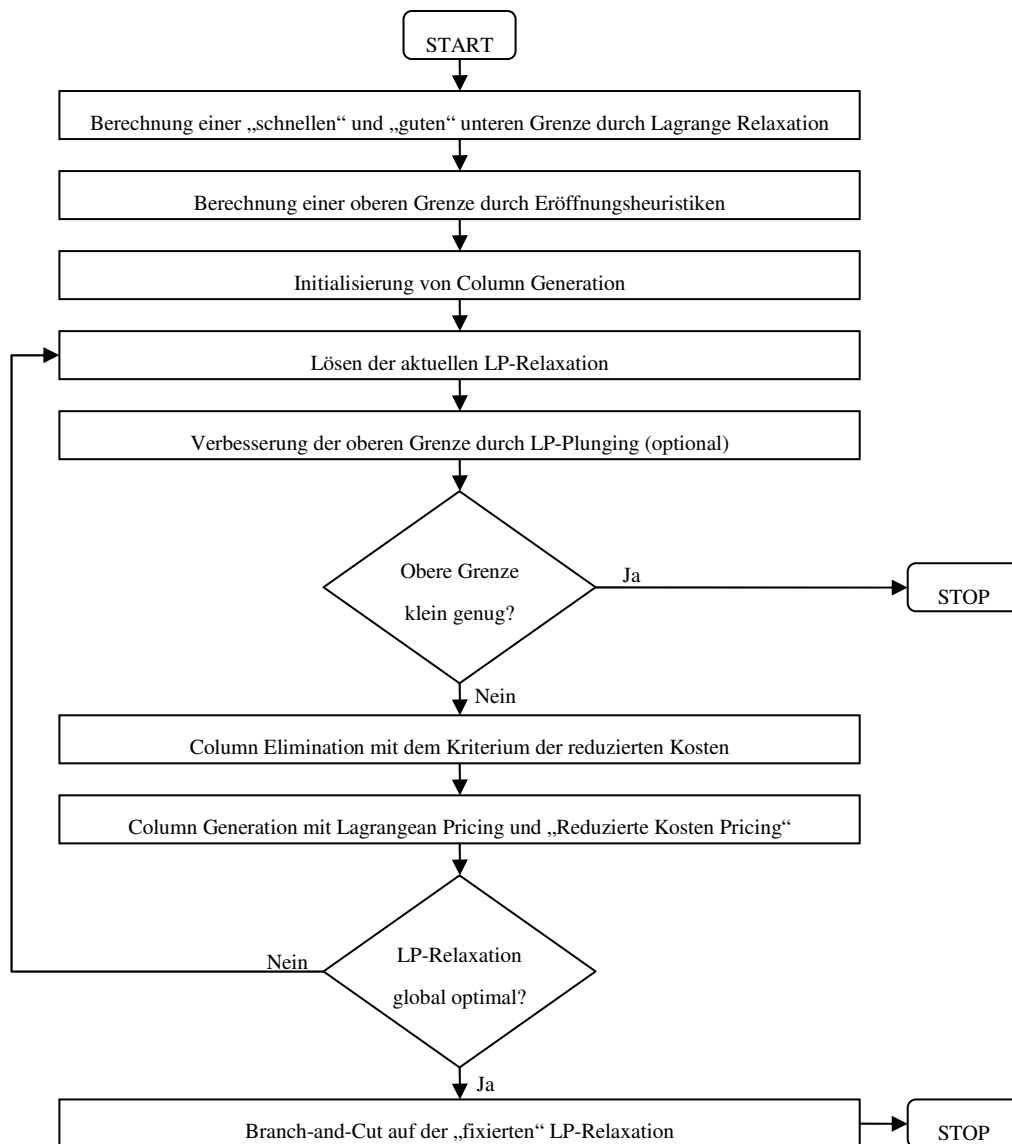


Abbildung 3.9: Ablaufdiagramm für die Lösung des MDVSP nach [Löbel, 1997]

stanzen die Berechnungen unter Umständen gestoppt werden, bevor eine optimale Lösung gefunden worden ist, ist eine Heuristik wichtig, um sehr früh fast optimale Lösungen zu erkennen. Dafür wird eine Heuristik von [Dell'Amico et al., 1993] benutzt, die auf der Berechnung kürzester Pfade beruht.

Der in FORTRAN 77 implementierte Algorithmus wurde auf einer Digital Alpha 533 MHz ausgeführt. Zur Lösung der LP-Relaxationen wurde CPLEX 6.0 verwendet. Es wurde eine Instanz mit zwei Depots und 500 Fahrplanfahrten in durchschnittlich (es wurden jeweils 10 Zufallsinstanzen jeder Größe generiert) 1.160,73 Minuten, eine zweite Instanz mit drei Depots und 400 Fahrplanfahrten in durchschnittlich 3.177,63 Minuten und eine dritte Instanz mit fünf Depots und 300 Fahrplanfahrten in durchschnittlich 3.130,49 Minuten gelöst. Eine Instanz mit realen Daten, die fünf Depots und 580 Fahrplanfahrten besitzt, wurde in 10.000 Minuten gelöst. Nach eigenen Angaben in [Fischetti et al., 1999] ist die verwendete Branch-and-Cut-Methode konkurrenzfähig mit den bis dahin publizierten Algorithmen, wenn zwei bis drei Depots vorhanden sind, was für mittelgroße Busunternehmen relevant wäre.

[Silva et al., 2000] beschreiben einen Kanten-Generierungs-Ansatz (engl.: arc generation approach), um eine Netzwerk-Repräsentation des SDVSP zu lösen. Das Prinzip ist dem des Column Generation ähnlich. Um eine Initillösung zu erhalten, werden nur einige wenige Kanten des Netzwerks beachtet, die mit hoher Wahrscheinlichkeit benutzt werden. Anhand der dualen Preise der nicht berücksichtigten Kanten werden Kanten ermittelt und dem Problem hinzugefügt, die den Ziel funktionswert verbessern können. Die Lösung wird so oft wiederholt, bis es keine Kante mehr gibt, die die vorherige Lösung verbessern kann. Um diese Optimierungsprobleme zu lösen, wird der *Out-of-Kilter-Algorithmus* verwendet. Es wurden Eindepot-Probleminstanzen mit realen Daten gelöst. Die größte gelöste Instanz der Stadt Reading (UK) weist 1.518 Fahrplanfahrten auf, während die größte gelöste Instanz der Stadt Sorocoba (Brasilien) 2.732 Fahrplanfahrten aufweist. Leider werden keine Angaben zur benötigten Rechenzeit gemacht. Es wird nur erwähnt, dass die Instanzen in einer angemessenen Rechenzeit gelöst werden können.

[Banihashemi and Haghani, 2000] geben eine exakte und zwei heuristische Lösungsprozeduren für das MDVSP mit Umlauflängebeschränkung an. Durch die Einbeziehung der Zeitbeschränkungen, ist die Anzahl der Bedingungen, die sicherstellen, dass ein Block seine maximal zulässige Routenzeit nicht überschreitet, zu hoch. Daher wird ein Constraint-Generation-Ansatz angewendet. Dem Problem werden also iterativ Routenzeitbedingungen hinzugefügt. Die Prozedur startet mit dem optimalen Lösen des MDVSP, das allerdings nicht genauer beschrieben wird. Dann werden die gefundenen Blöcke konstruiert, und es wird überprüft, ob sie die maximale Routenzeit überschreiten. Ist das nicht der Fall, ist bereits eine opti-

male Lösung für das MDVSPRTC gefunden worden. Andernfalls gibt es in jedem verletzten Block mit p Fahrplanfahrten $p + 1$ Variablen in der Lösung, die den Wert eins haben. Dementsprechend wird für jeden dieser Blöcke eine Bedingung zum aktuellen Problem hinzugefügt, die sicherstellt, dass höchstens p Variablen des entsprechenden Blocks den Wert eins haben. Diese Prozedur wird so oft wiederholt, bis eine optimale Lösung ohne Verletzungen der Routenzeitbedingungen gefunden worden ist. Die Berechnungen liefen auf einem PENTIUM Pro 200 PC mit 64 MB RAM mit ILOG CPLEX. Es wurden vier Probleminstanzen mit jeweils vier Depots und 400 Fahrplanfahrten in 47 bis 463 Minuten exakt gelöst. Die Lösung einer fünften Probleminstanz wurde nach 60 Iterationen und 4.230 Minuten abgebrochen.

[Freling et al., 2001] verwenden zur Lösung des *Ein-Depot*-Umlaufplanungsproblems einen speziellen Zuordnungsalgorithmus - den kombinierten Vorwärts- und Rückwärts-Auktions-Algorithmus aus [Bertsekas and Castañon, 1992]. Außerdem wird ein zwei-Phasen-Algorithmus beschrieben, in dem zuerst Blöcke gebildet werden, in denen ein Fahrzeug vom Depot startet, Fahrplanfahrten durchführt und wieder zum Depot zurückkehrt, und danach die Blöcke zu Umläufen kombiniert werden. Die Algorithmen sind in C programmiert und auf einem Pentium Processor mit 400 Mhz und 64 MB getestet worden. Es wurden Zufallsinstanzen wie in [Dell'Amico et al., 1993] generiert. Dabei sind Eindepot-Instanzen mit 1.500 Fahrplanfahrten in durchschnittlich unter einer Minute gelöst worden. Weitere Eindepot-Instanzen mit realen Daten und bis zu 1.328 Fahrplanfahrten wurden sekundenschnell gelöst.

3.4.2 Heuristische Ansätze

Heuristische Lösungsansätze werden meistens benutzt um Probleminstanzen lösen zu können, die aufgrund ihrer Größe oder wegen Berücksichtigung zusätzlicher Anforderungen, die das Problem schwieriger machen, nicht in annehmbaren Zeit oder gar nicht exakt optimal lösbar sind. So wurde in [Daduna and Mojsilovic, 1988] ein heuristischer Ansatz beschrieben, der nach dem Prinzip Schedule first - Cluster second vorgeht. Dabei wird das Entscheidungsproblem der Mehrdepot-Umlaufplanung in zwei Phasen aufgeteilt - die Bildung von Fahrtenketten und die Depotzuweisung, was zu einfacher zu lösenden Modellen mit Näherungslösungen führt.

[Dell'Amico et al., 1993] präsentieren für das MDVSP einen auf dem Ansatz der kürzesten Wege basierenden heuristischen Algorithmus. In jeder Phase des Algorithmus wird ein neuer Umlauf ermittelt. Dabei wird eine Menge von verbotenen Kanten im Netzwerk bestimmt und danach mit einem Kürzeste-Wege-Algorithmus ein zulässiger Zyklus im Netzwerk gesucht, in dem keine der verbotenen Kanten

benutzt wird. Eine Kante ist verboten, wenn ihre Einbeziehung in eine zulässige Lösung bedeuten würde, dass sich die Fahrzeuganzahl, die benötigt wird, um alle Fahrplanfahrten abzudecken, um eins erhöhen würde. Für das Finden dieser Kanten wird ein Algorithmus beschrieben, der einen Zeitaufwand von $O(n^3)$ benötigt, wobei n die Anzahl der Knoten im Graphen ist.

Zusätzlich werden vier Verbesserungsverfahren vorgestellt, die einen polynomiellen Zeitaufwand benötigen. Diese Verfahren beruhen jeweils auf der Idee, ein Unterproblem exakt zu lösen. Die erste Verfahren ordnet alle Fahrplanfahrten, die einem Depot zugeordnet sind, neu den Fahrzeugen zu und löst somit das SDVSP exakt. In der zweiten Verfahren werden die Umläufe neu den Depots zugeordnet. Mit der dritten Verfahren werden Fahrplanfahrten, die zwischen zwei vorgegebenen Zeitpunkten beginnen und enden, neu den Depots bzw. den Umläufen zugeordnet. In der vierten Verfahren werden die Fahrplanfahrten aller Paare von Umläufen, die unterschiedlichen Depots zugeordnet sind, neu den Umläufen zugeordnet, wobei dynamische Programmierung benutzt wird. Es wird vorgeschlagen, zuerst die erste Verfahren, danach die dritte Verfahren, dann die vierte Verfahren und schließlich die zweite Verfahren durchzuführen. Dieser Vorgang sollte wiederholt werden, bis keine Verbesserung des Zielfunktionswertes mehr erreicht werden kann.

Die Algorithmen wurden in FORTRAN 77 implementiert und auf einer Digital VAX-station 3100 unter VMS 5.3 ausgeführt. Es wurden Testinstanzen zufällig erzeugt, wobei versucht wurde, reale Daten zu simulieren. Eine Probleminstanz mit zehn Depots und 1.000 Fahrplanfahrten wird in durchschnittlich 9.037,6 Sekunden berechnet.

[[Banihashemi and Haghani, 2000](#)] geben neben einer exakten Lösungsverfahren für das MDVSPRTC zwei heuristische Lösungsverfahren an. Wie schon die exakte Lösungsverfahren, so sind auch die beiden heuristischen Lösungsverfahren Constraint-Generation-Ansätze.

Die erste heuristische Lösungsverfahren, ist eine etwas abgeänderte Form der exakten Lösungsverfahren. Statt nach dem Identifizieren „verletzter“ Blöcke, Bedingungen zum Problem hinzuzufügen, die so einen Block ausschließen, werden in jedem „verletzten“ Block solange die jeweils letzten Fahrplanfahrten gestrichen, bis der Block die Routenzeitbedingungen nicht mehr verletzt. Es werden weitere Bedingungen gebildet und dem Problem hinzugefügt, die dafür sorgen, dass die Fahrplanfahrten dieses Blocks durchgeführt werden. Mit dieser Verfahren wurden dieselben fünf Probleminstanzen mit jeweils vier Depots und 400 Fahrplanfahrten wie mit der exakten Verfahren gelöst. Die Lösungen wurden in 47 bis 322 Minuten errechnet. Die Rechenzeiten sind besser oder genau so gut wie bei der exakten Lösungsverfahren.

Basierend auf der ersten Lösungsverfahren wird in Kombination mit den in der

Modellbeschreibung angegebenen Techniken zur Verringerung der Problemgröße eine Lösungsstrategie angeben, um größere Probleme mit realen Daten zu lösen. Dabei wird eine Instanz mit Daten aus Baltimore City in ca. 23 Stunden gelöst. Diese Probleminstanz hat vier Depots und ursprünglich 5.650 Fahrplanfahrten, die allerdings durch das Zusammenfügen von einigen Fahrplanfahrten auf 2.218 Fahrplanfahrten reduziert wurden, was einen Großteil der Lösungszeit in Anspruch genommen hat (zwölf Stunden).

Mit der zweiten heuristischen Lösungsprozedur wird das Problem ohne Ganzzahligkeitsbedingung gelöst. In jeder Iteration werden dem Problem zwei Mengen von Bedingungen hinzugefügt, wobei zusätzlich eine minimale Routenzeit beachtet wird. Erstens werden Variablen von Blöcken, in denen alle Variablen den Wert eins haben, in dem Problem durch folgende Bedingungen auf den Wert eins fixiert:

- Liegt die Routenzeit des Blocks zwischen der minimalen und der maximalen Routenzeit, so werden alle Variablen dieses Blocks auf den Wert eins fixiert.
- Überschreitet die Routenzeit des Blocks die maximale Routenzeit, so werden solange die jeweils letzten Fahrplanfahrten des Blocks gelöscht, bis die Routenzeit im zulässigen Bereich liegt. Die verbleibenden Variablen werden auf den Wert eins fixiert.
- Unterschreitet die Routenzeit des Blocks die minimale Routenzeit, so werden alle Variablen dieses Blocks auf den Wert eins fixiert. Nur die letzte Variable, die mit der Einrückfahrt des Blocks korrespondiert, wird nicht auf den Wert eins fixiert.
- Ist die Anzahl der Blöcke, in denen alle Variablen den Wert eins haben, gleich der Anzahl aus der vorherigen Iteration, so werden die Variablen eines Blocks, in dem auch nicht ganzzahlige Variablen enthalten sind, nach denselben Regeln auf den Wert eins gesetzt.

Zweitens wird bei einer nichtganzzahligen Lösung eine Restriktion hinzugefügt, die verhindern soll, dass dieselben nicht ganzzahligen Variablen in der nächsten Iteration generiert werden. Es kann passieren, dass das MDVSP unlösbar wird, wenn die Differenz zwischen der Anzahl der verfügbaren Fahrzeuge und der Anzahl der in der exakten Lösung benötigten Fahrzeuge sehr gering ist. Dann müssen die in der letzten Iteration dem Problem hinzugefügten Bedingungen eliminiert werden und die Variablen, die das Problem unlösbar machen, auf den Wert null gesetzt werden. Mit dieser zweiten heuristischen Lösungsprozedur werden Instanzen mit vier Depots und 500 Fahrplanfahrten gelöst. Die Rechenzeit betrug zwischen 89 und 893 Minuten. Ein Vergleich für die Probleminstanzen mit vier Depots und

400 Fahrplanfahrten, die auch mit der exakten und der ersten heuristischen Prozedur gelöst wurden, zeigt, dass die zweite heuristische Prozedur schneller ist und meistens bessere Ergebnisse errechnet.

[Huisman et al., 2001] bieten einen heuristischen Lösungsansatz für das dynamische MDVSP, da eine exakte Lösung nicht in angemessener Rechenzeit erzielt werden kann. Es wird eine Cluster-Reschedule-Heuristik benutzt. Dabei werden zuerst alle Fahrplanfahrten den Depots zugeordnet, indem das statische Problem gelöst wird. Dann werden die dynamischen SDVSP gelöst. Zur Berechnung einer unteren Grenze wird Lagrange Relaxation benutzt. Eine Besonderheit des Lösungsansatzes für das dynamische MDVSP ist, dass jeweils eine Sequenz von Optimierungsproblemen gelöst werden muss. Dadurch können die möglichen Verspätungen gut berücksichtigt werden.

Auf einem Pentium 3 mit 450 MHz und 128 MB RAM wurden Testinstanzen mit realen Daten einer niederländischen Busgesellschaft gelöst. Das Problem hat 1.104 Fahrplanfahrten mit durchschnittlich 1,71 Depots pro Fahrt. Mit dem statischen MDVSP ist eine Lösung in 64 Sekunden berechnet worden. Das dynamische MDVSP wird mit unterschiedlichen Parametereinstellungen gelöst. Für den besten Lösungswert werden 6.088 Sekunden Rechenzeit benötigt.

3.5 Motivation für eine alternative Modellierung

Die Motivation für eine alternative Modellierung ergibt sich aus mehreren Gründen. Die betrachteten Arbeiten und die Testergebnisse zeigen, dass die Komplexität von MDVSP insbesondere die sehr großen Testinstanzen schwer lösbar macht. Deshalb existieren auch unterschiedliche Ansätze, um dieses Problem zu lösen. Teilweise sind fortgeschrittene Ansätze entwickelt worden, um spezielle Restriktionen wie eine maximale Umlauflänge oder eine Minimierung der Anzahl der Verspätungen zu beachten. Allerdings ist es schwierig, alle möglichen betrieblichen Vorgaben zu berücksichtigen, da die Verfahren meistens stark spezialisiert sind und keine Modellerweiterungen zulassen.

Für die Modellierung eines MDVSP scheint eine Mehrgüter-Fluss-Modellierung besonders geeignet zu sein, da in diesem Modell zum Beispiel unterschiedliche Fahrzeugtypen berücksichtigt werden können. Außerdem kann ohne Probleme gewährleistet werden, dass ein Umlauf in demselben Depot endet, von dem aus er auch gestartet ist. Bei der Belegung der Kanten mit entsprechenden Kosten, ist es wichtig darauf zu achten, dass Kanten, die während eines Umlaufs zum Depot und wieder zurück zur nächsten Fahrplanfahrt führen, nicht mit den Fixkosten oder einem ähnlich hohen Wert belegt werden. Denn sonst wird ein zwischenzeit-

liches Parken eines Fahrzeugs im Depot während eines Umlaufs verhindert. Allerdings geben nur [Löbel and Strubbe, 1996], [Grötschel et al., 1997], [Löbel, 1997], [Banihashemi and Haghani, 2000] und [Huisman et al., 2001] entsprechende Kostendefinitionen an, so dass ein zwischenzeitlicher Depotbesuch während eines Umlaufs möglich ist. Eine Beachtung von Pufferzeiten wie in [Silva et al., 2000] zwischen der Ausführung von zwei Fahrplanfahrten scheint sinnvoll zu sein, um bei verspäteten Fahrplanfahrten keinen Dominoeffekt zu erzielen. Aber dadurch werden unter Umständen kostengünstigere Zuordnungen von Fahrzeugen zu Fahrplanfahrten verhindert.

Die in diesem Kapitel betrachteten Modellierungsansätze zu MDVSP unterscheiden sich im Aufbau der Netzwerkmodelle, besitzen aber eine gemeinsame Eigenschaft - alle möglichen Fahrtenverknüpfungen werden explizit durch Kanten im Netzwerk repräsentiert (daher wird im Folgenden der Sammelbegriff *connection-basierte Modelle* benutzt). In einigen Ansätzen wurde diese Kantenmenge durch unterschiedliche Heuristiken reduziert, um die Modellgröße in Grenzen zu halten. Solche Reduktionen engen aber den Lösungsraum des Problems ein, so dass die Optimalität der gefundenen Umlaufpläne nicht mehr garantiert werden kann.

Die vorliegende Arbeit soll einen alternativen Modellierungs- und Lösungsansatz entwickeln, der die beschriebenen Nachteile existierender Modelle vermeidet. Als eine Alternative zur Modellierung und Lösung des MDVSP mit den betrachteten connection-basierten Modellen wird im nächsten Kapitel ein Time-Space-Netzwerk-basiertes Modell vorgeschlagen (beschrieben auch in [Kliwer et al., 2005], [Kliwer et al., 2002] und in [Mellouli and Kliwer, 2002]). Der Vorteil gegenüber der klassischen Modellierung des MDVSP ist, dass deutlich weniger Kanten benötigt werden. Bei der alternativen Modellierung werden folgende Zielsetzungen verfolgt:

1. Größere Probleminstanzen direkt mit Hilfe mathematischer Optimierungssoftware lösbar zu machen, ohne, wie das in den beschriebenen Ansätzen der Fall ist, spezielle Techniken entwickeln zu müssen.
2. Den absehbaren kontinuierlichen Fortschritt in den Optimierungsbibliotheken in den Dienst der Umlaufplanung zu stellen.
3. Einen flexiblen Ansatz zu entwickeln, mit dem sich spezielle Nebenbedingungen wie Kapazitäten, Outsourcing von Fahrplanfahrten, Fahrzeugankünfte in einem anderen Depot u.a. ohne großen Zusatzaufwand im Modell berücksichtigen lassen.

Kapitel 4

TSN-basierte Modellierung und Lösung

In diesem Kapitel wird alternativ zu den MDVSP-Ansätzen, die in Kapitel 3 beschrieben wurden, ein neuer Ansatz entwickelt. Der Ansatz ermöglicht eine Modellierung und Lösung der Umlaufplanungsaufgabe mit mehreren Depots, mehreren Fahrzeugtypen, mit Vorgaben von Fahrzeugtypgruppen für jede Fahrplanfahrt und von Kapazitäten verschiedener Art. Minimiert werden die Gesamtkosten des Umlaufplans, die aus den fixen Fahrzeugkosten (oder Fahrzeuganzahl) und operativen Einsatzzeit- und Kilometerkosten bestehen.

Die Modellierung basiert auf einem Time-Space-Netzwerk und erlaubt eine anschließende exakte Lösung mit einer konventionellen Branch-and-Cut Methode. Die Modellierung beruht auf der Idee, die Time-Space-Netzwerk-Technik, die bisher im Bereich Airline Scheduling üblich war, zu erweitern und auf das MDVSP anzuwenden. Dafür wird diese Technik um eine Berücksichtigung der für das ÖPNV-Bereich spezifischen Leerfahrtsverbindungen (*engl.: deadheads*), Depots, Fahrzeugtypen und Fahrzeugtypgruppen erweitert.

Die auf einer ähnlichen Idee basierende Modellierung ist zuvor bereits im Bereich der Umlaufplanung im Flugverkehr eingesetzt worden (z.B. in [Hane et al., 1995] oder in [Mercier and Soumis, 2005]). Ein Vorteil gegenüber der klassischen Modellierung des MDVSP mit einem Netzwerk-Fluss-Modell ist, dass deutlich weniger Kanten benötigt werden. Dies soll zu einer Reduktion der Größe des mathematischen Modells führen, so dass es mit Hilfe mathematischer Optimierungssoftware gelöst werden kann. Durch Verwendung universeller Optimierungsmethoden wird eine Berücksichtigung spezieller Nebenbedingungen ohne einer aufwändigen Anpassung des Lösungsverfahrens möglich.

Dieser Modellierung liegt das Transitivitäts-Prinzip für die Kompatibilität von

Fahrplanfahrten zugrunde: ist die Fahrplanfahrt B zeitlich kompatibel zu der Fahrplanfahrt A (kann von einem Fahrzeug nach A ausgeführt werden), und die Fahrplanfahrt C kompatibel zu B , dann muss C auch kompatibel zu A sein. Dann muss die Kompatibilität von C zu A nicht mehr, wie in den klassischen Modellen aus dem vorherigen Kapitel, explizit durch Kanten modelliert werden, sondern ist implizit durch den Weg von A zu C über B im Modell enthalten.

Der Hauptvorteil des TSN-Modells ist eine Reduktion der Anzahl von Variablen im mathematischen Modell. Nachteilig für manche Anwendungsfälle könnte die fehlende Möglichkeit sein, spezielle Kostenstrukturen, die auf individuellen Verknüpfungen von Fahrten basieren, zu berücksichtigen. Im Fall mit linearen Kostenstrukturen ist jedoch eine individuelle Berücksichtigung von Fahrtenverknüpfungen nicht notwendig, und daher ist TSN-Modell den traditionellen connection-basierten Modellen aus dem Kapitel 3 vorzuziehen.

Die Hauptunterschiede zu den Time-Space-Netzwerken aus der Flugplanung bestehen hier darin, dass

- die Kanten für Verbindungsfahrten eingeführt werden müssen und
- die Fahrzeugumläufe den Depots und den Fahrzeugtypen zuzuweisen sind.

Dies führt zu einer Aufgabe, die schwieriger zu modellieren und - durch eine höhere kombinatorische Komplexität - auch schwieriger zu lösen ist. Mit wachsender Problemgröße wächst die Lösungszeit exponentiell. Die Problemgröße wird durch die Anzahl der möglichen Verbindungsfahrten beeinflusst, die in den klassischen verbindungsbasierten Modellen quadratisch zu der Anzahl der Servicefahrten wächst (vgl. Kapitel 3). In dem Time-Space-Netzwerk-basierten Modell kann die Anzahl der entsprechenden Kanten jedoch, wie im Folgendem beschrieben wird, stark reduziert werden.

4.1 Modellierung

Nachfolgend werden zunächst alle relevanten Aspekte der Netzwerkmodellierung wie die Schichtenstruktur, die Mengen der Knoten und Kanten sowie die Reduktion der Anzahl von Verbindungen dargestellt. Nach der anschließenden Formulierung des mathematischen Modells für das Netzwerkflussproblem wird noch ein einfaches Beispiel für den Aufbau des Netzwerkmodells gegeben.

4.1.1 Das Netzwerk

Schichtenstruktur und Kanten

Im Time-Space-Netzwerk wird jede erlaubte Fahrzeugtyp-Depot-Kombination als eine Netzwerkschicht dargestellt. Eine solche Schicht beinhaltet unter anderem Kanten für alle Service-Fahrten, die mit dem entsprechenden Fahrzeugtyp bedient werden können. Das gesamte Netzwerkmodell besteht somit aus mehreren Netzwerken (im Weiteren: Netzwerk-Schichten), die weder gemeinsame Kanten noch gemeinsame Knoten besitzen. Die Entscheidungssituation wird also als ein Netzwerkflussproblem abgebildet, bei dem zusätzlich zu den normalen minimalen Kosten-Fluss-Restriktionen Cover Constraints berücksichtigt werden soll. Wäre nur eine Schicht vorhanden, wie es bei Ein-Depot-Problemen mit homogener Fahrzeugflotte der Fall ist, würde man auf den Fahrt-Kanten sowohl die untere, als auch die obere Flussschranke auf eins setzen, und die Cover Constraints wären somit überflüssig. Also hätte man ein gewöhnliches min cost flow Problem, das mit Standard-Netzwerkalgorithmen gelöst werden kann. Die entsprechenden Algorithmen sind z.B. in [Chvatal, 1983] und in [Ahuja et al., 1983] beschrieben.

Im Mehrdepot-Fall und/oder falls die Fahrzeugmenge heterogen ist, beinhaltet eine Netzwerkschicht Representationen aller Fahrzeugaktivitäten, die von den Fahrzeugen entsprechenden Typs aus dem entsprechenden Depot ausgeführt werden dürfen. Die einzelnen Wege mit Start und Ende in einem Depot stellen die zulässigen Fahrzeugumläufe dar.

Die Kanten einer Netzwerkschicht stellen mögliche Fahrzeugaktivitäten für gleichwertige, austauschbare Fahrzeuge dar. Es können folgende Aktivitäten sein:

- Ausführen einer Servicefahrt,
- Standzeit an einer Haltestelle,
- Ortswechsel durch eine Leerfahrt,
- Ausrücken vom Depot zu einem Einsatzort,
- Einrücken ins Depot am Ende eines Betriebstages oder für eine Pause im Laufe des Tages,
- eine Übernachtung im Depot.

Dementsprechend gibt es mehrere Arten von Kanten in jeder Netzwerk-Schicht. Diese sind an den Zeit- oder Anschlusslinien (Timelines) verankert, die die Haltestellen und die Depots für einen ganzen Tag im Netzwerkmodell präsentieren.

Servicefahrt-Kanten - repräsentieren jede mögliche Fahrplanfahrt, die mit dem dieser Schicht entsprechenden Fahrzeugtyp bedient werden kann. Die obere Flussschranke ist gleich eins, da jede Fahrplanfahrt höchstens von einem Fahrzeug bedient werden darf. Die untere Schranke ist gleich null, da die Fahrt eventuell in einer anderen Netzwerkschicht bedient wird. Die Kosten auf dieser Kante entsprechen den Kosten für die Ausführung dieser Fahrt mit dem besagten Fahrzeugtyp.

Warte-Kanten (Stand-Kanten) - modellieren die Möglichkeit für ein Fahrzeug an einer Haltestelle zu warten. Eine solche Kante wird mit den Kosten der Standzeit eines Fahrzeugs entsprechenden Typs versehen. Der Fluss auf diesen Kanten ist unbegrenzt. Im Normalfall betragen die Kosten die Kosten der Standzeit eines Fahrzeugs entsprechenden Typs für die Dauer der Wartezeit. Für den Sonderfall, dass die Haltestelle ein Depot ist, wird die Warte-Kante mit keinen Kosten belegt. Dies modelliert die Begünstigung einer Standzeit in einem Depot gegenüber einer normalen Endhaltestelle, da dort zum Beispiel bessere Ruhemöglichkeiten für die Busfahrer vorhanden sind, neue Einsatzunterlagen geholt werden können oder ein Fahrerwechsel möglich ist. Des Weiteren blockieren die im Depot verweilenden Fahrzeuge nicht die Haltestellen.

Verbindungsfahrt-Kanten (Leerfahrtkanten) - ermöglichen eine Verknüpfung kompatibler Fahrten, die in verschiedenen Haltestellen enden bzw. starten. Die Kosten auf diesen Kanten setzen sich aus den Kosten der entsprechenden Leerfahrt zzgl. der Kosten der notwendigen Standzeit eines Fahrzeugs entsprechenden Typs. Die Fahrtzeit sowie die Streckenlänge sind dabei abhängig von der Tageszeit. Später wird gezeigt, dass aufgrund der Modellierungsart diese Verbindungsfahrtkanten von mehreren Fahrzeugen verwendet werden dürfen. Also ist der maximale Fluss für diese Kanten unbeschränkt.

Depot-Kanten - werden vom Depot zu dem Anfang jeder Servicefahrt-Kante und vom Ende jeder Servicefahrt-Kante zum Depot eingefügt (s. auch die Abbildung 4.1). Diese Kanten bilden sowohl die Einsetz-/Aussetzfahrten am Anfang und am Ende eines Umlaufs, als auch die Depot-Einrucker und Depot-Ausrücker im Laufe des Tages. Des Weiteren ermöglichen die Depot-Kanten eine Rückfahrt in ein Depot, um lange Wartezeiten an den Haltestellen zu vermeiden. Die Kosten setzen sich wie bei den oben beschriebenen Verbindungsfahrt-Kanten aus den entsprechenden Kilometer- und Stundenkosten zusammen.

Der maximale Fluss für eine Depot-Kante beträgt im Normalfall 1, da für jede Servicefahrt-Kante zwei Depot-Kanten (eine Einrück- und eine Ausrückfahrt)

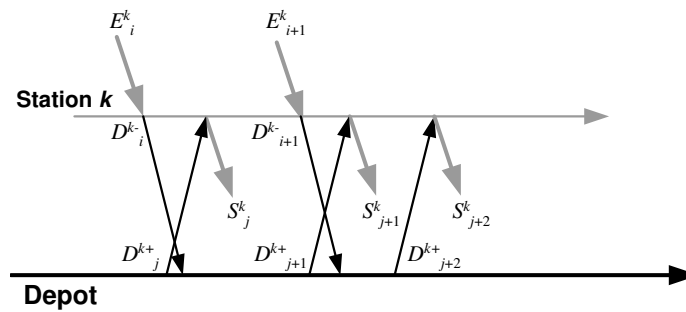


Abbildung 4.1: Timelines mit Kanten für potenzielle Service-, Depotausrück- und Depoteinrückfahrten

gebildet werden. Falls aber mehrere Depot-Kanten die gleichen Abfahrts- bzw. Ankunftszeiten sowie Start- bzw. Zielhaltestellen haben, wird nur eine dieser Kanten im Netzwerk eingefügt und ihr maximaler Fluss entsprechend der Anzahl solcher gleichwertigen Kanten begrenzt. Dieser Schritt ist nicht zwingend erforderlich, ermöglicht aber eine Reduzierung der Modellgröße.

Zirkulationsflusskante - eine pro Netzwerk-Schicht. Die Zeitlinie, die das Depot repräsentiert, hat zusätzlich eine Zirkulationsflusskante vom Ende der Zeitlinie zum Anfang der Zeitlinie. Das Flussvolumen auf dieser Kante gleicht der Anzahl einzusetzender Fahrzeuge der entsprechenden Fahrzeugtyp-Depot-Kombination. Die Kosten der Kante entsprechen den fixen (anteiligen) Anschaffungskosten eines Fahrzeugs entsprechenden Typs. Beispielsweise ist ein Standardbus günstiger als ein Gelenk- oder ein Niederflerbus. Da die Fahrzeugkosten auf dieser Kante viel größer als die sonstigen Kosten sind, benutzt die optimale Flusslösung immer die minimale Anzahl Fahrzeuge. Die obere Flusschranke ergibt sich aus der Kapazität des Depots für den Fahrzeugtyp. Falls diese nicht angegeben ist, ist der maximale Fluss für die Zirkulationsflusskanten unbeschränkt.

Die Gesamtkosten des Flusses ergeben sich als Summe der Kostenwerte auf den Kanten des Netzwerks, die mit den Flussgrößen auf den entsprechenden Kanten multipliziert werden. Die Fixkosten für die Fahrzeuge werden auf den Zirkulationsflusskanten der Depots beachtet. Da das Warten im Depot gegenüber dem Warten in einer Station bevorzugt werden soll, erhalten die Wartekanten der Depots niedrigere Kosten als die Wartekanten in den Stationen. Die Kanten, die die anderen Leerfahrten präsentieren, werden ebenfalls mit entsprechenden operativen Kosten belegt. Da in der Zielfunktion die Kosten minimiert werden, wird eine Lösung mit einer minimalen Anzahl an Fahrzeugen errechnet, wenn die Fixkosten auf den Zirkulationsflusskanten entsprechend hoch angegeben werden.

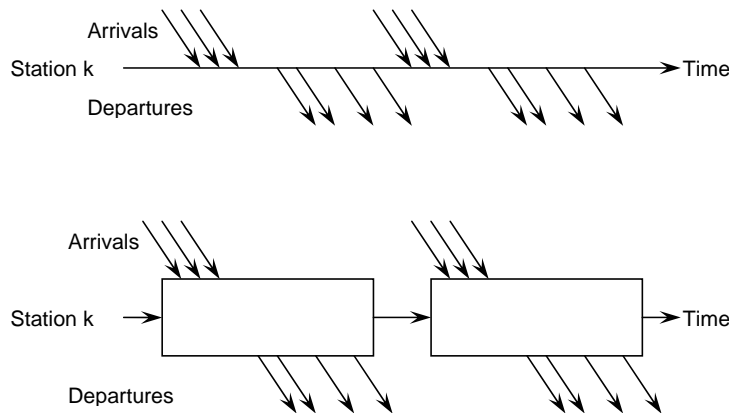


Abbildung 4.2: Anschlussknoten als aggregierte Gruppen von Ankünften und unmittelbar darauffolgenden Abfahrten

Anschlussknoten

Die Anschluss- oder Zeit-Ort-Knoten entsprechen einem Zeitraum oder einem Zeitpunkt an einem Ort (Haltestelle oder Depot). Sie aggregieren eine Menge potenzieller Anschlüsse, die gleichzeitig an einem Ort geschehen dürfen.

Ein Knoten gehört zu einer Haltestelle und wird beim Durchlaufen der Timelines gebildet, indem eine Gruppe der Abfahrten mit der vorangehenden Gruppe der Ankünfte vereinigt und ggf. mit der nächsten Gruppe durch Wartekanten verbunden wird. Die Abbildung 4.2 illustriert dieses Vorgehen. Somit sind Knoten verschiedener Stationen durch Servicefahrt- und Verbindungskanten und in gleicher Station durch Wartekanten verbunden. (Man bedenke dabei, dass die Fahrtkanten eine Zeit- und Ort-Transformation, die Wartekanten aber nur eine Zeit-Transformation verursachen.) Der erste bzw. der letzte Knoten einer Haltestelle enthält keine eingehenden bzw. ausgehenden Wartekanten. Im Depot ist der letzte Knoten mit dem ersten durch die Zirkulationsflusskante verbunden.

Aufbau und Reduktion von Verbindungsfahrt-Kanten

Die Konstruktion der Kanten für Verbindungsfahrten ist nicht trivial und verdient eine ausführlichere Erklärung. Im Vergleich zu den klassischen Modellen kann in einem Time-Space-Netzwerk-basierten Modell die Anzahl der entsprechenden Kanten reduziert werden, wie es im Folgenden gezeigt wird.

Zwei Fahrten sind kompatibel, wenn sie von einem Fahrzeug nacheinander aus-

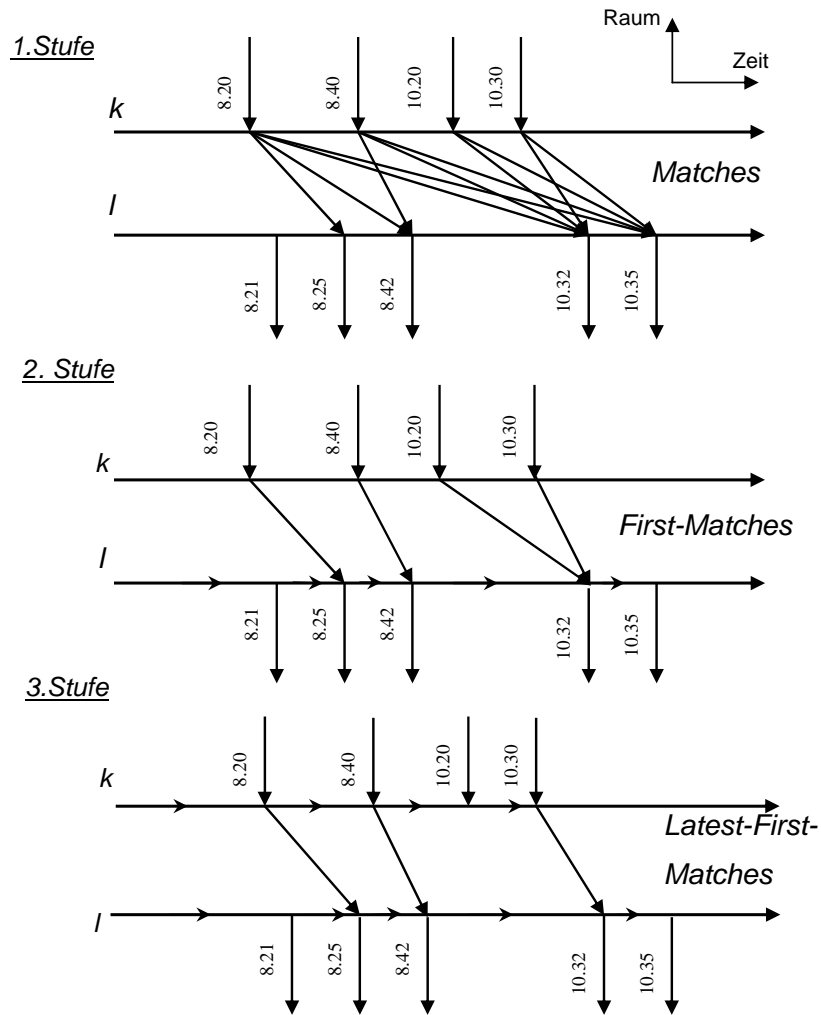


Abbildung 4.3: Eliminierung redundanter Verbindungsfahrt-Kanten

geführt werden können. Betrachten wir den Fall, wo ein Fahrzeug zwischen zwei kompatiblen Fahrten eine Leerfahrt ausführen müsste. Die Abbildung 4.3 zeigt einen Teilausschnitt aus einem größeren Netzwerk mit zwei Stationen k und l .

In der **ersten Stufe** sind alle möglichen Leerfahrten-Verknüpfungen zwischen den in der Station k ankommenden und den in der Station l abfahrenden Service-Fahrten gezeigt, wie sie in den klassischen Modellen aus dem Kapitel 3 benutzt werden.

Die erste in Haltestelle l abfahrende Fahrt (Abfahrt um 8:21) ist nicht kompatibel zu der ersten in k ankommenden Fahrt (Ankunft um 8.:20). Das heißt, ein Bus würde es zeitlich nicht schaffen, nachdem er in k angekommen war, zu l zu fahren und da die erste Fahrt auszuführen. (Anmerkung: das Bild stellt ja nur einen Ausschnitt aus dem Netzwerk dar, die Depot-Leerfahrten sind nicht eingezeichnet.

Diese „verwaiste“ erste Fahrt in Haltestelle l kann mit einem Bus bedient werden, der von einer anderen Station kommt oder direkt vom Depot ausrückt.)

Kompatibel zu der Fahrt, die in der Haltestelle k um 8:20 ankommt, ist beispielsweise die Fahrt, die um 8.25 in der Haltestelle l abfährt. Dementsprechend sind in der ersten Stufe diese beiden Fahrten durch eine Verbindungsfahrt-Kante verbunden und können so von einem Fahrzeug nacheinander ausgeführt werden. Solche Verbindungen werden nachfolgend *Matches* genannt.

In der **zweiten Stufe** verbleiben von dieser Menge der möglichen Verknüpfungen nur die *First-Matches*, die wie folgt definiert werden: Für eine in der Haltestelle k ankommende Fahrplanfahrt i wird in jeder Haltestelle $l \neq k$ die erste kompatible abfahrende Fahrplanfahrt ermittelt. Die entsprechenden Verbindungsfahrten werden *First-Match*(i, l) genannt.

Das Ausschließen der Matches, die keine First-Matches sind, würde bereits eine deutlich geringere Kantenanzahl als in dem klassischen Modellansatz verursachen. Denn im Gegensatz zu der Anzahl der Matches, hängt die Anzahl derartiger First-Matches nicht mehr quadratisch von der Anzahl der Servicefahrten, sondern beträgt höchstens $n * m/2$ (n und m sind die Anzahl der Servicefahrten bzw. die Anzahl der Haltestellen). Durch diese Reduktion bleiben in der zweiten Stufe alle möglichen Servicefahrten-Verknüpfungen aus der ersten Phase erhalten, da diese über Wartekanten in der Station l weiterhin erreicht werden können.

In der **dritten Stufe** wird die Anzahl der Verbindungsfahrt-Kanten weiter reduziert. Es wird jedes First-Match (i, l) von Station k zu Station l entfernt, wenn es ein anderes späteres First-Match (i', l) von Station k zu derselben Fahrt in Station l gibt. Die übrig gebliebenen Verbindungsfahrten werden *Latest-First-Matches* genannt. Auch in der dritten Stufe bleiben alle bis dahin mögliche Verbindungen zwischen Servicekanten durch nunmehr indirekte Wege (über Latest-First-Matches und eventuelle Wartekanten) erhalten.

Diese Möglichkeit kann am folgenden Beispiel erläutert werden: Drei letzte Ankünfte in der Station k sind mit den zwei letzten Abfahrten in der Station l paarweise kompatibel. Es soll also auch nach der Reduktion durch Wege im Netzwerk möglich sein, den Fluss von jeder dieser drei Ankünfte in k zu jeder dieser zwei Abfahrten in l fließen zu lassen.

Die Redundanz der drei First-Matches ist offensichtlich. Was passiert, wenn man nun, statt der letzten, die erste dieser drei Verbindungen lassen und die beiden letzten löschen würde? Dann würde das Netzwerk-Modell nicht mehr alle möglichen Verknüpfungen zwischen den kompatiblen Fahrten abbilden. Denn es gäbe dann keine Wege von den beiden letzten Ankünften in k zu den beiden letzten Abfahrten in l . Also hätten wir den Lösungsraum, der durch solche Möglichkeiten aufgespannt

ist, reduziert und dadurch, möglicherweise, die optimale Lösung abgeschnitten. Dies passiert nicht, wenn wir die letzte der drei First Matches im Netzwerk lassen und die beiden ersten löschen (bzw. sie werden in der Implementierung erst gar nicht gebildet). In diesem Fall bleiben Wege für alle paarweise Verknüpfungen vorhanden.

Die Latest-First-Matches im Netzwerk ermöglichen durch Wege im Netzwerk alle möglichen Verbindungen zwischen allen Paaren von kompatiblen Fahrten. Es werden lediglich redundante Wege eliminiert. Die Netzwerk-Reduktion besteht also in einer systematischen „Ausdünnung“ der redundanten Kanten.

In einer **vierten Stufe**, die in der Abbildung 4.3 nicht dargestellt ist, werden die Latest-First-Matches ermittelt, bei denen es zeitlich möglich ist, zwischen den entsprechenden beiden Fahrplanfahrten zum Depot zurückzukehren. Für jedes dieser Latest-First-Matches wird überprüft, ob die Kosten für das Latest-First-Match höher sind, als die Kosten für die zwischenzeitliche Rückkehr in das Depot und das anschließende Ausrücken zu der nächsten Fahrplanfahrt. Wenn das der Fall ist, wird die Kante, die dieses Latest-First-Match präsentiert, ebenfalls gestrichen.

Wie später anhand der numerischen Ergebnisse gezeigt wird, konnte in einigen realen Testinstanzen die Anzahl der Kanten, die die Verbindungsfahrten zwischen unterschiedlichen Stationen repräsentieren, auf diese Weise um über 95% reduziert werden (vgl. auch [Kliwer et al., 2002]).

4.1.2 Das Flussmodell

Der Lösungsvektor x_k beschreibt einen Fluss in der Netzwerkschicht k - wie die Flusseinheiten (Fahrzeuge) im ersten Depotknoten starten, sich über das Netzwerk verteilen und im letzten Depotknoten wieder zusammenkommen. Ein Fluss in einem Netzwerk ist ein Bündel von Wegen einzelner Flusseinheiten. Die einzelnen Wege sind die Fahrzeugumläufe und können durch Flussdekomposition aus dem Fluss extrahiert werden. Die Gesamtheit solcher extrahierter Wege ist der gesuchte Umlaufplan. Die Kosten dieses Umlaufplans sind gleich den Kosten des Flusses, aus dem der Umlaufplan extrahiert wurde. (Anmerkung: in unserem Fall ist diese Dekomposition wegen der oben erwähnten Aggregation von Verbindungskanten nicht eindeutig. Insbesondere können aus dem optimalen Fluss mehrere, allesamt optimale Umlaufpläne extrahiert werden. Mehr dazu in Kapitel 5.)

$x_{i,j,k}$ steht also für das Flussvolumen über die Kante ij in der Schicht k . k steht für eine zulässige Fahrzeugtyp-Depot-Kombination, oder auch eine Netzwerkschicht. $c_{i,j,k}$ sind die Kosten der Kante ij - je nach Art der entsprechenden realen Fahrzeugaktivität sind es Fahrtkosten, Standzeitkosten, Leerfahrtkosten oder Kombination dieser. E_k^s ist die Menge aller Fahrplanfahrt-Kanten in der Schicht k . E_k ist die Menge aller Kanten der Netzwerkschicht k .

Es handelt sich um ein Mehrgüterflussproblem (*engl.: multi-commodity flow problem*), in dem mehrere Güter mit minimalen Kosten durch das vorgestellte Netzwerk transportiert werden sollen. Eine Güterart ergibt sich aus Fahrzeugen eines bestimmten Fahrzeugtyps, die in einem bestimmten Depot stationiert sind. Repräsentiert wird eine solche Güterart durch die entsprechende Schicht des Netzwerks. Durch die Art der Konstruktion des Netzwerks ist eine „Vermischung“ der unterschiedlichen Güter ausgeschlossen - ein Fahrzeug, das ein Depot verlassen hat, wird zu demselben Depot zurückkehren.

4.1.3 Das mathematische Modell

Das auf dem oben beschriebenen Time-Space-Netzwerk basierte mathematische Modell wird in folgendem Abschnitt formal definiert. Folgende Definitionen sollen gelten:

H - die Menge aller Haltestellen,

D - die Menge aller Depots (wobei gilt: $D \subseteq H$),

T - die Menge aller Fahrzeugtypen und

F - die Menge aller Servicefahrten.

Für alle Fahrzeugtypen werden die benötigten Kostensätze wie folgt definiert:

c_{std}^t - die Stundenkosten des Fahrzeugtyps $t \in T$,

c_{km}^t - die Kilometerkosten des Fahrzeugtyps $t \in T$ und

c_{fix}^t - die Fixkosten des Fahrzeugtyps $t \in T$.

Zusätzliche Kapazitäten für Depot und Fahrzeugtypen werden ebenfalls im Modell berücksichtigt:

u_d - die Kapazität des Depots $d \in D$,

u_t - die Kapazität des Fahrzeugtyps $t \in T$ und

l_d^t, u_d^t - untere und obere Schranke für die Anzahl der Fahrzeuge vom Typ $t \in T$ im Depot $d \in D$.

Für die Kostenberechnung werden Distanzen zwischen den Haltestellen benötigt. Deshalb seien

d_{ij}^t - die Distanz zwischen der Haltestelle $i \in H$ und $j \in H$ zum Zeitpunkt t und

d_f - die Streckenlänge der Servicefahrt $f \in F$.

Durch Kombination von Fahrzeugtypen und Depots werden die Schichten des Netzwerks gebildet. Dabei werden keine Schichten gebildet, falls für Fahrzeugtyp-Depot-Kombinationen keine Servicefahrt existiert, die von der Kombination bedient werden kann.

Sei $S \subseteq D \times T$ - die Menge aller Schichten im Netzwerk, wobei gelte, dass

$t_s \in T$ - der Fahrzeugtyp der Schicht $s \in S$ und

$d_s \in D$ - das Depot der Schicht $s \in S$ sei.

Knoten und Kanten des Netzwerks werden für jede Schicht definiert:

V_h^s - die Menge aller Knoten der Haltestelle $h \in H$ in der Schicht $s \in S$,

ES^s - die Menge aller Servicefahrt-Kanten in der Schicht $s \in S$,

ES_f^s - die Menge aller Servicefahrt-Kanten der Fahrt $f \in F$ in der Schicht $s \in S$,

ED^s - die Menge aller Depot-Kanten in der Schicht $s \in S$,

EV^s - die Menge aller Verbindungs-Kanten in der Schicht $s \in S$,

EW^s - die Menge aller Warte-Kanten in der Schicht $s \in S$,

ER^s - die Menge aller Zirkulationsfluss-Kanten in der Schicht $s \in S$.

Dann ist E^s - die Menge aller Kanten in der Schicht $s \in S$:

$$E^s = ES^s \cup ED^s \cup EV^s \cup EW^s \cup ER^s$$

und V^s - die Menge aller Knoten in der Schicht $s \in S$:

$$V^s = \bigcup_{h \in H} V_h^s$$

. Für eine Kante e im Netzwerk soll gelten:

ta_e - die Startzeit von $e \in E^s, s \in S$,

te_e - die Endzeit von $e \in E^s, s \in S$ und

$f_e \in F$ - die Servicefahrt, die der Kante $e \in ES^s, s \in S$ entspricht (falls e eine Servicefahrt-Kante ist).

Für eine vereinfachte Darstellung des mathematischen Modells sei die Menge ESG_f als Gesamtheit aller Servicefahrt-Kanten der Servicefahrt $f \in F$ im gesamten Netzwerk definiert:

$$ESG_f = \bigcup_{s \in S} ES_f^s \quad \forall f \in F$$

Für alle Kanten im Netzwerk werden nun Schranken und Kosten definiert. Dabei seien:

c_e - die Kosten von $e \in E^s, s \in S$,

l_e - die untere Schranke von $e \in E^s, s \in S$ und

u_e - die obere Schranke von $e \in E^s, s \in S$.

Die Schranken und Kosten werden nach den folgenden Regeln berechnet:

$$l_e = 0 \quad \forall e \in (ES^s \cup ED^s \cup EV^s \cup EW^s), \forall s \in S \quad (4.1)$$

$$u_e = 1 \quad \forall e \in (ES^s \cup ED^s), \forall s \in S \quad (4.2)$$

$$u_e = u_d^t \quad \forall e \in (EV^s \cup EW^s \cup ER^s), \forall s \in S \quad (4.3)$$

$$l_e = l_d^t \quad \forall e \in ER^s, \forall s \in S \quad (4.4)$$

$$c_e = (te_e - ta_e) \cdot c_{std}^{ts} \quad \forall e \in (ES^s \cup EW^s), \forall s \in S \quad (4.5)$$

$$c_e = (te_e - ta_e) \cdot c_{std}^{ts} + d_{i,j \in H}^{ta_e} \cdot c_{km}^{ts} \quad \forall e \in (ED^s \cup EV^s), \forall s \in S \quad (4.6)$$

$$c_e = c_{fix}^{ts} \quad \forall e \in ER^s, \forall s \in S \quad (4.7)$$

Alle Kanten im Netzwerk werden in 4.1 mit einer unteren Schranke von 0 versehen. Somit ist sichergestellt, dass kein negativer Fluss im Netzwerk möglich ist. Für alle Servicefahrt-, Depot- und Linien-Kanten wird der maximale Fluss in 4.2 auf eine Flusseinheit beschränkt (siehe Abschnitt 4.1.1). Dagegen wird die obere Schranke für der Fluss auf den Verbindungs-, Warte, und Zirkulationsfluss-Kanten durch 4.3 auf u_d^t gesetzt. Der Fluss auf der Zirkulationsflussschranke wird zusätzlich durch 4.4 nach unten beschränkt.

In 4.5 werden die Kosten für Warte- und Servicefahrt-Kanten gesetzt. Dabei werden nur die Stunden-Kosten des Fahrzeugs berücksichtigt (da keine Entfernung zurückgelegt wird bzw. diese bei den Servicefahrt-Kanten nicht berücksichtigt werden muss, da sie auf jeden Fall gefahren wird). In 4.6 werden dagegen für alle Depot- und Verbindungs-Kanten sowohl die Stunden- als auch die Entfernungskosten der Fahrzeuge berücksichtigt. In 4.7 werden die Zirkulationsfluss-Kanten mit den fixen Anschaffungskosten der Fahrzeuge entsprechenden Typs belegt.

Für die Variante der Problemstellung, bei der unabhängig von Betriebskosten nach einem Umlaufplan mit minimaler Anzahl von Fahrzeugumläufen gesucht wird,

werden die Kostenwerte auf Kanten wie folgt gesetzt: alle Kanten bis auf die Zirkulationsflusskanten erhalten einen Kostenwert von null. Jede Zirkulationsflusskante erhält einen Kostenwert von eins. Der Wert der so definierten Zielfunktion entspricht der Anzahl der Fahrzeuge.

Für die Formulierung des MDVSP als ein gemischt-ganzzahliges Programm wird für jede Kante $e \in E^s, s \in S$ im Netzwerk eine ganzzahlige Flussvariable x_e definiert. Dabei stellt der Wert der Variable x_e den Fluss der Kante e dar. Um einen optimalen Lösungsvektor x im oben definierten Time-Space-Netzwerk zu berechnen, wird das mathematische Modell des Mehrdepot-Umlaufplanungsproblems somit wie folgt formuliert:

$$\min \sum_{s \in S} \sum_{e \in E^s} c_e \cdot x_e \quad (4.8)$$

$$\sum_{i|(i,j) \in E^s} x_{ij} - \sum_{i|(j,i) \in E^s} x_{ji} = 0 \quad \forall j \in V^s, \forall s \in S \quad (4.9)$$

$$\sum_{e \in ESG_f} x_e = 1 \quad \forall f \in F \quad (4.10)$$

$$\sum_{s \in S | d_s = d} \sum_{e \in ER^s} x_e \leq u_d \quad \forall d \in D \quad (4.11)$$

$$\sum_{s \in S | t_s = t} \sum_{e \in ER^s} x_e \leq u_t \quad \forall t \in T \quad (4.12)$$

$$l_e \leq x_e \leq u_e \quad \forall e \in E^s, \forall s \in S \quad (4.13)$$

$$x_e \in \mathbb{N} \quad \forall x_e \in \bigcup_{s \in S} E^s \quad (4.14)$$

Mit der Zielfunktion 4.8 werden die Gesamtkosten der Lösung minimiert. Wenn die Minimierung der benötigten Fahrzeuge als Hauptziel angestrebt ist, müssen die Fixkosten der Fahrzeugtypen deutlich höher als die operativen Kostensätze (Stunden- und Kilometerkosten) gewählt werden.

Die Flusserhaltungsgleichungen (*flow-conservation-constraints*) 4.9 stellen sicher, dass jeder Knoten gleich viele ein- und ausgehende Flusseinheiten hat, bzw. dass die Summe aller eingehenden minus der Summe aller ausgehenden Flüsse gleich Null ist.

Durch die Restriktionen 4.10 (*cover-constraint*) wird die Bedienung von jeder Servicefahrt im Fahrplan sichergestellt. Von der Menge aller im Netzwerk vorhandenen Servicefahrt-Kanten einer Fahrt (ESG_f) muss genau eine in der Flusslösung vorhanden sein. Durch diese Restriktion wird überhaupt erst ein Fluss im Netzwerk erzwungen, da ohne sie die Trivial-Lösung mit einem Nullfluss zulässig und somit optimal wäre.

Fahrt	Von	Nach	Abfahrtszeit	Ankunftszeit	Linie	Depots
<i>a</i>	h_1	h_2	1	3	1	d_1
<i>b</i>	h_2	h_1	4	6	2	d_1
<i>c</i>	h_2	h_1	8	10	1	d_1
<i>d</i>	h_2	h_1	9	11	3	d_1, d_2
<i>e</i>	h_2	d_1	13	15	4	d_1
<i>f</i>	h_1	h_2	8	10	5	d_2

Tabelle 4.1: Fahrplan mit Servicefahrten

Falls Kapazitäten in den Depots vorgegeben sind, wird in den Restriktionen 4.11 durch Aufsummierung aller Flüsse über alle einem Depot zugehörigen Zirkulationsfluss-Kanten sichergestellt, dass die Zahl der von dem Depot eingesetzten Fahrzeuge kleiner oder gleich der vorgegebenen Kapazität (u_d) ist. In der gleichen Form werden in den Restriktionen 4.12 für alle Fahrzeugtypen deren zugehörige Zirkulationsfluss-Kanten behandelt und sichergestellt, dass die Summe ihrer Flüsse nicht größer als die Kapazität des Fahrzeugtyps (u_t) ist. Kapazitäten von Depots für bestimmte Fahrzeugtypen (u_d^t) werden implizit durch die Schranken auf den entsprechenden Zirkulationsflusskanten berücksichtigt.

Die Flussgrößen auf allen Kanten werden durch die unteren und oberen Schranken l_e und u_e durch die Bedingung 4.13 begrenzt und die Bedingung 4.14 verlangt schließlich die Ganzzahligkeit des Flussvektors.

4.1.4 Einfaches Beispiel für den Netzwerkaufbau

Im Folgenden werden an einem einfachen Beispiel der Aufbau des Netzwerks und die verschiedenen Arten der Kanten weiter verdeutlicht.

In der Tabelle 4.1 ist ein Beispiel für die Servicefahrten eines Fahrplans gegeben. Der Fahrplan beinhaltet die beiden Haltestellen h_1 und h_2 sowie die beiden Depots d_1 und d_2 . Im Fahrplan sind sechs verschiedene Servicefahrten gegeben, wobei die Servicefahrten *a* und *c* die gleiche, alle anderen unterschiedliche Linien haben. Die Servicefahrten *a* bis *e* können von Fahrzeugen aus dem Depot d_1 befahren werden. Fahrzeuge aus dem Depot d_2 können auch die Servicefahrt *d* befahren sowie als einziges Depot die Servicefahrt *f*.

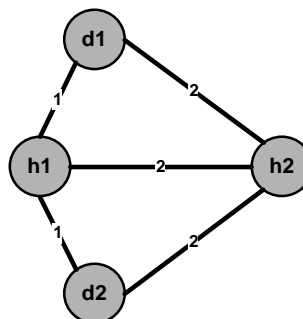


Abbildung 4.4: Netztopologie

Alle Fahrten werden von einem einzigen Fahrzeugtyp bedient. Dies ist so gewählt, damit das Beispiel nicht zu unübersichtlich wird. Sollten verschiedene Fahrzeugtypen behandelt werden, so vergrößert sich die Anzahl der Schichten; an dem Prinzip des Netzwerkaufbaus allerdings ändert sich nichts. Eine topologische Darstellung der Haltestellen ist in der Abbildung 4.4 zu sehen. Die auf den Verbindungen angegebenen Werte stellen die Entfernungen zwischen zwei Stationen sowohl für die Servicefahrten wie auch für alle Leerfahrten (Verbindungs- und Depot-Fahrten) dar. So beträgt die Entfernung von der Haltestelle h_2 zu allen anderen Stationen zwei Streckeneinheiten, alle anderen Entfernungen eine Streckeneinheit.

Anhand des vorgestellten Beispiels soll der Aufbau des Time-Space-Network für die Modellierung des MDVSP illustriert werden. Die einzelnen Phasen des Aufbaus werden im Folgenden erläutert.

Timelines erstellen und Servicefahrt-Kanten einfügen

In der Phase 1 des Netzwerkaufbaus wird für alle Depot-Fahrzeugtyp-Kombinationen eine Netzwerkschicht angelegt. In jeder Netzwerkschicht werden für alle Haltestellen sowie für das der Schicht zugehörige Depot Timelines gebildet. An die Timelines werden im weiteren Netzwerkaufbau alle Kanten geknüpft. Die Timelines dienen zunächst der Fixierung der Kanten im Time-Space-Network. In einem späteren Schritt werden dann aus den Timelines Knoten gebildet sowie diese mit Wartekanten verbunden.

Für jede Servicefahrt im Fahrplan wird je eine Servicefahrt-Kante in jede Netzwerkschicht eingefügt, deren Fahrzeugtyp-Depot-Kombination die Servicefahrt ausführen kann. Diese Kanten werden gemäß ihrer Daten aus dem Fahrplan (Abfahrts- und Endzeit) an die Timelines ihrer Start- und Zielhaltestellen gehängt.

In Abbildung 4.5 sehen wir wie für jede Netzwerk-Schicht Timelines für die Haltestellen des vorliegenden Beispiels sowie für das Depot der jeweiligen Schicht angelegt sind.

Für die Servicefahrten a bis e wird je eine Kante in die Schicht vom Depot 1 eingefügt. In die zweite Schicht werden nur Kanten für die Servicefahrten aufgenommen, die vom Depot 2 befahren werden können. Dies sind in diesem Beispiel Kanten für die Fahrten d und f .

Der Übersicht halber soll in den weiteren Beschreibungen des Netzwerkaufbaus die zweite Schicht nicht mehr betrachtet werden, da alle Schritte anhand der ersten Schicht beschrieben werden können. Der fertige Netzwerkaufbau wird anschließend wieder für beide Schichten dargestellt.

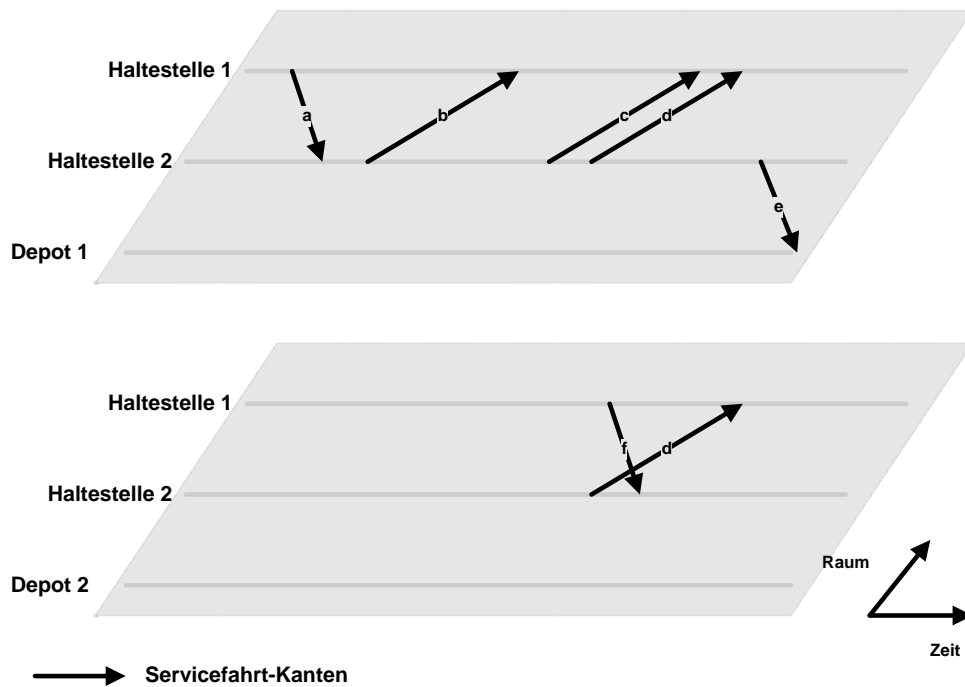


Abbildung 4.5: Erstellen der Timelines und Einfügen der Servicefahrt-Kanten

Einfügen nicht-redundanter Verbindungs-Kanten

In der Phase 2 des Netzwerkaufbaus werden alle Verbindungs-Kanten in das Netzwerk eingefügt. Sie dienen der Möglichkeit, dass Fahrzeuge, die eine Servicefahrt abgeschlossen haben, nicht nur von der Haltestelle aus, an der sie stehen, eine neue Fahrt beginnen können. Vielmehr lohnt es sich in vielen Fällen, eine Leerfahrt in Kauf zu nehmen, mit der man eine andere Haltestelle anfährt, um dort die nächste Servicefahrt zu starten.

Das Einfügen dieser Verbindungs-Kanten wird entsprechend der Beschreibung in Abschnitt 4.1.1 vorgenommen. Dieses Verfahren ermöglicht eine drastische Reduzierung der Modellgröße, so dass auch die Lösung großer Instanzen möglich wird.

Das Verfahren wird zum besseren Verständnis in der Abbildung 4.6 in drei Schritten gezeigt, wobei zu betonen ist, dass in der Implementierung nur die Kanten des letzten Schrittes wirklich dem Netzwerk hinzugefügt werden.

Im oberem Bereich der Abbildung 4.6 ist von jeder Ankunft eine Verbindungs-Kante zu jeder erreichbaren Abfahrt gezeichnet. In der betrachteten Schicht können zum Beispiel von der Ankunft der Servicefahrt *b* zur Zeitpunkt 6 (siehe Tabelle 4.1) alle abgehenden Fahrten der Haltestelle 2 erreicht werden, die zur (oder nach) der Zeiteinheit 8 beginnen, da die Entfernung zwischen den Haltestellen in unserem Beispiel 2 beträgt (siehe Abbildung 4.4). Dies sind in dem vorliegenden Beispiel

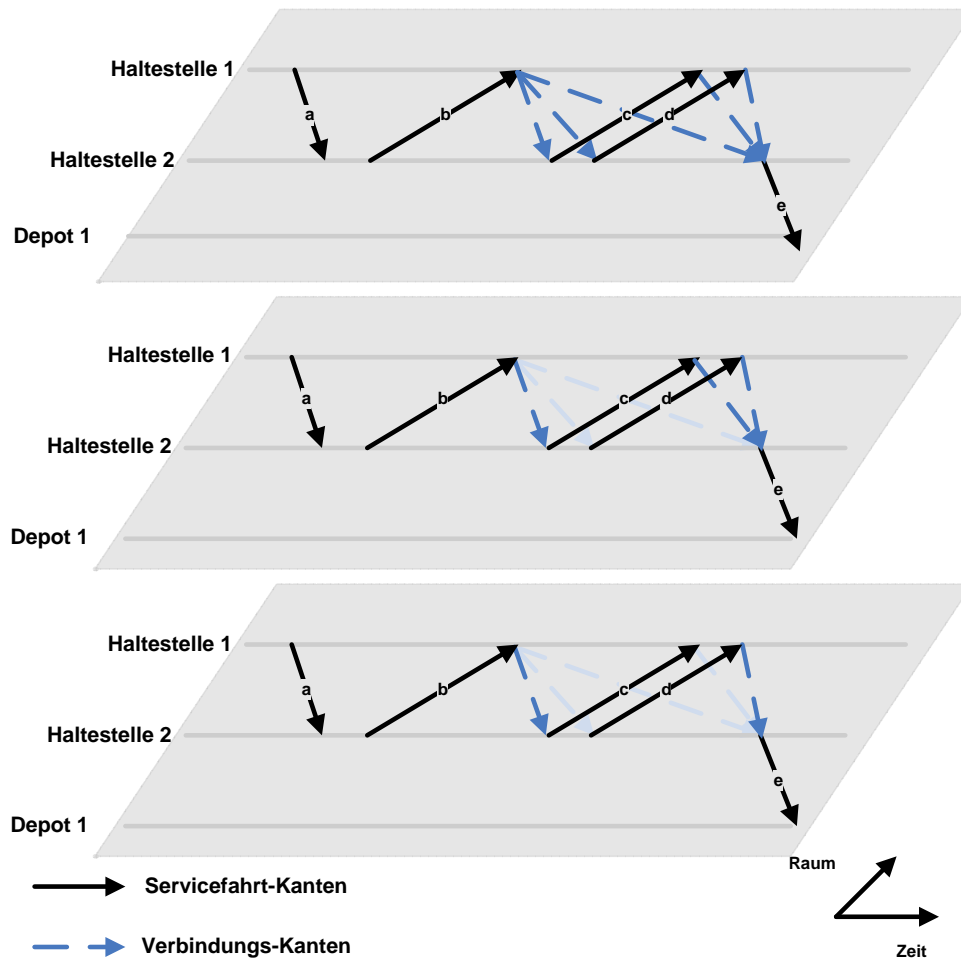


Abbildung 4.6: Matches-Reduktion

also die Fahrten c , d und e , so dass für diese je eine Verbindungs-Kante eingefügt wird. Weitere Verbindungs-Kanten können zwischen den Fahrten c und e sowie d und e eingefügt werden.

Der mittlere Bereich der Abbildung 4.6 enthält die First-Matches. Dabei handelt es sich jeweils um die Verbindungs-Kante zu einer Servicefahrt-Kante in einer anderen Haltestelle mit dem frühesten Startzeitpunkt von allen möglichen Matches in dieser Haltestelle. Alle späteren Servicefahrt-Kanten, zu denen potentielle Verbindungs-Kanten geführt hätten, können durch Warten auf der Timeline erreicht werden. Somit fallen sämtliche Verbindungs-Kanten außer der First-Matches aus der Betrachtung heraus, was in den praktischen Instanzen eine drastische Reduzierung der Modellgröße ermöglicht.

In der Abbildung kann man gut erkennen, dass so für das Beispiel die Verbindungs-Kanten von b nach d sowie von b nach e wegfallen (im Bild nur noch angedeutet),

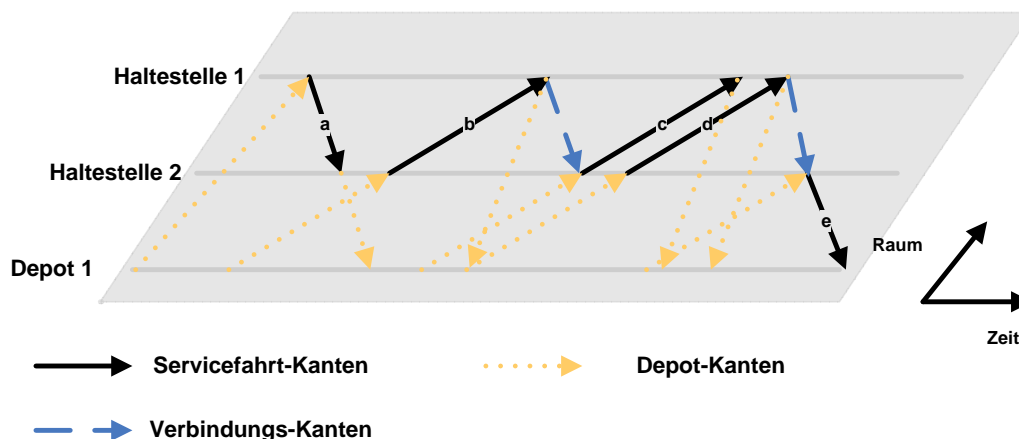


Abbildung 4.7: Einfügen der Depot-Kanten

da für eine Verbindung auch die Kante von b nach c benutzt werden kann und dann auf der Timeline der Haltestelle 2 bis zu den Startzeitpunkten der anderen gewartet werden kann. Im unteren Bereich der Abbildung 4.6 werden noch alle First-Matches, die nicht zu den so genannten Latest-First-Matches zählen, herausgefiltert: Wenn es unter den First-Matches zwischen zwei bestimmten Haltestellen mehrere Verbindungskanten gibt, die zu der gleichen Servicefahrt-Kante führen, reicht die letzte von ihnen bereits aus, um alle Verbindungen zu ermöglichen. Die Verbindungen der anderen First-Matches können über entsprechende Wartenkanten in der Timeline erreicht werden.

Im Beispiel kann das Netzwerk also um die Verbindungs-Kante von c nach e reduziert werden, da ein Bus nach Ausführung der Servicefahrt-Kante c und nach entsprechendem Warten in der Haltestelle die Verbindungs-Kante von d nach e nutzen kann, um die nächste Servicefahrt-Kante zu erreichen. Der Fluss auf den verbleibenden Verbindungs-Kanten wird nicht mehr durch eins begrenzt, da ein Latest-First-Match für mehrere Fahrtenverknüpfungen benutzt werden darf. (Bei einer Flusschranke von eins dürfte nur eine der Verknüpfungsmöglichkeiten genutzt werden und das reduzierte Netzwerk wäre nicht mehr äquivalent zu dem Netzwerk mit allen Verbindungs-Kanten.)

Depot-Kanten einfügen

In der Phase 3 des Netzwerkaufbaus wird für jede Servicefahrt-Kante im Netzwerk eine Depot-Kante eingefügt, die vom Depot der jeweiligen Schicht zu der Starthaltestelle der Servicefahrt-Kante hinführt, so dass die Depot-Kante genau zur Abfahrt der Servicefahrt ankommt. Zusätzlich wird für jede Service-Kante im Netzwerk eine Depot-Kante eingefügt, die zum Zeitpunkt der Ankunft der Servicefahrt-Kante

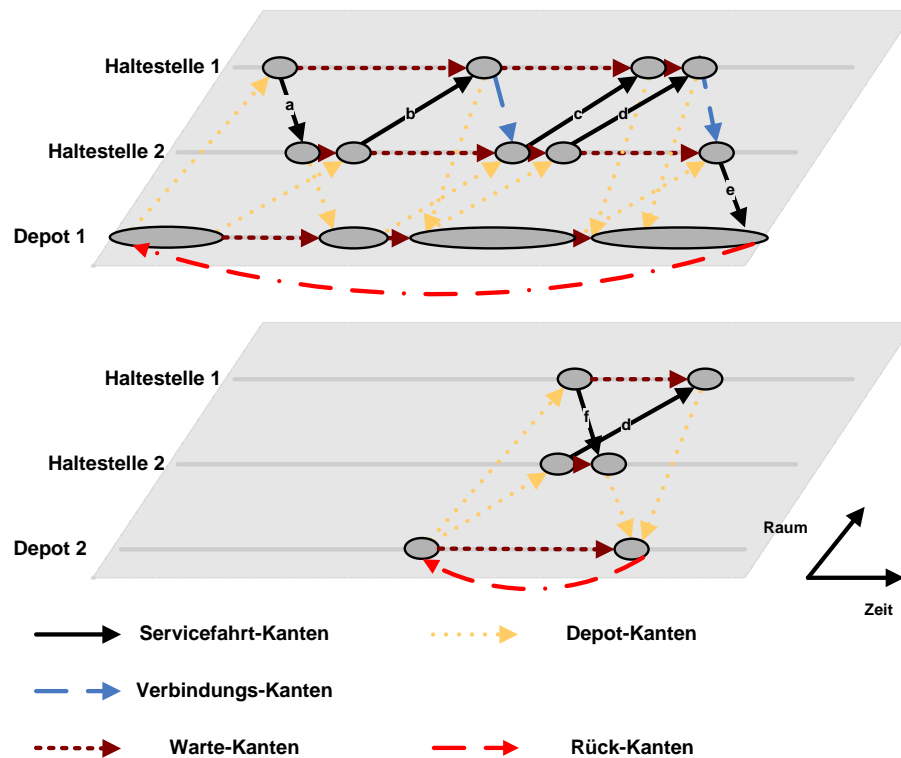


Abbildung 4.8: Knotenbildung, Einfügen der Warte- und Zirkulationsfluss-Kanten

eine Fahrt in das Depot ermöglicht.

Somit garantieren Depot-Kanten sowohl das Erreichen einer Haltestelle zu Beginn eines Umlaufs, als auch die Rückfahrt in das Depot am Ende des Umlaufs. Zusätzlich ermöglichen die Depot-Kanten auch eine Rückfahrt in das Depot innerhalb des Umlaufs, um gegebenenfalls lange Wartezeiten im Depot zu verbringen, falls dies kostengünstiger ist als ein Warten an der Haltestelle.

In Abbildung 4.7 können wir erkennen wie für jede Servicefahrt-Kante eine Depot-Kante vom Depot zum Startzeitpunkt, sowie eine Depot-Kante vom Endzeit zum Depot zurück eingefügt werden.

In vielen realistischen Fahrplänen gibt es Konstellationen, so dass mehrere Fahrten von der gleichen Haltestelle zum gleichen Zeitpunkt starten (oder ankommen). In diesem Fall würden mehrere Depot-Kanten eingefügt werden, die zur gleichen Zeit starten und ankommen. Diese sich überlagernden Kanten können durch eine einzige ersetzt werden. Allerdings muss die obere Schranke ihrer Flussgröße um die Anzahl der auf diese Weise eingesparten Depot-Kanten erhöht werden, damit ggf. alle verbundenen Abfahrten (oder Ankünfte) angeschlossen sind.

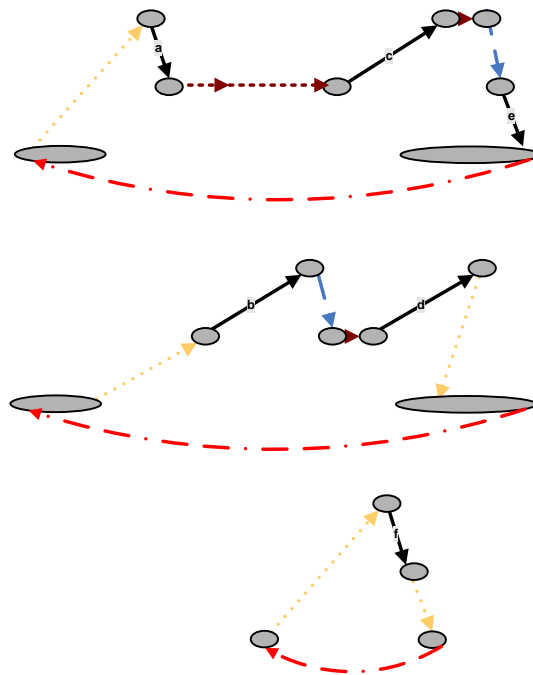


Abbildung 4.9: Ein optimaler Fluss von drei Flusseinheiten

Knoten bilden, Warte- und Zirkulationsfluss-Kanten einfügen

In der fünften und letzten Phase des Netzwerkaufbaus muss aus der Struktur der Timelines ein Netzwerk mit Knoten erstellt werden. Zudem muss der wiederholte Fluss durch das Netzwerk durch Zirkulationsfluss-Kanten sichergestellt werden. Für das Beispiel wird diese Phase durch die Abbildung 4.8 dargestellt.

Im ersten Schritt werden aus den Ankünften und Abfahrten aller Timelines Knoten gebildet. Um das Verweilen auf einer Timeline mit Kanten zu modellieren, werden im zweiten Schritt dieser Phase Warte-Kanten entlang der Timelines eingefügt. Eine Warte-Kante verbindet jeden Knoten des Netzwerks mit dem nächsten Knoten in seiner Timeline. Zuletzt müssen dem Netzwerk noch die Zirkulationsfluss-Kanten hinzugefügt werden. Diese Kanten werden vom letzten Knoten der Depot-Timeline zum ersten derselben geführt. Die Flussgrößen auf den Zirkulationsflusskanten in einer Lösung entsprechen der Anzahl der benötigten Fahrzeuge.

In Abbildung 4.8 kann man die Knotenbildung und das Einfügen der Warte- und Zirkulationsfluss-Kanten für das Beispiel-Netzwerk sehen. Alle Kanten sind nun nicht mehr an die Timelines, sondern an die neu erstellten Knoten gebunden. Das Time-Space-Netzwerk für das Beispiel-MDVSP ist vollständig aufgebaut. Als weitere wichtige Restriktion für die Knoten im Netzwerk muss für jede gültige Lösung gelten, dass die Anzahl der eingehenden Flusseinheiten jedes Knotens gleich

der Anzahl der ausgehenden Flusseinheiten sein muss. Diese Bedingung stellt die Flusserhaltung im Netzwerk sicher. Eine mögliche Flusslösung mit drei Umläufen ist in der Grafik 4.9 zu sehen. Da es keine Möglichkeit gibt das Netzwerk mit weniger als drei Flüssen zu durchlaufen, kann die mit negativen Kosten belegte Linien-Kante benutzt werden und somit die angesprochene Reduzierung der Linien-Wechsel in dem Umlauf gefördert werden.

4.2 Lösungsansatz

Zum Lösen des formulierten gemischt-ganzzahligen Programms wird zunächst eine optimale Lösung für die LP-Relaxation des mathematischen Modells 4.8-4.14 berechnet. Basierend auf dieser Lösung wird die exakte optimale ganzzahlige Lösung mit einem Branch-and-Cut Verfahren ermittelt.

In der Literatur zu MDVSP sind viele heuristische Ansätze beschrieben worden, die das „klassische“ Problem aus dem Kapitel 3 vereinfachen, um eine akzeptable Lösung zu bekommen. Andere Ansätze lösen das Problem exakt optimal, benötigen aber fortgeschrittene komplexe Methoden, um die Problemkomplexität in den Griff zu bekommen. Da die in diesem Kapitel beschriebene kompakte Problemmodellierung zu einer enormen Reduktion der Problemgröße führt, ohne die Optimalität zu gefährden, ist es möglich geworden, die meisten praxisrelevanten Problemstellungen exakt optimal mit Hilfe der Standard-Optimierungssoftware zu lösen.

Zur Lösung des für das Mehrdepot-Umlaufplanungsproblem aufgestellten Programms 4.8-4.14 können verschiedene mathematische Optimierungsbibliotheken benutzt werden. In Rahmen der vorliegenden Arbeit wurden Schnittstellen zu den Optimierungsbibliotheken MOPS ([Suhl, 2000]) und ILOG CPLEX ([ILOG, 2004]) implementiert. Bei der Lösung eines gemischt-ganzzahligen Optimierungsproblems wird sowohl von MOPS als auch von CPLEX erst die LP-Relaxation des Problems gelöst und dann mit Hilfe spezieller Branch-and-Cut-Techniken die optimale ganzzahlige Lösung bestimmt.

4.3 Numerische Ergebnisse

Nachfolgend werden die numerischen Ergebnisse für das in diesem Kapitel beschriebene Modell und den Lösungsansatz vorgestellt. Die Implementierung wurde unter der .NET-Umgebung in der Programmiersprache C# vorgenommen. Näheres zu der Programmiersprache oder der .NET-Klassenbibliothek findet sich in [Hanisch, 2002]. Als Entwicklungsumgebung wurde „Microsoft Visual Studio 2003“ verwendet (vgl. [Schwichtenberg and Eller, 2004]). Sämtliche im Rahmen dieser Ar-

beit vorgenommenen Tests wurden auf einem XEON 2,20 GHz Prozessor mit 2 GB RAM unter Microsoft Windows XP Professional ausgeführt.

4.3.1 Probleminstanzen und Testergebnisse

Es wurden zahlreiche Tests auf realen Instanzen durchgeführt. Im Folgenden werden die Testergebnisse für neun unterschiedlich komplexe Instanzen untersucht. Die Instanzen stammen aus mehreren deutschen Städten. Die Münchener Instanz *mun1* ist auch in der kapazitierten Variante (*mun1-kap*) vorhanden. Die Kapazitäten sind durch eine minimale und eine maximale Anzahl der verfügbaren Fahrzeuge pro Fahrzeugtyp-Depot-Kombination ausgedrückt. Die Vorgaben, dass dem Depot 47 mindestens 10 Fahrzeuge vom Typ 1 zugewiesen werden müssen und höchstens 120 zugewiesen werden dürfen wird durch folgende Relation ausgedrückt:

```
VEHYPECAPTOSTOPPOINT:
VehTypeID;StoppointID;Min;Max
1;47;10;120
```

Die verwendeten Instanzen sind im Bezug auf die Lösbarkeit unterschiedlich schwierig. Wie komplex eine Instanz ist, hängt von mehreren Faktoren ab. Die wichtigsten Kennzahlen für die Komplexität einer Problemstellung sind:

- die Größe des Fahrplans, die durch die Anzahl der Fahrten vorgegeben ist,
- die Anzahl der Depots und der Fahrzeugtypen sowie
- die Anzahl der Haltestellen.

Diese Kennzahlen sind für die untersuchten Instanzen in der Tabelle 4.2 angegeben. Eine weitere Kennzahl für die Komplexität der jeweiligen Instanz ist die durchschnittliche Anzahl der erlaubten Depot-Fahrzeugtyp-Zuordnungen pro Fahrplanfahrt. Diese Kennzahl wird die *durchschnittliche Gruppengröße* bezeichnet. Sie ist in der Spalte $\circlearrowleft G$ der Tabelle 4.2 für jede Instanz angegeben und bedeutet die Anzahl der Schichten, in denen eine Fahrplanfahrt durchschnittlich abgebildet wird. Berechnet wird sie als die Gesamtanzahl der Fahrplanfahrt-Kanten in allen Netzwerkschichten geteilt durch die Anzahl der Fahrplanfahrten:

$$\circlearrowleft G = \sum_{f \in F} |S_f| / |F| = \sum_{s \in S} |F_s| / |F|$$

Die Tabelle 4.3 beschreibt die Eigenschaften der Netzwerkmodelle für die untersuchten Instanzen. Neben der Gesamtanzahl von Knoten und Kanten im Time-Space-Netzwerk ist auch die Anzahl der Verbindungskanten angegeben. Ein Vergleich der Anzahl von Matches (Spalte $\#Matches$) und der Anzahl von Latest First

Instanz	#Fahrten	#Depots	#Fahrzeugtypen	$\odot G$	#Haltestellen
Saar	1296	2	3	2,0	89
hal	2047	2	3	3,2	21
mun14	2452	19	1	2,0	61
mun15	2825	19	1	2,0	90
mun28	1808	19	1	13,0	76
mun18	2394	19	1	12,0	103
mun1	11062	19	12	6,5	161
mun1-kap	11062	19	12	6,5	161
mun2	10710	3	12	7,1	140

Tabelle 4.2: Eigenschaften der Probleminstanzen

Instanz	Knoten	Kanten	Layers	Matches	First M.	Latest FM	ND-LFM	%
Saar	5751	40121	6	636642	77828	37444	2689	0,4%
hal	12947	56334	6	2115896	69390	34955	26412	1,2%
mun14	10511	100048	2	5014262	124686	76254	75215	1,5%
mun15	11861	124191	2	5810252	235548	105492	95714	1,6%
mun28	50489	478823	13	15039882	801736	385593	358940	2,4%
mun18	59735	675333	12	23542944	1285584	564120	530974	2,3%
mun1	151852	1479833	56	51114069	2560914	1194266	1117291	2,2%
mun1-kap	151852	1479833	56	51114069	2560914	1194266	1117291	2,2%
mun2	148913	2027743	15	161452729	3632528	1930432	1657410	1,0%

Tabelle 4.3: Eigenschaften der Netzwerkmodelle

Instanz	Zeilen	Spalten	Nonzeros	CPU(sec)	LP-Solver
Saar	7053	40121	82840	2	Dual
hal	15000	56334	119242	45	IPM
mun14	12982	100048	205002	11	Dual
mun15	14705	124191	254034	14	Dual
mun28	52316	478823	981163	2154	IPM
mun18	62148	675333	1379406	3292	IPM
mun1	163142	1479833	3031285	30765	Dual
mun1-kap	163142	1479833	3031285	18345	Dual
mun2	159659	2027743	4131436	>36000	Dual

Tabelle 4.4: Exakte Optimierung: Ergebnisse

Instanz	#Fahrzeuge	#Blöcke	Operative Kosten	Fixe Fahrzeugkosten
Saar	49	95	153469	32300000
hal	115	824	6557	10050
mun14	107	129	238702	18029500
mun15	94	124	243467	11797000
mun28	56	121	113022	11928000
mun18	82	210	152669	17466000
mun1	417	706	891551	76771496
mun1-kap	440	700	900175	81650992

Tabelle 4.5: Eigenschaften der optimalen Lösungen

Matches (Spalte #Latest FM) zeigt den Ausmaß der Modellreduktion gegenüber der klassischen Fluss-Modellierung aus der Literatur. Die Spalte ND-LFM enthält die Anzahl der Leerfahrtskanten, die explizit im Netzwerk eingefügt wurden (Latest First Matches nicht über Depot). Wie die letzte Tabellenspalte zeigt, führte der Time-Space-Netzwerk basierte Ansatz zu einer Reduktion der Anzahl von Verbindungskanten auf 0,4 bis 2,4%.

Durch eine derartige Reduktion wurde auch die Größe des zu lösenden mathematischen Modells und somit die Lösungszeit für das Problem im Rahmen gehalten. Die Tabelle 4.4 beinhaltet die Anzahl der Restriktionen, Variablen und Nichtnull-elemente im mathematischen Modell sowie die Optimierungszeit zur Lösung dieses Modells. Zur Lösung der gemischt-ganzzahligen Probleme wurde ILOG CPLEX 9.0 (vgl. [ILOG, 2004]) verwendet. Die Tests wurden mit dem dualen Simplex-Algorithmus und mit der Interior-Point-Methode durchgeführt. In der Spalte für die Laufzeit (CPU(sec)) ist die jeweils beste Zeit eingetragen, in der letzten Spalte - das entsprechende Verfahren.

Für die Optimierung jeder Instanz wurde ein Zeitlimit von 10 Stunden gesetzt. Alle Instanzen bis auf *mun2* konnten in dieser Zeit optimal gelöst werden. Aufgrund der beobachteten Laufzeiten kann man schließen, dass die durchschnittliche Gruppengröße einen starken Einfluss auf die Komplexität des Problems hat. Die Instanzen *mun18* und *mun28* mit einer relativ kleiner Anzahl von Fahrten haben sich als schwierig lösbar erwiesen; sie weisen auch die höchsten Werte für die Gruppengröße auf.

In der Tabelle 4.5 sind die Eigenschaften der optimalen Lösungen für die gelösten Instanzen zusammengefasst. Neben der Anzahl Fahrzeugumläufe ist auch die Anzahl der Depotblöcke angegeben. Ein Block ist ein Umlaufteil zwischen zwei Depotaufenthalten. Je mehr solcher Blöcke generiert wurden, desto öfter halten sich die Fahrzeuge im Laufe eines Betriebstages in ihrem Home-Depot auf. Neben den minimalen Kosten ist es eine weitere erwünschte Eigenschaft eines Umlaufplans, da im Depot ein Fahrerwechsel stattfinden kann beziehungsweise die Pausen für den Fahrer besser als an normalen Haltestellen gestaltet werden können.

Instanz	MIP-Zeit (\simeq root-LP-Zeit)				Best Solver
	CPLEX IPM	CPLEX Dual	MOPS IPM	MOPS Dual	
Saar	10	2	18	5	CPLEX Dual
hal	45	80	70	114	CPLEX IPM
mun14	81	11	154	34	CPLEX Dual
mun15	121	14	197	35	CPLEX Dual
mun28	2154	20931	1972	17919	MOPS IPM
mun18	3292	20929	4304	14350	CPLEX IPM
mun1	Out of Mem	30765	Out of Mem	>36000	CPLEX Dual

Tabelle 4.6: Laufzeiten-Vergleich: IPM (Barrier) versus dualer Simplex

Instanz	#Fahrzeuge	MIP-Zeit	davon root-LP
Saar	49	48	29
hal	115	1185	556
mun14	107	717	352
mun15	94	583	255

Tabelle 4.7: Laufzeiten für die Minimierung der Fahrzeuganzahl

4.3.2 Wahl des LP-Verfahrens: Laufzeitanalyse

Eine Analyse des MIP-Verhaltens hat gezeigt, dass bei fast allen Instanzen der Großteil der Laufzeit mit der Lösung der LP-Relaxation im Wurzelknoten von Branch-and-Cut verbracht wurde. Daher ist es wichtig, aus den verfügbaren LP-Algorithmen vom jeweiligen Solver den schnellsten auszuwählen. Die Entscheidung, welches LP-Verfahren sich besser für eine Problem Instanz eignet, ist nicht trivial. Zum Lösen der LP-Relaxationen der MDVSP-Instanzen im Rahmen des Branch-and-Bound Verfahren bieten sich grundsätzlich zwei Algorithmen an: die duale Simplex-Methode und die Interior-Point-Methode. Für jede Instanz wurden vier Optimierungsläufe durchgeführt: mit CPLEX IPM (Barrier Optimizer) und Dual, MOPS IPM und Dual. Sowohl MOPS-, als auch CPLEX-Ergebnisse in der Tabelle 4.6 zeigen, dass die Instanzen mit höheren Werten für die durchschnittliche Gruppengröße und vielen Fahrten am besten von Interior-Point-Methode gelöst werden und die mit wenigen Schichten bzw. geringerer durchschnittlichen Gruppengröße - mit dem dualen Simplex. Allerdings braucht die Interior-Point-Methode fast doppel so viel Speicher wie der duale Simplex. Das führt dazu, dass die ganz großen Instanzen wegen Begrenzung des Hauptspeichers nicht mehr mit dieser Methode lösbar sind (s. auch Zeile mun1 in der Tabelle 4.4). Der primale Simplex wies schon bei kleineren Problem Instanzen nicht akzeptable Laufzeiten auf und wurde daher nicht weiter verwendet.

4.3.3 Minimierung der Fahrzeuganzahl

Interessant ist auch der Unterschied in den Laufzeiten, der sich aus den verschiedenen Zielsetzungen in den ansonsten gleichen Problemstellungen ergibt. Die Tabelle 4.7 zeigt exemplarisch, die Ergebnisse der Testläufe, in denen nur die Anzahl der Fahrzeuge minimiert werden sollte. Dafür wurden in den Daten die Kostensätze wie folgt verändert: die Kilometer- und Stundenkostensätze wurden auf null gesetzt, die fixe Fahrzeugkosten - auf eins. Diese ergibt die Flusskosten von eins auf jeder Zirkulationsflusskanten. Damit wird mit den gleichen Verfahren nun die Anzahl der Fahrzeuge statt der Gesamtkosten minimiert. Dies wird auch durch die Ergebnisse in der Tabelle 4.7 bestätigt. Die Anzahl der Fahrzeuge (Spalte Fleetsize) ist minimal für jede untersuchte Probleminstanz.

Was die Laufzeiten dieser Tests betrifft - könnte man annehmen, dass die fahrzeuganzahlminimale Lösungen schneller berechnet werden, als die mit minimalen Gesamtkosten. Wie die Laufzeiten in der Tabelle 4.7 zeigen, sind die Laufzeiten entgegen dieser Erwartung bei allen Instanzen stark gestiegen. Eine Erklärung dafür ist, dass eine Differenzierung der Kosten auf den Kanten im Falle einer Gesamtkostenminimierung zu einer Art Kostenperturbation im gemischt-ganzzahligen Modell führt, so dass der Optimierer schneller die optimale Lösung finden kann.

Kapitel 5

Flussdekomposition für TSN

Mit dem Ansatz aus Kapitel 4 können nun optimale Flusslösungen für die Netzwerkfluss-Formulierung des Mehrdepot-Umlaufplanungsproblems berechnet werden. Eine optimale Flusslösung wird dabei durch die ganzzahligen Flusswerte auf den Netzwerkkanten beschrieben. In Kapitel 4 wurde aber noch nicht beschrieben, wie man von einer solchen Flusslösung zu den gesuchten Fahrzeugumläufen kommt.

Die Art der Modellierung, die in dieser Arbeit für das MDVSP gewählt wurde, erlaubt es, mehrere unterschiedliche Umlaufpläne aus einem zulässigen Netzwerkfluss zu bilden. Dies geschieht durch eine geeignete Strategie bei der Flussdekomposition. Werden in der Flussdekomposition bestimmte Ziele verfolgt, wie die Homogenität der Umläufe bezüglich Linien oder günstigere Pausenverteilungen, so handelt es sich um ein Vorgehen in die nächste Planungsaufgabe (Dienstplanung) und bedeutet eine teilweise Integration dieser beiden Planungsschritte.

5.1 Dekomposition der Flüsse in Wege

Allgemein kann bei der Formulierung der Netzwerkflussprobleme auf einen der beiden äquivalenten Modellierungsansätze zurückgegriffen werden. Es können Flüsse auf Kanten definiert werden (wie in Kapitel 4), oder alternativ Flüsse entlang der Pfade und gerichteter Zyklen (*engl.: arc flow formulation vs. path and cycle flow formulation* - s. auch [Ahuja et al., 1983]).

In dieser Arbeit wird durchgehend die kantenbasierte Formulierung verwendet. Manchmal ist es aber notwendig von einer zu der anderen Darstellungsart zu wechseln. Falls eine pfad- und zyklenbasierte Beschreibung eines Flusses gegeben ist, wie in der Abbildung 5.1 rechts, können die Flussgrößen auf Kanten einfach berechnet werden, indem die Flussgrößen auf entsprechenden Pfaden und Zyklen addiert werden. Die umgekehrte Fragestellung ist viel schwieriger zu lösen. Dies liegt daran, dass eine kantenbasierte Flussbeschreibung auf verschiedene Weise in eine pfad-

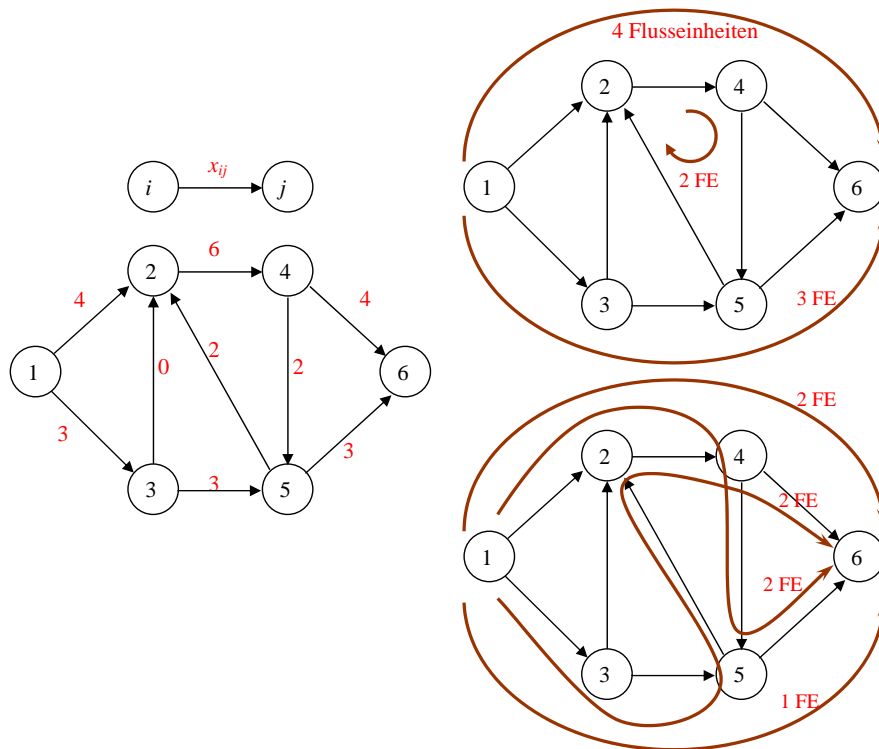


Abbildung 5.1: Dekomposition eines S-T-Flusses in Wege

und zyklenbasierte überführt werden kann. Dies gilt auch für den kostenoptimalen Fluss, da dieser zulässig sein muss.

Betrachtet man ein S-T-Netzwerk, oder einen Zirkulationsfluss, dann entspricht die Anzahl der Wege und der Zyklen von einzelnen Flusseinheiten dem Flussvolumen. Dies wird am Beispiel in der Abbildung 5.1 verdeutlicht (in Anlehnung an [Ahuja et al., 1983]).

Die Besonderheit der Dekomposition eines zulässigen Flusses in einem Time-Space-Netzwerk eines Mehrdepotumlaufproblems besteht darin, dass der Fluss in N Zyklen dekomponiert wird. N entspricht dabei der Anzahl der benötigter Fahrzeuge. Jeder Zyklus stellt also einen Busumlauf dar, wie es auch an der optimalen Lösung des Beispiels aus Kapitel 4 in der Abbildung 4.9 zu sehen war.

5.2 Die Dekompositionsstrategien

Nachdem eine optimale Lösung für das Netzwerkflussmodell in Form eines Flusses berechnet wurde, ergibt sich eine Anzahl von Möglichkeiten, dieses Netzwerk zu durchlaufen, d.h., die einzelnen Fahrzeugumläufe zu bilden. In Begriffen der Netzwerkmodellierung besteht an dieser Stelle die Notwendigkeit von der kanten-

basierten zu der wegbasierten Flussbeschreibung zu wechseln um den gegebenen Netzwerkfluss zu dekomponieren.

Der Weg einer Flusseinheit stellt in dem MDVS-Netzwerk die Aktivitäten eines Tages für ein Fahrzeug dar. Die endgültigen Umläufe werden nun mit Hilfe einer geeigneten Dekompositionsstrategie konstruiert. Hierfür stehen mehrere Möglichkeiten zur Verfügung, die in der Implementierung wie folgt bezeichnet wurden:

- FIFO - in jedem Knoten in der Ankunftsreihenfolge verknüpfen (First In - First Out),
- LIFO - in jedem Knoten in der umgekehrter Ankunftsreihenfolge verknüpfen (Last In - First Out),
- MinAlternation - in jedem Knoten zuerst die Aktivitäten gleicher Linie verknüpfen (Minimal Alternation),
- XMinAlternation - in jedem Knoten ein Summenzuordnungsproblem lösen (eXtended Minimal Alternation),
- EngPassMinAlternation - in jedem Knoten ein Engpasszuordnungsproblem lösen,
- SPP - ein Set Partitioning Problem lösen.

Die ersten vier Dekompositionsstrategien verknüpfen die Flusseinheiten an den einzelnen Haltestellen/Zeitknoten zu Umläufen, die je nach Methode bestimmte Eigenschaften aufweisen. Die Gesamtheit der konstruierten Umläufe stellt den Umlaufplan dar.

Die fünfte Methode besteht darin, den kostenoptimalen Fluss durch das Lösen eines Set Partitioning Problems zu dekomponieren. Hierbei handelt es sich um eine Methode, die sämtliche mögliche kostenoptimale Umlaufpläne betrachtet, und einen auswählt, der in Bezug auf die anderen Kriterien (hier, z.B. - Linienreinheit) die besten Eigenschaften aufweist. Nachfolgend wird hauptsächlich die Minimierung der Linienwechsel im Rahmen von SPP betrachtet. Es ist aber nicht weiter schwierig dieselbe Technik für die anderen Ziele anzuwenden. Denkbar ist z.B. die Länge der Umläufe oder auch bestimmte Verteilungen von Pausen und Pausendauer.

Eins ist den im Folgenden beschriebenen lokalen Dekompositionsstrategien gemeinsam: es werden sequentiell alle Knoten im Netzwerk abgearbeitet. Ein Knoten ist ein Ort (eine bestimmte Haltestelle) zu einem Zeitpunkt, an dem ein oder mehrere Fahrzeuge, die gerade bestimmte Aktivitäten ausgeführt haben, nun auf die folgenden, anstehenden Aktivitäten warten. Dementsprechend müssen die anstehenden Aufgaben (die ausgehenden Flusseinheiten) mit den gerade erledigten Aufgaben (den eingehenden Flusseinheiten) verknüpft werden.

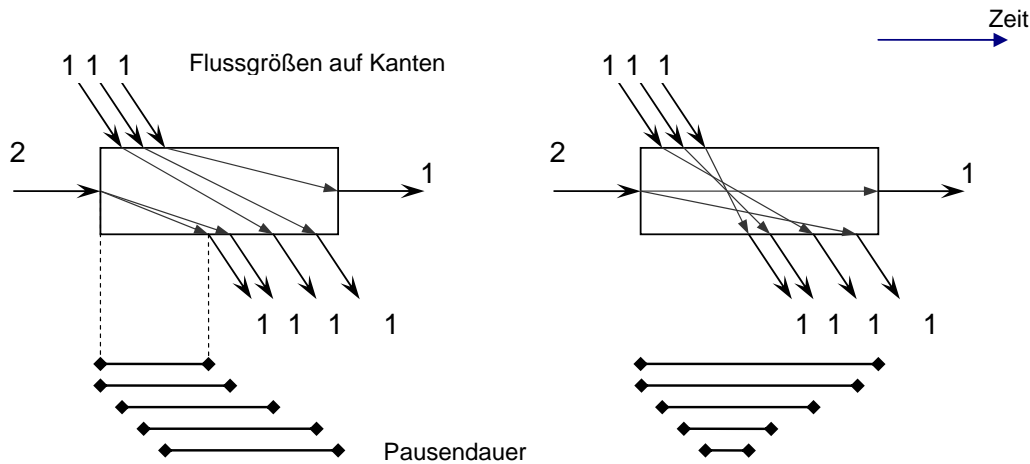


Abbildung 5.2: Dekomposition eines Flusses nach FIFO/LIFO

In einem Knoten wird also jede anstehende Aktivität eines Fahrzeugs - sei es Fahren, Warten oder Standortwechseln - einer der bereits erledigten Aktivitäten zugeordnet. Die Freiheitsgrade für diese Zuordnung erlauben eine Auswahl von Verfahren, die allesamt kostenoptimale Umlaufpläne konstruieren, dabei aber bestimmte umlaufbezogene Aspekte mitberücksichtigen.

5.2.1 *FIFO*: First In First Out

Diese Strategie verbindet in jedem Knoten des Netzwerks die zuerst eingehende Kante mit der zuerst ausgehenden. Von der praktischen Seite gesehen bedeutet das folgendes: falls zu einem Zeitpunkt mehrere gleichwertige Fahrzeuge in einer Haltestelle stehen, wird die nächste Fahrt mit dem Fahrzeug bedient, der am längsten an dieser Haltestelle gewartet hat.

Dadurch wird eine sehr einfache und dadurch extrem schnelle Strategie geschaffen, die allerdings sämtliche praktische Aspekte, wie die Anzahl der Linien in den Umläufen und die Anzahl der Linienwechsel unberücksichtigt lässt. Für Problemstellungen, die keine zusätzlichen Anforderungen enthalten, liefert diese Strategie dennoch brauchbare Ergebnisse.

5.2.2 *LIFO*: Last In First Out

Bei dieser Strategie wird die letzte eingehende Kante mit der ersten ausgehenden verbunden. Wiederum von der praktischen Seite gesehen bedeutet das, dass von mehreren gleichwertigen Fahrzeugen, die an einer Haltestelle stehen, die nächste

Fahrt mit dem Fahrzeug bedient wird, das zuletzt an dieser Haltestelle angekommen ist.

Auch bei dieser Strategie, analog zu FIFO, bleibt die Anzahl der Linien in einem Umlauf und die Anzahl der Linienwechsel unberücksichtigt, auch hier sind die Ergebnisse stark vom Zufall abhängig. Aber auch diese Strategie ist algorithmisch sehr einfach und ermöglicht somit extrem kurze Rechenzeiten.

Es ist nahe liegend, dass eine Dekomposition nach FIFO-Prinzip tendenziell für alle Fahrer und Fahrzeuge gleichmäßige Pausen generiert. Bei dem LIFO-Verfahren werden dagegen viele kurze und viele lange Unterbrechungen erzeugt. Dies wird durch das Beispiel in der Abbildung 5.2 deutlich.

5.2.3 *MinAlternation*: Greedy-Verknüpfung

Die Strategie *MinAlternation* ist die einfachste lokale Strategie, die die Anzahl der Linien, sowie der Linienwechsel in einem Umlauf berücksichtigt. Bei dieser Strategie wird in jedem Knoten geprüft, ob eine eingehende und eine ausgehende Kante Fahrten der gleichen Linie repräsentieren und diese, im Fall einer Liniengleichheit, verbindet. Diese Strategie ist allerdings nur wenig besser als FIFO und LIFO, da hier zwar eine Auswahl stattfindet, diese aber nur in den einzelnen Knoten geschieht. Allerdings unterscheidet sich die Laufzeit kaum von denen der beiden vorgenannten Strategien und die Ergebnisse sind immer gleich oder besser als LIFO oder FIFO. *MinAlternation* stellt die einfachste Strategie mit einem Greedy-Vorgehen und dementsprechend nur geringfügig verbesserten Linienberücksichtigung dar.

5.2.4 Zuordnungsproblem-Strategien

Auch hier werden in jedem Knoten die eingehenden und die ausgehenden Kanten verglichen und in Bezug auf die Liniengleichheit verbunden, allerdings wird hierbei auch das bisher schon durchlaufende Netzwerk gemerkt und berücksichtigt. Es wird in dem aktuell betrachteten Knoten immer die ausgehende Kante verbunden, die entweder liniengleich mit einer eingehenden Kante ist oder mit einer Kante, die schon einmal eine liniengleiche Kante auf ihrem Weg enthielt. Hierbei wird die Anzahl der Linien in jedem Umlauf deutlich minimiert, da die Strategie immer versucht sich mit Linien fortzusetzen, die schon einmal in dem Teilumlauf vorgekommen sind.

Technisch wurde es dadurch realisiert, dass jede mögliche Zuordnung von Fahrten mit einer Bewertungszahl versehen wird. Diese Zahl gibt wieder, wie viele verschiedene Linien das resultierende Umlaufstück beinhalten würde, falls diese konkrete Zuordnung vorgenommen wird. Um eine optimale Zuordnung zu finden, muss in jedem Knoten ein Zuordnungsproblem gelöst werden.

Exkurs: das Zuordnungsproblem

Das Zuordnungsproblem (*engl.: assignment problem*) bildet eine Entscheidungssituation ab, in der eine bestimmte Anzahl von Objekten der gleichen Anzahl von Positionen zugeordnet werden soll. Dabei ist für jede Kombination eine Bewertungszahl vorgegeben. Wird als Zielsetzung eine Minimierung der Summe der Bewertungszahlen der gewählten Kombinationen angestrebt, so handelt es sich um den klassischen Fall des Zuordnungsmodells - das Engpasszuordnungsproblem (s. [Dürr and Kleibohm, 1992]).

Seien in einem Zuordnungsproblem n Objekte den n Positionen eindeutig zuzuordnen und sei die Bewertung jeder Kombination von Objekt und Position durch die Kosten c_{ij} vorgegeben. In einer Lösung x beschreibt der Wert x_{ij} die Entscheidung, ob Objekt i der Position j zugeordnet wurde. Und zwar, ist $x_{ij} = 1$, falls Objekt i der Position j zugeordnet wurde und $x_{ij} = 0$ sonst.

Mathematisches Modell Wird eine Minimierung der *Summe in der Lösung auftretenden Bewertungszahlen* angestrebt, so handelt es sich um das *Summenzuordnungsproblem*, das durch das folgende mathematische Modell beschrieben wird:

$$\min \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \quad (5.1)$$

$$\sum_{j=1}^n x_{ij} = 1 \quad \forall i = 1 \dots n \quad (5.2)$$

$$\sum_{i=1}^n x_{ij} = 1 \quad \forall j = 1 \dots n \quad (5.3)$$

$$0 \leq x_{ij} \quad \forall i = 1 \dots n, j = 1 \dots n \quad (5.4)$$

Im Gegensatz zu dem vorher beschriebenen Summenzuordnungsproblem, soll bei einem *Engpasszuordnungsproblem* (*engl.: bottleneck assignment problem*) die *größte bei der Zuordnung auftretende Bewertung* minimiert werden (in Anlehnung an [Dürr and Kleibohm, 1992]).

Sei nun die größte in der Lösung auftretende Bewertungszahl zu minimieren. Das mathematische Modell verändert sich wie folgt. Die Restriktionen eines mathematischen Modells für die Engpasszuordnung gleichen den Restriktionen im Fall der Summenzuordnung. Als Zielfunktion wird nun die Minimierung der maximalen Bewertungszahl der Zuordnung angestrebt:

$$\min \max_{ij} c_{ij} x_{ij} \quad (5.5)$$

Lösungsansätze Das Zuordnungsproblem ist ein Sonderfall des klassischen Transportproblems, das wiederum ein Sonderfall des klassischen Flussproblems ist (*engl.: min cost flow problem*). Es könnten dementsprechend Methoden angewandt werden, die für diese allgemeineren Problemklassen entwickelt wurden. Aufgrund der einfacheren Struktur dieser Aufgabenstellung existieren aber auch sehr effiziente problemspezifische Algorithmen zur Lösung des Zuordnungsproblems.

Für die Summenzuordnung ist eine Reduktion der Bewertungsmatrix notwendig. Diese Reduktion dient dazu, die gegebene Matrix in eine Bewertungsmatrix zu überführen, die die gleiche minimale Zuordnung hat, wie die Ursprüngliche, in der jedoch so viele Bewertungszahlen null sind, dass eine Minimalzuordnung direkt angegeben werden kann.

Dazu wird die Ausgangsmatrix sowohl einer Zeilen- als auch einer Spaltenreduktion unterzogen. In welcher Reihenfolge die beiden Einzelreduktionen vorgenommen werden ist dabei irrelevant. Nach der Zeilen- und Spaltenreduktion dürfen keine negativen Bewertungen bestehen. Diese Voraussetzung sollte erfüllt sein, um zu gewährleisten, dass auch der Zielfunktionswert immer nichtnegativ ist. Es sollten jedoch in beiden Verfahren möglichst große Werte gefunden werden, um die Matrix effizient zu reduzieren.

Bei der Zeilenreduktion wird daher im ersten Schritt der kleinste Wert jeder Zeile ermittelt und in einer zusätzlichen Spalte mit der Bezeichnung q_i gespeichert. Jedes q_i wird anschließend von jedem Wert der zugehörigen Zeile abgezogen und in der zeilenreduzierten Matrix eingetragen. Die Vorgehensweise der Spaltenreduktion verläuft analog. Hierbei wird für jede Spalte der kleinste Wert ermittelt und in der zusätzlichen Zeile p_j vermerkt. Dieser wird ebenfalls von jedem zugehörigen Spaltenwert abgezogen und in einer neuen Matrix eingetragen.

Sind Spalten- und Zeilenreduktion vorgenommen, so liegt die vollständig reduzierte Matrix vor. Eine minimale Zuordnung besteht, wenn ein Zielfunktionswert von null gefunden werden kann. Die Zielfunktionen der Ausgangsmatrix und der reduzierten Matrix (K und K_R) unterscheiden sich um einen konstanten Wert, so dass mit der Minimallösung der reduzierten Matrix auch die Minimallösung für die Ausgangsmatrix gefunden ist.

Größere Probleme lassen sich mit dem Verfahren der Reduktion der Bewertungsmatrix jedoch nicht vernünftig bearbeiten. Die *Ungarische Methode* basiert auf einer erweiterten Reduktion der Bewertungsmatrix. Zunächst erfolgt eine Zeilen- und Spaltenreduktion. Die Zielfunktion wird dabei gegenüber der Ausgangsmatrix um eine Konstante reduziert. Anschließend wird eine Nullzuordnung vorgenommen.

Um zu zeigen, ob eine Nullzuordnung existiert und um diese anzugeben kann wie folgt vorgegangen werden. Vorerst werden alle Zuordnungen getätigt, die zwangsläufig sind. Dieses sind genau die Nullbewertungen, die alleine in einer Zeile- oder Spalte stehen. Lässt sich erkennen, dass keine Zuordnung besteht, in der nur Null-

bewertungen vorkommen (ist $K_R > 0$), so dient die Ungarische Methode der nun notwendigen weiteren Reduktion. Voraussetzung ist, dass in der bisher getätigten Reduktion maximal r (mit $r < n$) Nullzuordnungen möglich sind.

Die Reduktionsvorschrift der Ungarischen Methode lautet:

1. In der reduzierten Matrix r Zeilen und Spalten so überdecken, dass keine Nullbewertungen mehr in der Restmatrix vorhanden sind.
2. Das kleinste Element aus der Restmatrix ermitteln und
 - (a) von allen Elementen der Restmatrix subtrahieren und
 - (b) zu doppelt überdeckten Elementen addieren (d.h. von Spalten und Zeilen überdeckte Elemente).
3. Erhaltene Werte als weiter reduzierte Matrix ausgeben.
4. Ist keine Zuordnung mit $K_R = 0$ möglich, wende die Ungarische Methode auf die neue Matrix erneut an.

Die Verfahren zur Lösung des Engpasszuordnungsproblems basieren auf der Idee, die Zuordnungsmöglichkeiten zu sperren, deren Bewertungszahlen eine gegebene Schranke überschreiten. Nach [Dürr and Kleibohm, 1992] ist ein Beispiel für eine Schranke das Maximum der Zeilenminima, da in jeder Zeile eine Zuordnung vorgenommen werden muss. Nach dem Sperren der unzulässigen Bewertungszahlen muss überprüft werden, ob eine Nullzuordnung möglich ist. Eine Möglichkeit dies vorzunehmen ist das Anwenden der Ungarischen Methode.

XMinAlternation: Summenzuordnung

Es wird also ein Zuordnungsproblem in jedem Netzwerkknoten formuliert. Jede mögliche Kombination von eingehenden und ausgehenden Flusseinheiten wird mit einer Bewertungszahl versehen. Diese gibt an, wie viele unterschiedliche Linien ein Umlaufstück enthalten würde, das durch die Verknüpfung dieser Flusseinheiten entstehen würde.

Die eigentliche Dekomposition nach XMinAlternation geschieht nun, indem sequentiell in jedem Knoten die Flusseinheiten mit Hilfe der ungarischen Methode verknüpft werden. Da die Ungarische Methode in jedem Knoten aufgerufen wird, müssen bei den realistischen Instanzen hunderttausende Zuordnungsprobleme gelöst werden. Aus diesem Grund wird eine sehr effiziente Implementierung benötigt.

Die Anzahl der Linienwechsel wird hierdurch im Vergleich zu den vorher vorgestellten Strategien verkleinert. Das Ergebnis des Algorithmus ist aber im Bezug

auf Linienreinheit bei gegebenen minimalen Kosten des Umlaufplans nicht gesamt-optimal. Er handelt lokal in der Hinsicht, dass nicht in die „Zukunft“ geschaut wird, sondern nur das bereits durchlaufene Netzwerk betrachtet wird. Die neuen Entscheidungen basieren somit nur auf den bisher getroffenen Entscheidungen.

***EngPassMinAlternation*: Engpasszuordnung**

In der Praxis wird oft verlangt, dass die Anzahl verschiedener Linien in einem Fahrzeugumlauf oder in einem Umlaufteil zwischen zwei Depot-Aufenthalten, einen bestimmten Grenzwert nicht übersteigt. Dieser Wert ist aber in der Regel mehr oder weniger willkürlich festgelegt, vielmehr geht es darum, die maximal auftretende Anzahl verschiedener Linien pro Umlauf zu reduzieren. Dies wird durch eine geringfügig modifizierte Dekompositionsstrategie angestrebt. Dabei wird in jedem Knoten ein Engpasszuordnungsproblem (im Gegensatz zum Summenzuordnungsproblem bei XMinAlternation) gelöst.

5.2.5 Vorsortierung der Haltestellen

Aufgrund der verschiedensten Ergebnisse der Tests wurde eine Sortierung der Netzwerkknotten nach der Größe der Timelines implementiert. Dies ermöglicht eine chronologische Ausführung der Dekomposition, so dass zu Anfang die Knoten mit trivialen Problemen und nachfolgend größere Knoten mit komplexen Problemen gelöst werden.

Um den Nachteil der Lokalität der Entscheidungen zu überwinden, können weitere Dekompositionsstrategien eingesetzt werden. Im folgenden Abschnitt wird eine globale Flussdekomposition vorgestellt, die durch das Lösen eines Set Partitioning Problems realisiert wird. Es müssen dabei nicht mehr, wie bei den letzten beiden Strategien, die Zuordnungsprobleme in jedem Knoten einer Netzwerkschicht gelöst werden, sondern nur ein Set Partitioning Problem pro Netzwerkschicht.

5.2.6 Globale Dekompositionsstrategie: *SPP*

Die Abkürzung SPP steht im weiteren Text für das Set Partitioning Problem. Bei dieser Strategie wird zunächst eine Menge möglicher Fahrzeugumläufe durch die Wege im optimalen MDVSP-Fluss gebildet. Diese Umläufe stellen die Spalten des Set Partitioning Problems dar. Daraus soll eine Teilmenge der möglichen Umläufe mit Hilfe eines MIP-Optimierers ausgewählt werden. Diese Teilmenge soll jede abzudeckende Fahrzeugaktivität genau einmal enthalten und dabei den bestmöglichen Wert einer vordefinierten Zielfunktion liefern. Im Folgenden wird eine Optimierung im Bezug auf die Linienreinheit beschrieben, um die Ergebnisse mit den Ergebnissen der zuvor beschriebenen lokalen Dekompositionsstrategien vergleichen zu

können. Die Zielfunktion des Set Partitioning Problems in der Flussdekomposition kann aber auch anders definiert werden - z.B. als Optimierung der Pausenverteilungen oder anderer Anforderungen.

Aus allen kostenoptimalen Umlaufplänen, die durch einen optimalen Fluss im MDVSP-Netzwerk vorgegeben sind, wählt diese Strategie den Umlaufplan, der am besten im Bezug auf Linienwechsel ist.

Die mathematische Formulierung ist ähnlich aufgebaut wie das Set Partitioning Problem 3.11-3.14 aus dem Kapitel 3. Für das Set Partitioning Problem (SPP) wird die SPP-Matrix gebildet, deren Spalten alle Möglichkeiten eines Umlaufs darstellen und deren Zeilen jede Fahrzeugaktivität repräsentieren, die in den Umläufen vorkommen muss. Jede Zeile, die eine Fahrt darstellt ist in dem LP-Modell nun als Restriktion anzusehen, deren rechte Seite immer die Anzahl des Vorkommens dieser einen Fahrt in dem Modell beschreibt. Auf der linken Seite der Restriktion steht in jeder Spalte, d.h. in jedem Umlauf, in dem die Fahrt vorkommt, eine Eins, sonst eine Null. Durch diese Anordnung der Restriktionen wird sichergestellt, dass jede Fahrt genau so oft bedient wird, wie sie im Modell enthalten ist. Der Optimierer berechnet nun die optimalen Kombinationen von Umläufen, die in dem Modell möglich sind, unter der Voraussetzung, dass jede Fahrt genau einmal bedient wird.

Weiterhin benötigt der Optimierer noch die Zielfunktion des LP-Modells. In dieser werden für jeden Umlauf die Kosten bestimmt und vom Optimierer minimiert. Mithilfe dieses Modells ist es möglich, aus einem Satz gegebener Umläufe die herauszufinden, die bezüglich der in der Zielfunktion angegebenen Kosten optimal sind.

5.3 Numerische Ergebnisse

Die Ergebnisse zahlreicher Testläufe belegen, dass die auf der Lösung von Zuordnungsproblemen basierte Dekompositionstrategien deutlich bessere Berücksichtigung von Linien ermöglichen, als die anderen lokalen Strategien, die einen höheren Grad an „Kurzichtigkeit“ aufweisen.

Dekomp.-Strategie	city 1	tp9	tp17	tp30
FIFO	00:00.03	00:00.00	00:00.00	00:00.00
LIFO	00:00.03	00:00.01	00:00.01	00:00.00
MinAlternation	00:00.15	00:00.01	00:00.01	00:00.01
XMinAlternation	00:14.19	00:00.11	00:00.21	00:00.09
SPP	22:58.56	00:04.15	00:07.86	00:02.32

Tabelle 5.1: Laufzeitvergleich der verschiedenen Dekompositionsstrategien

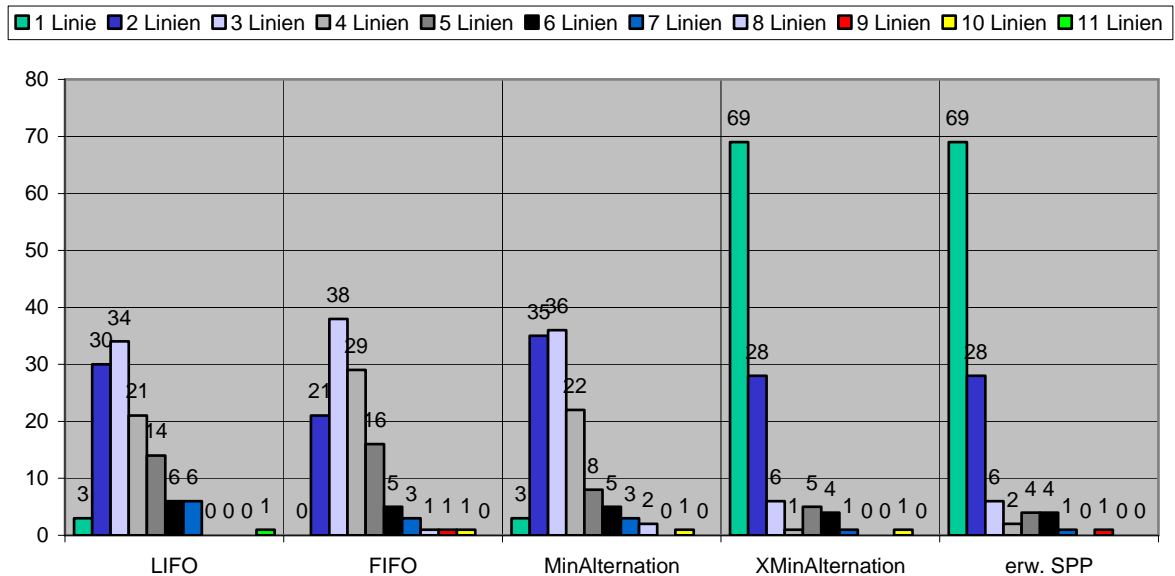


Abbildung 5.3: Ergebnisvergleich der verschiedenen Dekompositionsstrategien

Im Folgenden sollen die Laufzeiten der fünf Dekompositionsstrategien miteinander verglichen werden.

Die in der Tabelle 5.1 enthaltenen Laufzeiten (Min:Sec) der verschiedener Dekompositionsstrategien zeigen, dass die SPP-Dekomposition momentan noch viel höhere Laufzeiten als die anderen Dekompositionsstrategien hat.

Nachfolgend sollen die Ergebnisse, die für den Datensatz city 1 gewonnen wurden, bezüglich der Linienberücksichtigung verglichen werden.

Die Strategie MinAlternation liefern deutlich schlechtere Ergebnisse als die beiden XMinAlternation Strategien und die SPP-Dekomposition. Die Ergebnisse von XMinAlternation und der SPP-Dekomposition sind ziemlich ähnlich, jedoch wird ersichtlich, dass einzelne Werte verbessert werden konnten. Durch die Nonzeros- und Zeitbegrenzung kann die Laufzeit der SPP-Dekomposition gesteuert werden. Durch die Vorberechnung durch XMinAlternation liefert der Algorithmus immer ein Ergebnis, das allerdings, falls die Zeit für die Lösung begrenzt wird, suboptimal sein kann. Insgesamt betrachtet ist das Ergebnis der SPP-Dekomposition immer das Beste von allen Strategien. Die Effizienz dieses Verfahrens kann insbesondere im Spaltengenerierungsteil noch erheblich gesteigert werden, wenn die Wegesuche im SPP-Netzwerk optimiert wird.

Im nächsten Kapitel werden Möglichkeiten für die Berücksichtigung verschiedener Anforderungen aus der Praxis der Umlaufplanung erarbeitet. Unter anderem wird eine Berücksichtigung von Linien im Netzwerkmodell beschrieben, die zu weiteren Verbesserungen im Bezug auf die Linienreinheit führt. Die Ergebnisse dieser

Modellerweiterung werden mit den Ergebnissen der Dekompositionsstrategien verglichen. Daher wird in diesem Abschnitt auf eine ausführlichere Zusammenfassung der Testergebnisse verzichtet - sie werden im nächsten Kapitel diskutiert.

Kapitel 6

Einbeziehung praktischer Randbedingungen und Zielsetzungen

In diesem Kapitel werden Modellerweiterungen zur Berücksichtigung verschiedener praktischer Aspekte im Aufbau des Netzwerkmodells vorgestellt. Dazu zählen die Linienberücksichtigung, die Verschiebeintervalle für Fahrten, das Konzept der gleitenden Depot-Fahrzeugtypzuordnung sowie verschiedene Arten von Kapazitätsanforderungen.

6.1 Multikriterielle Optimierung: Kosten und Liniereinheit

Unterschiedliche Anforderungen in der Praxis der Umlaufplanung können erfordern, dass bestimmte Anschlüsse begünstigt oder verhindert werden sollen. Insbesondere wurden von Planungsverantwortlichen der Verkehrsunternehmen Wünsche geäußert, bestimmte Pausenverteilungen bei einem Aufenthalt des Fahrzeugs im Depot anzustreben und möglichst Fahrten gleicher Linie an den Haltestellen zu verknüpfen.

Dies sind die Anforderungen, die in der Regel erst bei der Erstellung von Fahrerdiensten wichtig werden. Trotzdem lohnt es sich, diese Aspekte bereits in die Planung der Fahrzeugumläufe einzubeziehen, weil auf diese Weise eine Teilintegration der beiden Planungsphasen ermöglicht wird, was zu Einsparungen in den Gesamtkosten führen soll.

Im Folgenden wird eine Erweiterung des Netzwerkmodells aus Kapitel 4 vorgestellt, die eine solche Teilintegration ermöglicht. Die Modellierungstechnik wird am Beispiel der Berücksichtigung von Linien erläutert und kann auf ähnliche Weise für die Pausen-Problematik verwendet werden.

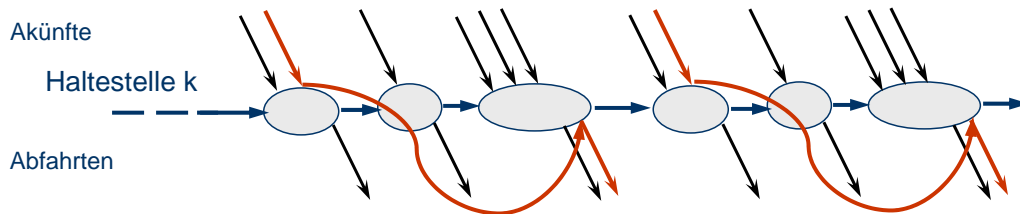


Abbildung 6.1: Line Arcs in Time-Space Netzwerk

6.1.1 Linienkanten im Netzwerkmodell

Neben den in Kapitel 5 beschriebenen spezialisierten Dekompositionsstrategien kann die Anzahl der Linien pro Umlauf durch eine erweiterte Modellierung mit zusätzlichen Linienkanten im MDVSP-Netzwerk geregelt werden. Wie in Abbildung 6.1 an einem einfachen Beispiel gezeigt, verbindet in einer Haltestellen-Timeline eine Linienkante jede Ankunft mit der nächsten Abfahrt gleicher Linie. Diese Kanten werden mit negativen Kosten versehen, die als Bonus für „Linienreinheit“ gelten. Dieser Bonus ist relativ klein im Vergleich zu den anderen Kosten, so dass die Linienkanten zwar bevorzugt verwendet werden, allerdings die Kostenoptimalität der Lösung nur minimal beeinflussen und insbesondere die Anzahl der Fahrzeuge unverändert lassen.

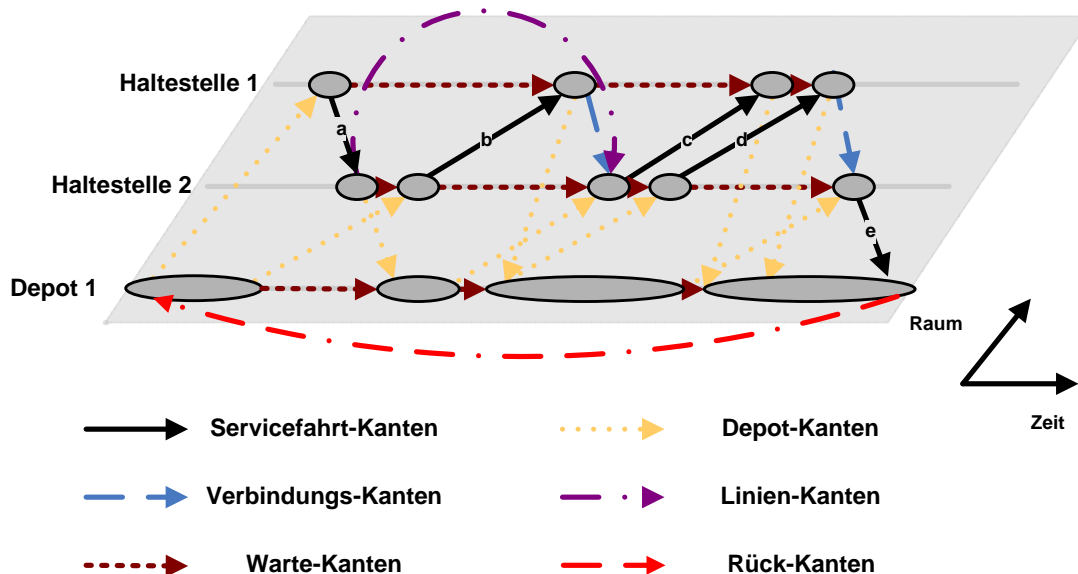


Abbildung 6.2: Einfügen der Linienkanten

Die *Linienkanten* werden also optional in das Netzwerk aufgenommen und stellen eine Möglichkeit dar, Servicefahrten mit gleicher Linie bevorzugt in einen Um-

lauf aufzunehmen. Dadurch kann die Anzahl der verschiedenen Linien in einem Umlauf reduziert werden. Linienkanten haben einen maximalen Fluss von Eins, da nur die betroffenen Servicekanten verbunden werden sollen. Eine Linienkante wird im Netzwerk zwischen jedem Endpunkt einer Servicefahrt-Kante und dem in der gleichen Timeline liegenden Startpunkt der nächsten Servicefahrt-Kante eingefügt, wenn beide Servicefahrt-Kanten die gleiche Linien bedienen. Die Abbildung 6.2 zeigt am Beispiel aus Abschnitt 4.1.4 wie eine Linienkante (violett) eingefügt wird, die die beiden Servicefahrt-Kanten der Linie 1 verbindet. In der abgebildeten Netzwerkschicht wird also eine Linienkante zwischen dem Ende der Servicefahrt-Kante *a* und dem Anfang der Servicefahrt-Kante *c* eingefügt, da beide Servicefahrten die Linie 1 bedienen (siehe Abschnitt 4.1). Für alle anderen Servicefahrt-Kanten in dem Beispiel können keine weiteren Linienkanten eingefügt werden. Durch ihre negativen Kosten wird die neue Kante bevorzugt in die Lösung aufgenommen und sorgt somit dafür, dass die beiden Kanten der Linie 1 hintereinander von einem Fahrzeug werden. Eine optimale Flusslösung, in der die Service-Fahrten *a* und *c* durch die Linienkante verknüpft werden ist in der Abbildung 6.3 gezeigt.

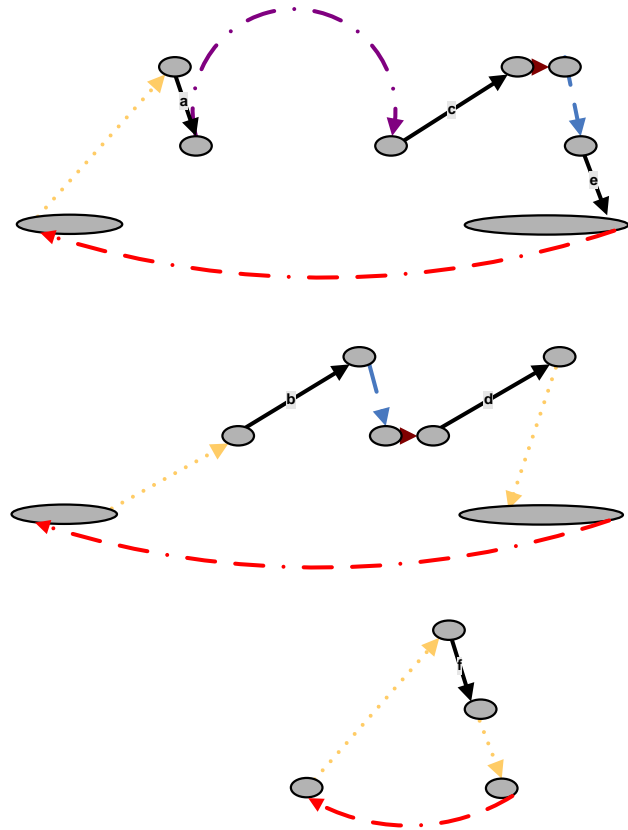


Abbildung 6.3: Einfügen der Linienkanten

6.1.2 Erweiterung des mathematischen Modells

Da Linienkanten negative Kosten haben, werden sie bei der Findung eines optimalen Flusses bevorzugt bedient. Deshalb muss durch einen maximalen Fluss von eins für Linienkanten sowie durch weitere Restriktionen sichergestellt werden, dass die Linienkanten nur den gewünschten Effekt der Minimierung von Linienwechseln ermöglichen und das Modell nicht anderweitig beeinflusst wird. Durch Restriktionen muss sichergestellt werden, dass eine Linien-Kante nur benutzt werden kann, wenn beide Servicefahrt-Kanten, die sie verbindet, benutzt werden. Sollte eine der beiden Fahrten also von einer Fahrzeugtyp-Depot-Kombination einer anderen Schicht bedient werden, so darf die Linien-Kante nicht benutzt werden. Folgende Definitionen

sollen zusätzlich zum mathematischen Modell 4.8-4.14 gelten:

- $s1_e$ vorausgehende Kante der Linien-Kante e , $e \in EL^s, s \in S$
- $s2_e$ nachfolgende Kante der Linien-Kante e , $e \in EL^s, s \in S$

Linienkanten haben geringe negative Kosten (6.1), die eine bevorzugte Befahrung von gleichen Linien garantieren. Sie werden vom Benutzer festgelegt, um einen Trade-Off zwischen der Kostenoptimalität und der Linienreinheit zu steuern.

$$c_e = -\epsilon \quad \forall e \in EL^s, \forall s \in S \quad (6.1)$$

Die IP-Formulierung aus dem Kapitel 4 soll nun mit zusätzlichen Restriktionen versehen werden:

$$x_e \leq x_{s1_e} \quad \forall e \in EL^s, \forall s \in S \quad (6.2)$$

$$x_e \leq x_{s2_e} \quad \forall e \in EL^s, \forall s \in S \quad (6.3)$$

Restriktion 6.2 besagt, dass eine Linien-Kante nur benutzt werden darf, wenn ihre vorausgehende Kante benutzt wird und die Restriktion 6.3 garantiert, dass eine Linien-Kante dann benutzt wird, wenn ihre nachfolgende Kante benutzt wurde. Diese Restriktionen garantieren also, dass eine Linienkante nur dann in einer Lösung benutzt werden darf, wenn beide Servicefahrt-Kanten, die sie verbindet, ebenfalls in dieser Lösung benutzt wurden (Flusswert ist größer als null). Die obere Flussschranke auf einer Linienkante wird implizit durch die Restriktionen 6.2 und 6.3 auf eins begrenzt, da der Fluss auf den beiden entsprechenden Servicefahrt-Kanten bereits eine obere Schranke von eins erfüllen muss.

6.1.3 Ergebnisse

Der beschriebene Ansatz zur Berücksichtigung von Linien im Netzwerkmodell wurde auf einer Teilmenge der Instanzen getestet, die schon in Kapitel 4 benutzt wurden. Nachfolgend sollen die Ergebnisse der Testläufe im Bezug auf die Lösungsqualität und Laufzeiten beschrieben werden.

Die Tabelle 6.1 beschreibt die Änderungen in der Modellgröße und die Optimierungszeit für die Modelle mit Linienkanten. Die Kantenanzahl wächst dabei um die Anzahl der Linienkanten, die im Netzwerk eingefügt wurden. In der Tabelle 6.1 ist diese Anzahl sowie die prozentuale Veränderung der Gesamtanzahl von Kanten in der Spalte #Linienkanten angegeben. Die Variablen im mathematischen Modell entsprechen den Flussgrößen auf den Kanten im Netzwerk. Daher wächst auch die

Instanz	#Linienkanten	Modellgröße			CPU(sec)	
		#Zeilen	#Spalten	#Nonzeros	IPM	Dual
Saar	1154 (2,8%)	9361	41275	89764	14	3
hal	5272 (25%)	26308	62370	152402	98	279
mun14	4076 (4%)	21134	104124	229458	273	28
mun15	4892 (3,9%)	24489	129083	283386	274	45
mun28	19812 (4,1%)	91940	498635	1100035	3305	> 39600
mun18	23004 (3,4%)	108156	698337	1517430	4851	4619

Tabelle 6.1: Modellgröße und Laufzeit mit Linienkanten

Instanz	#Fahrzeuge	Fixe Kosten	#Blöcke			operative Kosten		
			mit LK	ohne LK	Abw.	mit LK	ohne LK	Abw.
Saar	49	32300000	84	95	-11,6%	156305	153469	1,8%
hal	115	10050	817	824	-0,8%	6700	6557	2,2%
mun14	107	18029500	122	129	-5,4%	245246	238702	2,7%
mun15	94	11797000	111	124	-10,5%	249259	243467	2,4%
mun28	56	11928000	84	121	-30,6%	118219	113022	4,6%
mun18	82	17466000	120	210	-42,9%	160257	152669	5,0%

Tabelle 6.2: Vergleich der Lösungen mit und ohne Linienkanten (LK)

Anzahl der Variablen um die gleichen Anteile. Die entsprechende Vergrößerung des mathematischen Modells führt tendenziell zu längeren Optimierungszeiten.

Eine Verlängerung der Optimierungszeit tritt bei allen Instanzen auf und ist am stärksten bei der Instanz *hal* ausgeprägt. Dies ist dadurch zu erklären, dass die Anzahl der Kanten und somit der Variablen bei dieser Instanz um 25% durch das Hinzufügen der Linienkanten gewachsen ist.

Die Tabelle 6.2 beschreibt die Eigenschaften der optimalen Lösungen mit Berücksichtigung der Linien und einen Vergleich mit den Lösungen der Basismodelle aus dem Kapitel 4 ohne Linienkanten (vgl. Tabelle 4.5). Die neuen Lösungen verursachen die gleichen Fixkosten für die Fahrzeuge, da die minimale Anzahl der Fahrzeuge durch die Berücksichtigung von Linien unverändert geblieben ist. Die neuen Lösungen verursachen aber höhere operative Kosten. Dies liegt in der Natur der verwendeten Modellierung. Durch die speziellen Dekompositionsstrategien in Kapitel 5 wurde die Anzahl der Linien pro Umlauf reduziert, ohne auf die Kostenoptimalität zu verzichten. Das primäre Ziel war also eine Kostenminimierung. Die hier vorgenommene Modellerweiterung verfolgt hingegen zwei konkurrierende Ziele. Zum einen werden weiterhin die Gesamtkosten des Fahrzeugeinsatzes minimiert, zum anderen - die Anzahl der Linien pro Umlauf. Es handelt sich also um ein multikriterielles Optimierungsproblem ohne Festlegung einer Hierarchie der Ziele. Wie stark das jeweilige Ziel gewichtet sein soll wird durch einen Bonusfaktor für die Benutzung einer Linienkante geregelt.

Wie der Tabelle 6.2 zu entnehmen ist, wird die bessere Verteilung der Linien auf die Umläufe durch ein nur geringfügiges Anwachsen der operativen Kosten (1,8 bis 5%) erkauft. An den Gesamtkosten gemessen ist der Kostenanstieg vernachlässigbar klein.

Die zweite Kennzahl der berechneten Umlaufpläne, die sich im Vergleich zum Basismodell ohne Linienkanten geändert hat ist die Anzahl der Depot-Blöcke. Depot-Blöcke sind Umlaufstücke zwischen zwei aufeinander folgenden Aufenthalten im Depot. Die Ergebnisse im mittleren Bereich der Tabelle 6.2 zeigen, dass die Anzahl solcher Blöcke gesunken ist. Dies bedeutet, dass ein Fahrzeug im Laufe des Tages nicht mehr so oft an einem Depot vorbeikommt.

Die Abbildung 6.4 visualisiert für drei Instanzen die Qualität der Linienberücksichtigung in jeweils fünf auf unterschiedliche Weise berechneten Umlaufplänen. Die ersten vier Umlaufpläne sind dabei durch verschiedene Dekompositionsstrategien aus einem kostenoptimalen Fluss extrahiert worden (FiFo, MinAlt, XMinAlt und SPP - vgl. Kapitel 5). Ein fünfter Umlaufplan wurde für jede Instanz mit Verwendung der Linienkanten im Modell berechnet. Diese fünf Lösungen werden in jeweils drei Grafiken auf unterschiedliche Weise in Bezug auf die Linienreinheit verglichen.

Die jeweils mittlere Grafik der Abbildung 6.4 zeigt die maximal auftretende Linienanzahl pro Umlauf für die jeweilige Strategie. Die rechte Grafik zeigt die Verteilung der Umläufe in Abhängigkeit von der Linienanzahl. Die Säulen-Darstellung rechts wird wie folgt interpretiert: Für die Instanz *hal* liefert die Anwendung der Linienkanten-Strategie Umläufe mit folgenden Eigenschaften: 96 Umläufe sind linienrein, 11 Umläufe beinhalten Fahrten von zwei Linien, fünf Umläufe beinhalten Fahrten von drei verschiedenen Linien und drei Umläufe beinhalten vier unterschiedliche Linien. Der Linienverlauf in der linken Grafik zeigt die kumulierte Anzahl Umläufe mit einer bestimmten Qualität. Ein Datenpunkt gibt an wie viele Umläufe nicht mehr als die gegebene Anzahl von Linien beinhalten. Anhand dieser Grafik kann man am besten die Strategien bewerten. Je weiter links sich eine Kurve befindet und je früher der Sättigungszustand von dieser Kurve erreicht ist, desto besser ist die jeweilige Strategie in Bezug auf die Linienberücksichtigung. Für alle untersuchte Instanzen lieferte die Strategie „Linienkanten“ die besten Ergebnisse. Da die operativen Kosten der Umlaufpläne und die Laufzeiten dabei nur geringfügig schlechter sind und die Anzahl Fahrzeuge sowie die fixen Fahrzeugkosten optimal bleiben, ist diese Strategie für den praktischen Einsatz zu empfehlen.

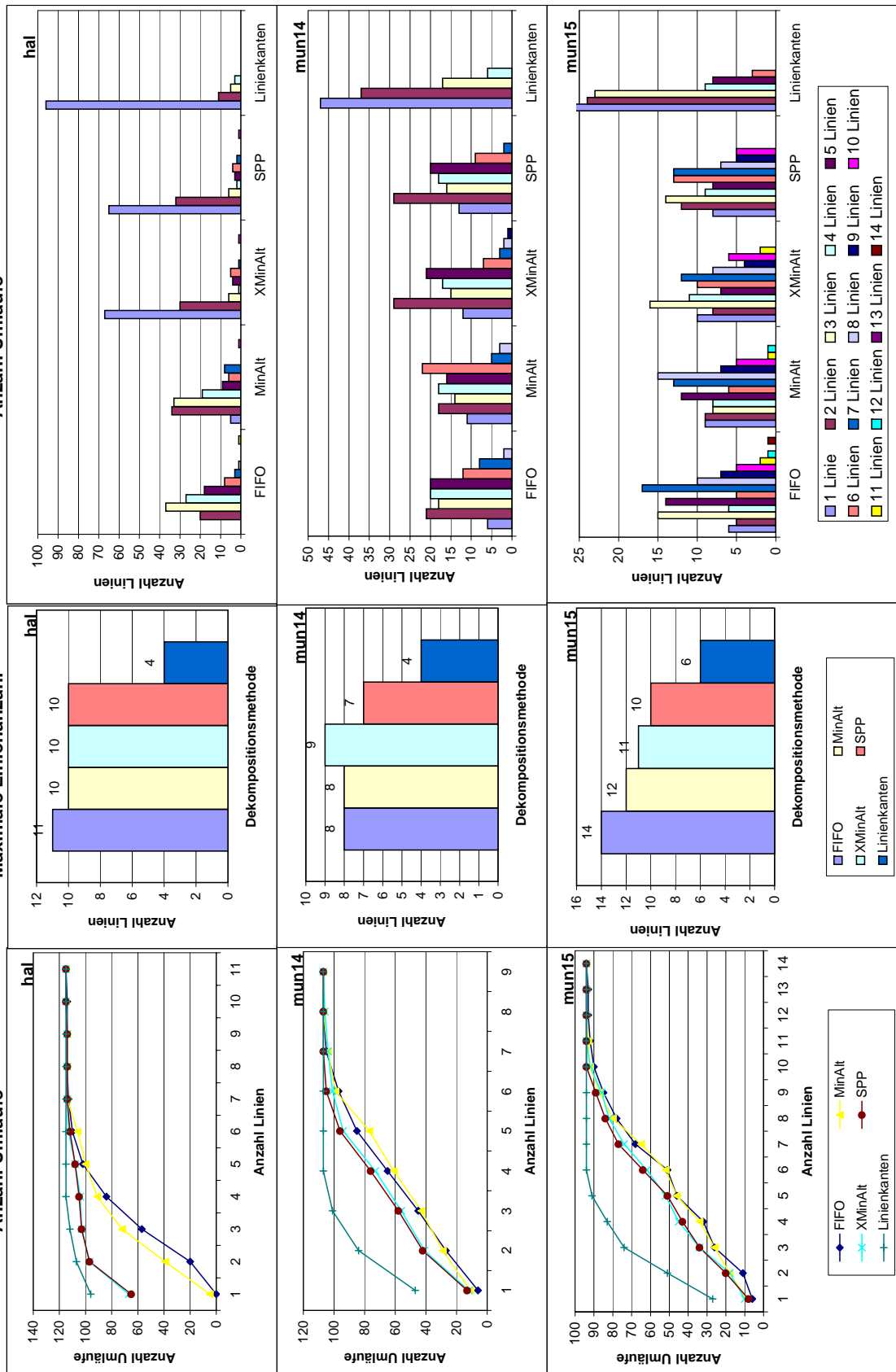


Abbildung 6.4: Ergebnisse für drei Instanzen

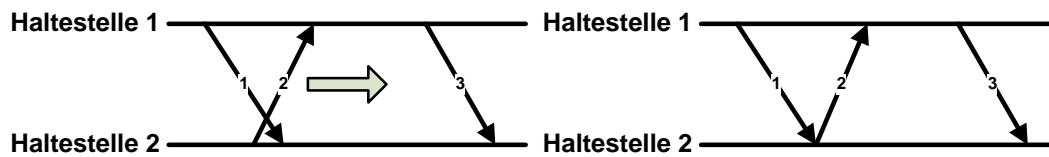


Abbildung 6.5: Potential von Zeitfenstern

6.2 Verschiebeintervalle für Fahrten

Durch geringfügiges zeitliches Verschieben von Fahrten kann unter Umständen eine Einsparung im Fahrzeugbedarf erzielt werden. Dabei kann zum einen das zulässige Verschiebeintervall unterschiedlich groß sein und zum anderen kann die Menge der verschiebbaren Fahrten unterschiedlich definiert werden. So dürfen in der Regel die Schulfahrten um längere Zeit verschoben werden als normale Taktfahrten. Der Wunsch, eine Fahrt nicht zu verschieben, kann verschieden gewichtet sein. Das Potential der zeitlichen Fahrtenverschiebungen liegt in den besseren Anschlussmöglichkeiten, so dass Fahrtenverknüpfungen zustande kommen können, die ohne Verschiebung nicht möglich wären. Dies wird in der Abbildung 6.5 verdeutlicht. Während die links dargestellten Servicefahrten nur von zwei verschiedenen Fahrzeugen ausführbar wären, können sie mit der rechts dargestellten zeitlichen Verschiebung von einem Fahrzeug nacheinander gefahren werden.

Das Optimierungsziel ist hierbei, mit möglichst geringen Verschiebungen möglichst viele Fahrzeuge einzusparen. Die Einsparung in den operativen Kosten ist zweitrangig bzw. überhaupt nicht erstrebenswert. Wenn keine Fahrzeuge eingespart werden können, sollten auch keine Fahrten verschoben werden. In der Modellierung muss die Zielfunktion entsprechend angepasst werden, so dass die Strafkosten für die Verschiebung von Fahrten möglichst nicht durch die eingesparten operativen Kosten kompensiert werden.

Die gängigen Modellierungsansätze aus der Literatur wie [Desaulniers et al., 1998] und [Suhl, 1995] bilden die möglichen Verschiebungen von Fahrten bzw. Flügen durch die Einführung kontinuierlicher Variablen für die Abfahrts- bzw. Abflugszeiten. Diese Lösung ist für die hier beschriebene Entscheidungssituation in der Fahrzeugumlaufplanung nicht sinnvoll, denn die Abfahrts- und Ankunftszeiten werden in diskreten Zeitintervallen festgelegt. Außerdem sind die Verschiebeintervalle relativ klein und werden in der Praxis nur für eine Teilmenge der Fahrplanfahrten zugelassen. Da die Ankunfts- und Abfahrtsfahrten der Fahrten in diskreten Zeitintervallen geplant sind, betragen auch die zeitlichen Fahrtenverschiebungen immer ein vielfaches der entsprechenden Zeiteinheit, wie z.B. in [Daduna and Völker, 1997] angenommen wurde. Ohne Beschränkung der Allgemeinheit nehmen wir in den folgenden Ausführungen an, dass die Planung minutengenau erfolgt und somit die Fahrten nur um eine ganze Anzahl von Minuten verschoben werden dürfen. Im

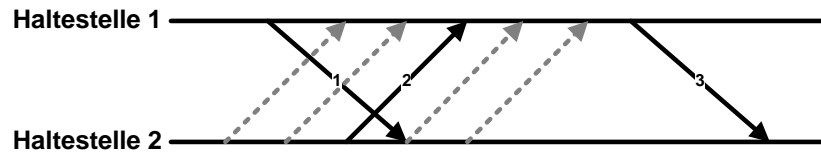


Abbildung 6.6: Verwendung von Zeitfenster-Kanten

nächsten Abschnitt wird die Erweiterung des Netzwerkmodells aus dem Kapitel 4 beschrieben.

6.2.1 Erweiterung des Netzwerkmodells

Um die möglichen Fahrtenverschiebungen im Netzwerkmodell abzubilden, wird ein neuer Typ von Kanten für das Netzwerk definiert: die so genannten **Zeitfenster-Kanten**. Wenn eine Servicefahrt-Kante in einem vorgegebenen Zeitfenster verschiebbar sein soll, werden mehrere Zeitfenster-Kanten in die Umgebung der Servicefahrt-Kante eingefügt. Für eine Zeitfenster-Kante, die eine Verfrühung der Servicefahrt ermöglichen soll, werden die Abfahrts- und Ankunftszeiten entsprechend vermindert; bei einer potentiellen Bedienung zu einem späteren Zeitpunkt entsprechend erhöht.

Dabei ist ein Einfügen der Zeitfenster-Kanten in einem Minutenraster ausreichend, wenn der Fahrplan minutengenau geplant ist. Das Beispiel in Abbildung 6.6 zeigt die Modellierung eines Zeitfensters von ± 2 Minuten für die Servicefahrt 2. Innerhalb des Netzwerks werden die Zeitfenster-Kanten genau wie die normalen Servicefahrt-Kanten behandelt. Bei der Optimierung kann zwischen einer Servicefahrt-Kante und einer der dazugehörigen Zeitfenster-Kanten gewählt werden. Dadurch ist es möglich, eine bessere Lösung im Bezug auf die Anzahl benötigter Fahrzeuge zu finden.

Anhand des in Kapitel 4 gegebenen Beispiel-Fahrplans sollen an dieser Stelle die Änderungen des Netzwerkaufbaus gezeigt werden. Da sich mit den gewählten Beispiel-Verschiebeintervallen keine Änderungen an der ersten Schicht des Netzwerks ergeben, wird nachfolgend nur die zweite Netzwerkschicht behandelt. Dazu wird die Fahrt f des Fahrplans mit einem Verschiebeintervall versehen, wie in der Tabelle 6.3 gezeigt wird. Die Fahrt darf bis zu einer Minute nach vorne verschoben oder bis zu einer Minute nach hinten verschoben werden.

Fahrt	Von	Nach	Abf.-zeit	Ank.-zeit	Intervall	Linie	Depots
f	h_1	h_2	8	10	± 1 Minute	5	d_2

Tabelle 6.3: Servicefahrt mit Zeitfenstern



Abbildung 6.7: Netzwerkaufbau mit Zeitfenstern (Schritt 1)

In Abbildung 6.7 werden für das angegebene Beispiel die Timelines mit den Servicefahrt-Kanten der gegebenen Fahrplanfahrten eingefügt. Zusätzlich werden Zeitfenster-Kanten eingefügt, die relativ zu der Original-Servicefahrt-Kante entsprechend der zeitlichen Verschiebung nach vorne bzw. nach hinten versetzt werden. In diesem Beispiel wird eine Zeitfenster-Kante eine Minute vor und die Zweite - eine Minute nach der originalen Servicefahrt-Kante eingefügt. Im allgemeinen wird eine Zeitfenster-Kante für jede Minute, um die eine Fahrt nach hinten bzw. vorne verschoben werden darf, in das Netzwerk eingefügt.

Die beiden Zeitfenster-Kanten werden im weiteren Verlauf des Netzwerkaufbau genau wie die originale Servicefahrt-Kante behandelt. Dies gilt zum Beispiel für obere und untere Flusschranken und für die Berücksichtigung im mathematischen Modell. Beim Einfügen von Verbindungs-Kanten werden sie auch wie die normalen Servicefahrt-Kanten beachtet.

Da für jede eingefügte Zeitfenster-Kante das Netzwerk um weitere zwei Depot-Kanten sowie eine häufig große Anzahl von Verbindungs- und Warte-Kanten erweitert werden müsste, würde die Modellgröße des aus dem Netzwerk gebildeten mathematischen Modells sehr schnell anwachsen. Außerdem führt das Einfügen der Zeitfenster-Kanten zu einer zusätzlichen Komplexität des zu lösenden Problems, da gerade die „schwierigen“ cover constraints nunmehr die Variablen für die Zeitfenster-Kanten mitberücksichtigen müssen. Das resultierende größere und komplexere mathematische Modell wäre damit unter Umständen nicht mehr mit Hilfe der Standardoptimierungssoftware in annehmbarer Zeit lösbar. Im nächsten Abschnitt wird eine Reduktion der überflüssigen Kanten in dem um die Zeitfenster-Kanten erweiterten Netzwerkmodell vorgestellt.

Netzwerkreduktion

Um die Größe des mathematischen Modells für das MDVSP zu reduzieren, müssen überflüssige Kanten aus dem Netzwerk entfernt werden. Auf der Suche nach den überflüssigen Kanten wird untersucht, wo das Potential der zusätzlich eingefügten Zeitfenster-Kanten liegt. Eine Verschiebung kann nur dann sinnvoll sein, wenn dadurch ein Anschluss ermöglicht wird, der vorher nicht möglich war.

Nachfolgend werden alle im Bezug auf die originale Servicefahrt-Kante vorgelagerten Zeitfenster-Kanten die *verfrühten Zeitfenster-Kanten* genannt und alle nachgelagerten - die *verspäteten Zeitfenster-Kanten*. Eine verfrühte Zeitfenster-Kante ermöglicht eine frühere Ausführung der Fahrt. Das Potential liegt in einer früheren *Ankunft* an der Endhaltestelle dieser Fahrt. Dadurch kann es eine weitere Fahrt bedienen, deren Abfahrt nach der normalen aber vor den verspäteten *Ankunft* der betrachteten Fahrt liegt. Eine verspätete Zeitfenster-Kante ist ebenfalls nur dann sinnvoll, wenn durch die Verschiebung der Abfahrt ein Anschluss möglich wird, der vorher nicht möglich war. Dabei kann die verspätete *Abfahrt* ein Einsparungspotential bringen. Dadurch wird ermöglicht, dass diese Fahrt von einem Fahrzeug bedient wird, das erst nach der normalen Abfahrtszeit an der Starthaltestelle der zu untersuchenden Fahrt ankommt.

Für alle verfrühten Zeitfenster-Kanten werden ihre Ankunftsereignisse in der Endhaltestelle untersucht. Dabei wird folgende Regel für die Überflüssigkeit einer verfrühten Zeitfenster-Kante abgeleitet: *zwischen der zu untersuchenden verfrühten Ankunft und der normalen Ankunft (bzw. der nächsten Ankunftsöglichkeit für die entsprechende Fahrt) liegt keine Abfahrt.*

Für alle verspäteten Zeitfenster-Kanten werden ihre Abfahrtsereignisse in der Starthaltestelle untersucht. Dabei ist eine verspätete Zeitfenster-Kante überflüssig, wenn gilt: *zwischen der normalen Abfahrt (bzw. einer vorherigen Abfahrtsmöglichkeit für diese Fahrt) und der zu untersuchenden verspäteter Abfahrt keine Ankunft.*

Das Beispiel in der Abbildung 6.8 soll diese Reduktion verdeutlichen. Die Abbildung zeigt einen Netzwerk-Ausschnitt mit einer Servicefahrt-Kante, vier Zeitfenster-Kanten sowie weiteren Abfahrten und Ankünften. Für die erste verfrühte Zeitfenster-Kante wird untersucht, ob eine Abfahrt zwischen ihrer Ankunft und der Ankunft der nächsten Zeitfenster-Kante liegt. Da dies nicht der Fall ist, kann sie aus dem Netzwerk gelöscht werden. Für die nächste verfrühte Zeitfenster-Kante liegt eine Abfahrt im kritischen Bereich, so dass sie einen neuen Anschluss ermöglicht. Deshalb bleibt diese Kante im Netzwerk.

Bei der ersten verspäteten Zeitfenster-Kante sehen wir, dass keine Ankunft in dem kritischen Bereich (zwischen der Abfahrt der Servicefahrt-Kante und der verspäteter Abfahrt) liegt. Somit ist diese Zeitfenster-Kante überflüssig und kann entfernt werden. Die letzte verspätete Zeitfenster-Kante ist nicht überflüssig, da eine Ankunft kurz vor ihrer Abfahrt stattfindet und sich somit andere Anschlussmög-

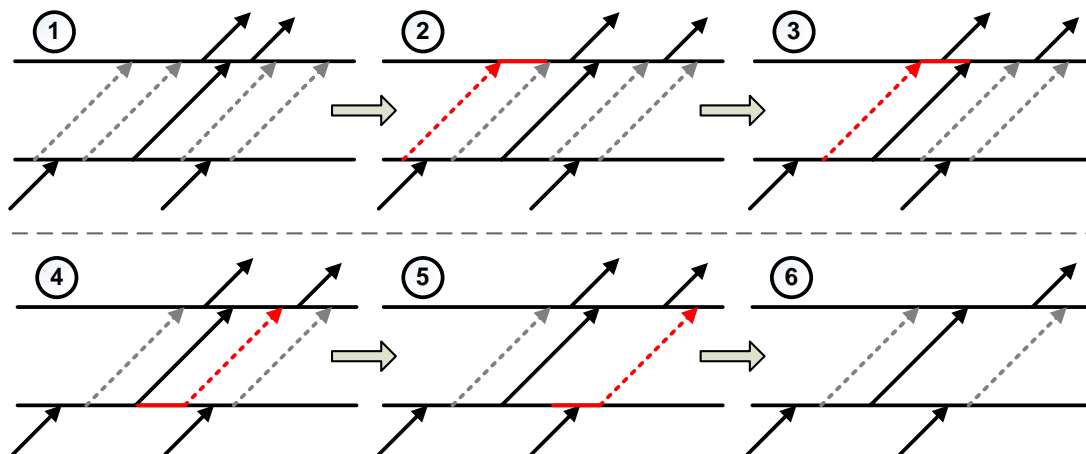


Abbildung 6.8: Netzwerk-Reduktion durch Löschen überflüssiger Zeitfenster-Kanten

lichkeiten durch die Zeitfenster-Kante ergeben. Somit wurde das Netzwerk um zwei von vier nicht notwendigen Zeitfenster-Kanten bereinigt. Dies führt dazu, dass auch das spätere Hinzufügen weiterer überflüssiger Kanten (z.B. Warte-Kanten) verhindert wird. Die weiteren Schritte des Netzwerkaufbaus (siehe Kapitel 4) bleiben unverändert.

Eine weitere mögliche Reduktion der Anzahl von Kanten im erweiterten Netzwerkmodell ist die Entfernung überflüssiger Depot-Kanten für verschobene Fahrten. Wie die linke Abbildung in 6.9 zeigt, werden für eine Servicefahrt-Kante im Normalfall zwei Depot-Kanten eingefügt. Die erste ermöglicht eine Ausrückfahrt eines Fahrzeugs aus dem Depot zu der Abfahrtshaltestelle hin und die zweite ermöglicht eine Einrückfahrt in das Depot.

Für alle Zeitfenster-Kanten müssten also jeweils auch zwei Depot-Kanten eingefügt werden. Wie der rechte Teil der Abbildung 6.9 zeigt, ist es allerdings ausreichend, wenn für die gesamte Gruppe von Kanten, die eine Fahrt darstellen (also alle Zeitfenster- sowie die Servicefahrt-Kante) insgesamt nur zwei Depot-Kanten eingefügt werden. Die Einrückfahrt führt vom Depot zu der frühesten dieser Kanten und die Ausrückfahrt – von der spätesten zum Depot. Dies ist ausreichend, da alle anderen Kanten über die Timelines bzw. über die Warte-Kanten erreicht werden können.

Abbildung 6.10 zeigt den Netzwerkaufbau für das Beispiel aus dem Kapitel 4 mit Zeitfenstern aus der Tabelle 6.3. Von den zwei Zeitfenster-Kanten, die zuvor in die zweite Schicht eingefügt wurden, war eine überflüssig, so dass sie aus dem Netzwerk entfernt werden konnte. Auch wurden wie oben beschrieben nur insgesamt zwei Depot-Kanten für die Kantengruppe der Fahrt f eingefügt. Die Zeitfenster-Kante ermöglicht eine Verknüpfung von den Servicefahrten f und d . Damit wird eine Flusslösung möglich, die lediglich zwei im Gegensatz zu zuvor drei Umläufen

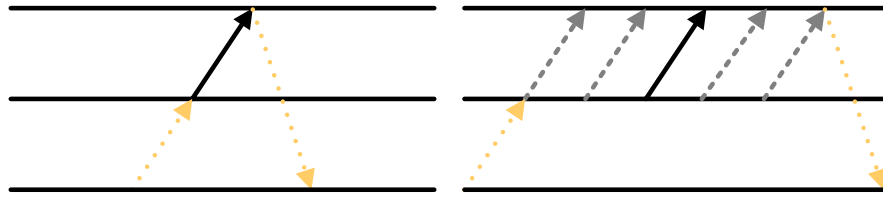


Abbildung 6.9: Reduktion von Depot-Kanten für Zeitfenster

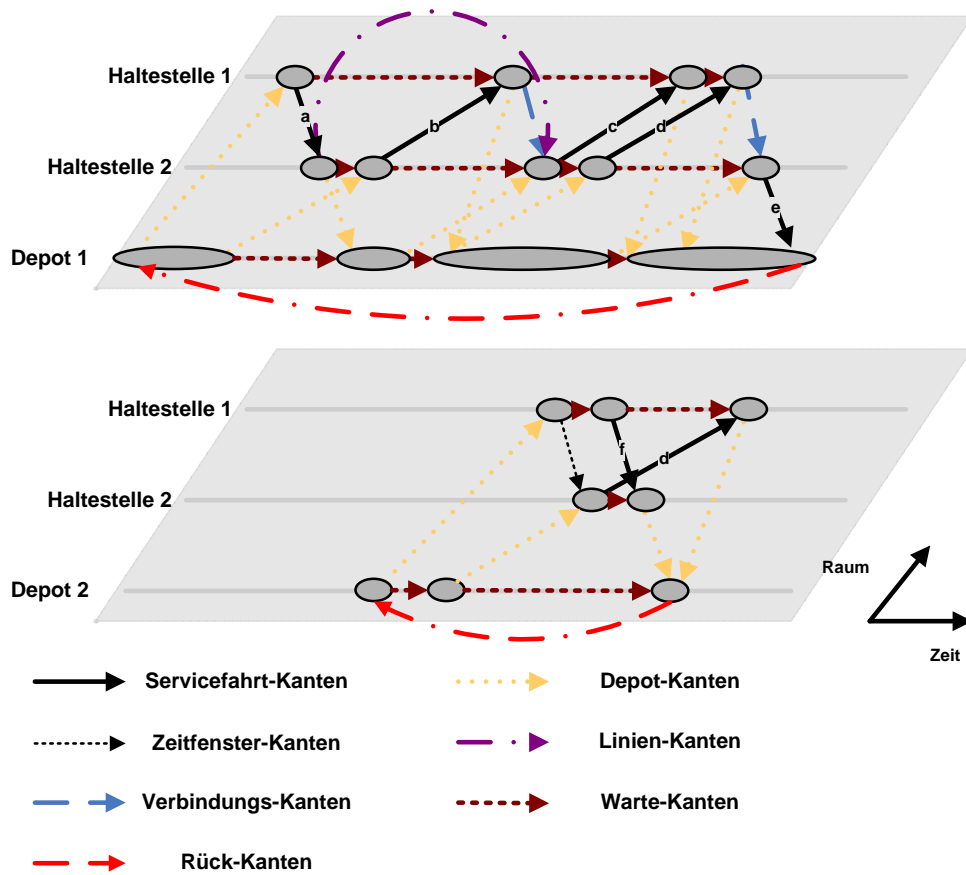


Abbildung 6.10: Netzwerk nach Reduktion überflüssiger Zeitfenster-Kanten

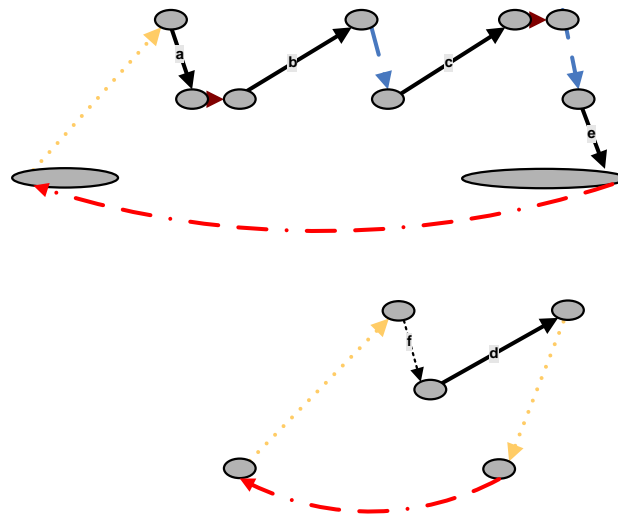


Abbildung 6.11: Optimaler Fluss im Netzwerk mit Zeitfenster-Kanten

enthält. Die Abbildung 6.11 zeigt die entsprechende Flusslösung. In diesem Beispiel hat eine Verschiebung einer Fahrt um eine Minute zur Einsparung eines Fahrzeugs geführt.

Interoperabilität mit den Linienkanten

Die im folgenden beschriebenen Anpassungen müssen für die gleichzeitige Berücksichtigung der Linien und der Verschiebeintervalle für Fahrten vorgenommen werden.

Eine Linien-Kante verbindet zwei Servicefahrt-Kanten, die zu der gleichen Linie gehören. Damit werden diese bei der Lösung des Modells bevorzugt (durch die negativen Kosten der Linien-Kante) von einem Fahrzeug hintereinander gefahren. Durch Restriktionen im mathematischen Modell wurde sichergestellt, dass die Linien-Kante nur benutzt wird, wenn beide Servicefahrt-Kanten benutzt werden.

Im Modell mit Zeitfenstern sind für jede Fahrplanfahrt nicht nur die Servicefahrt-Kanten, sondern eventuell noch die Zeitfenster-Kanten vorhanden. Beim Netzwerkaufbau muss die Linien-Kante nicht mehr die beiden originalen Servicefahrt-Kanten verbinden, sondern die letzte Kante, die der ersten Fahrplanfahrt entspricht mit der ersten Kante, die der zweiten Fahrplanfahrt entspricht. Alle anderen Kombinationen von Kanten aus betroffenen Gruppen können über eventuelles Warten auf der Timeline diese Linien-Kante nutzen. In Abbildung 6.12 ist dieser Fall verdeutlicht.

Zusätzlich zu dieser Regel müssen die Restriktionen für die Linien-Kanten angepasst werden. Statt sicherzustellen, dass die Linien-Kante nur benutzt werden darf,

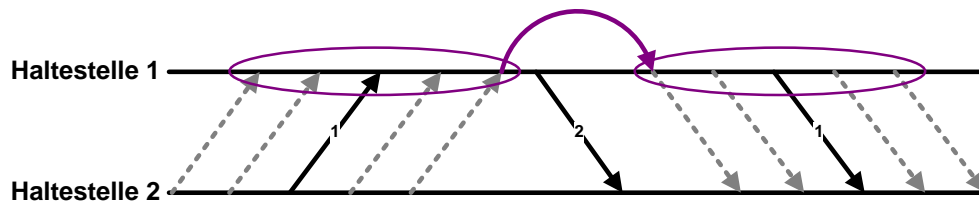


Abbildung 6.12: Linien-Kanten und Zeitfenster-Kanten: Normalfall

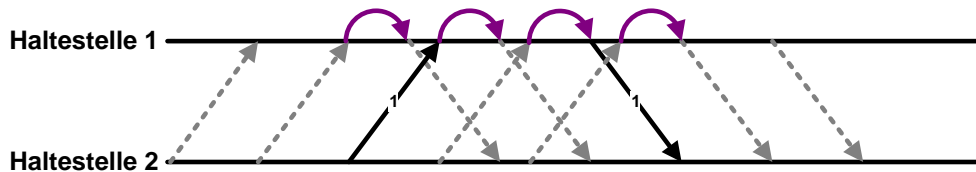


Abbildung 6.13: Linien-Kanten und Zeitfenster-Kanten: Spezialfall

wenn beide Kanten, die sie verbindet, gesetzt sind, müssen nun beide Gruppen von Zeitfenster- und Servicefahrt-Kanten Berücksichtigung finden: eine Linien-Kante darf nur dann benutzt werden, wenn eine Kante aus der Gruppen der Kanten erster Fahrt und eine aus der Gruppen der Kanten zweiter Fahrt benutzt werden. In der Abbildung sind die beiden Gruppen von Kanten durch ein Oval gekennzeichnet.

Neben dieser Grundform der Interaktion kann insbesondere bei größeren Verschiebeintervallen ein Sonderfall auftreten, bei dem sich Zeitfenster-Kanten der beiden zu verbindenden Fahrpläne zeitlich überschneiden. In diesem Fall muss zusätzlich für jede Kante, die in die Menge der vorgehenden Kanten hineinragt eine weitere Linien-Kante eingefügt werden, um auch diese Kombination der Verbindung zu ermöglichen. Die Restriktionen müssen dabei für jede Linien-Kante so angepasst werden, dass nur die Kanten der vorhergehenden Fahrplanfahrt, die vor ihr sind (bzw. nur die Kanten der nachfolgenden Fahrplanfahrt, die hinter ihr sind) in der Restriktion beachtet werden. Dies wird am Beispiel in der Abbildung 6.13 dargestellt. Die erste Linien-Kante ist beispielsweise notwendig, wenn die erste Zeitfenster-Kante der ersten Fahrt mit der ersten Zeitfenster-Kante der zweiten Fahrt benutzt werden. In diesem Fall wäre die Verbindung über keine andere Linien-Kante möglich. Die zahlreichen Tests haben gezeigt, dass dieser Spezialfall in der Praxis so gut wie nie auftritt.

6.2.2 Erweiterung des mathematischen Modells

Für die Integration des Zeitfenster-Konzeptes ist neben Änderungen im Netzwerkmodell eine entsprechende Erweiterung des mathematischen Modells aus dem Kapitel 4 notwendig. In diesem Abschnitt werden die notwendigen Änderungen bzw. Erweiterungen des mathematischen Modells beschrieben.

Für die Beschreibung der Zeitfenster-Kanten sowie ihrer Zugehörigkeit zu einer Servicefahrt sollen folgende Definitionen gelten:

EZ^s - die Menge aller Zeitfenster-Kanten in der Schicht $s \in S$,

EZ_f^s die Menge aller Zeitfenster-Kanten der Servicefahrt $f \in F$ in der Schicht $s \in S$ ($EZ_f^s \subseteq EZ^s$).

Die Zeitfenster-Kanten werden zu der Menge aller Kanten in einer Schicht $s \in S$ hinzugefügt: $E^s = ES^s \cup ED^s \cup EV^s \cup EL^s \cup EW^s \cup ER^s \cup EZ^s$.

Für die Bestimmung der Menge von Kanten, die eine bestimmte Servicefahrt repräsentieren, wird die Menge $erwESG_f$ eingeführt, die alle Servicefahrt- und Zeitfenster-Kanten der Fahrt $f \in F$ im Netzwerk beinhaltet. Gebildet wird diese Menge aus der bereits vorhandenen Menge aller Servicefahrt-Kanten einer Servicefahrt ESG_f und der Vereinigung aller zu der Fahrt zugehörigen Zeitfenster-Kanten über alle Schichten: $erwESG_f = ESG_f \cup \bigcup_{s \in S} EZ_f^s$

Für die Linien-Kanten Restriktionen wird die Menge $EPre_e^s$ definiert. Sie beinhaltet die vorausgehende Kante der Linien-Kante $e \in EL^s$ sowie alle Zeitfenster- und Servicefahrt-Kanten der Schicht $s \in S$, die zu der gleichen Servicefahrt gehören und eine **frühere** Ankunftszeit haben: $EPre_e^s = \bigcup_{n \in EZ_{f_{s1e}}^s | te_n < te_{s1e}} n \quad \forall e \in EL^s, \forall s \in S$

Zusätzlich wird die Menge $EPost_e^s$ definiert. Sie beinhaltet die vorausgehende Kante der Linien-Kante $e \in EL^s$ sowie alle Zeitfenster- und Servicefahrt-Kanten der Schicht $s \in S$, die zu der gleichen Servicefahrt gehören und eine **spätere** Abfahrtszeit haben:

$$EPost_e^s = \bigcup_{n \in EZ_{f_{s2e}}^s | ta_{s2e} < ta_n} n \quad \forall e \in EL^s, \forall s \in S$$

Da die Zeitfenster-Kanten im Netzwerk eine Alternative für die Startzeit repräsentieren, werden sie im mathematischen Modell wie Servicefahrt-Kanten behandelt. Deshalb ist der maximale Fluss auf Eins beschränkt. Die Kosten werden mit einem einstellbaren Parameter $WindowCost$ skaliert, um die unnötigen Fahrtenverschiebungen zu vermeiden. Die Zeilen (6.4) und (6.5) aus dem ursprünglichen Modell müssen deshalb entsprechend angepasst werden:

$$u_e = 1 \forall e \in (ES^s \cup ED^{s1} \cup EL^s \cup EZ^s), \quad \forall s \in S \quad (6.4)$$

$$c_e = WindowCost((te_e - ta_e) \cdot c_{std}^{ts}) \quad \forall e \in (ES^s \cup EW^s \cup EZ^s), \forall s \in S \quad (6.5)$$

Die Zielfunktion des mathematischen Modells kann unverändert bleiben, allerdings müssen einige Restriktionen für die Betrachtung von Zeitfenstern angepasst

werden:

$$\sum_{e \in \text{erwESG}_f} x_e = 1 \quad \forall f \in F \quad (6.6)$$

$$x_e \leq x_{s1_e} + EPre_e^s \quad \forall e \in EL^s, \forall s \in S \quad (6.7)$$

$$x_e \leq x_{s2_e} + EPost_e^s \quad \forall e \in EL^s, \forall s \in S \quad (6.8)$$

Bei den cover constraints (6.6) muss nun sichergestellt werden, dass die Summe aller Flüsse der zu einer Servicefahrt gehörigen Zeitfenster- und Servicefahrt-Kanten genau gleich eins ist. Somit wird sichergestellt, dass bei der Lösung dieses Modells zwischen den Servicefahrt- und Zeitfenster-Kanten alternativ gewählt werden kann und somit die Verschiebung von Fahrten bei kostengünstigeren Konstellationen zustande kommt. Die Restriktionen (6.7) und (6.8) stellen sicher, dass eine Linien-Kante nur dann benutzt werden darf, wenn entsprechend die Servicefahrt- bzw. Zeitfensterkanten benutzt wurden.

6.2.3 Verschiebbare Fahrten

Die Menge der verschiebbaren Fahrten kann auf unterschiedliche Weise definiert werden. Die erste Möglichkeit, Zeitfenster zu definieren, besteht darin, Angaben für eine **globale Zeitfensterverschiebung** zu machen. Diese Angabe gilt für sämtliche Servicefahrten des Fahrplans und wird sowohl auf die in die Verschiebemöglichkeit nach vorne wie auch nach hinten angegeben. Eine Angabe über die globale Verschiebemöglichkeit für beide Richtungen kann durch einen Parameter angegeben werden. Da - wie zu Beginn dieses Kapitels erläutert wurde - nur ganze Minuten bei der Zeitfensterintegration berücksichtigt werden, könnte dieser Parameter als Anzahl von Minuten definiert werden. Diese Art der Definition der verschiebbaren Fahrtenmenge hat einen entscheidenden Nachteil - die Problemgröße explodiert mit wachsender Länge des Verschiebeintervalls. Um dies zu vermeiden werden nachfolgend einige Möglichkeiten vorgestellt, die Menge der verschiebbaren Fahrten zu begrenzen, um die Größe des zu lösenden Problems beherrschbar zu halten.

Benutzerdefinierte Fahrten

Eine benutzerdefinierte Definition der Menge verschiebbarer Fahrten erfolgt durch eine explizite Angabe der Zeitfenster - sowohl für eine Verschiebung nach vorne als auch nach hinten. Dadurch kann der Benutzer Angaben über die Verschiebeintervalle für jede Fahrten machen und hat die Möglichkeit, nur eine Teilmenge der Fahrplanfahrten für die Verschiebung freizugeben. Im Normalfall wird die Modellgröße durch diesen Umstand gegenüber der globalen Verschiebung verkleinert, da nur Kanten für die explizit ausgezeichneten Fahrten in das Netzwerk eingefügt werden

müssen. Deshalb ist die Laufzeit für die meisten praxisrelevanten Instanzen mit benutzerdefinierten Zeitfensterangaben nicht viel größer als die ohne Zeitfenster. Dies ist natürlich abhängig von der Menge und Höhe der angegebenen Zeitfensterwerte.

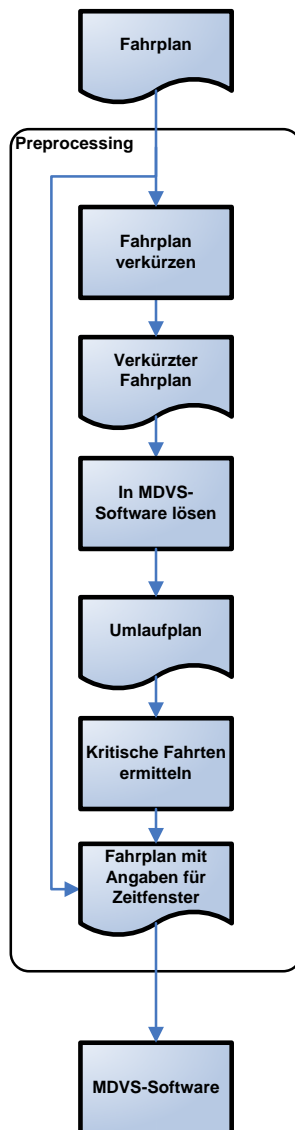


Abbildung 6.14: Ablauf der Verkürzungsheuristik

Allerdings besteht in vielen Fällen das Problem, dass die Menge der zur Verschiebung freizugebenden Fahrten nicht festgelegt werden kann. Daher wäre es wünschenswert, diese Menge von verschiebbaren Fahrten, die das Potential für die Einsparung von Umläufen beinhalten, automatisch ermitteln zu können. Diese Menge wird nachfolgend auch die Menge der kritischen Fahrten bezeichnet. Um sie zu ermitteln, werden in dieser Arbeit verschiedene heuristische Methoden entwickelt. Die nachfolgend vorgestellten heuristischen Methoden dienen der Bestimmung einer kritischen Fahrtenmenge und deren Verschiebungsintervalle. Im Gegensatz zu globalen Angaben könnten auf diese Weise bei erheblicher Modellverkleinerung (und somit kürzerer Laufzeit) gleiche Ergebnisse in Bezug auf die Einsparung von Umläufen erzielt werden. Zwei unterschiedliche heuristische Verfahren werden in folgenden Abschnitten vorgestellt.

Verkürzungsheuristik zur Ermittlung verschiebbarer Fahrten

Die Optimierung mit Verschiebeintervallen erfolgt bei der Benutzung der Verkürzungsheuristik zur Ermittlung verschiebbarer Fahrten in mehreren Schritten, die in der Abbildung 6.14 dargestellt sind. Als Ausgangsbasis dient der zu optimierende Fahrplan. Im ersten Schritt der Optimierung mit Verschiebeintervallen werden alle Servicefahrten des Fahrplans FP verkürzt und es entsteht ein Fahrplan FP' . Dies geschieht dadurch, dass die Abfahrt aller Servicefahrten um eine durch den Benutzer anzugebende Zeit verspätet wird. Die Ankunftszeiten aller Servicefahrten bleiben dagegen unverändert. Für diesen verkürzten Fahrplan FP'

wird im nächsten Schritt der optimale Umlaufplan UP' berechnet, der aus den verkürzten Fahrten besteht. Der Umlaufplan UP' ist aufgrund der verkürzten Fahrten nicht für den praktischen Gebrauch bestimmt, sondern stellt vielmehr ein Hilfskonstrukt dar, das dazu dient herauszufinden, bei welchen Fahrten sich Verschiebungen

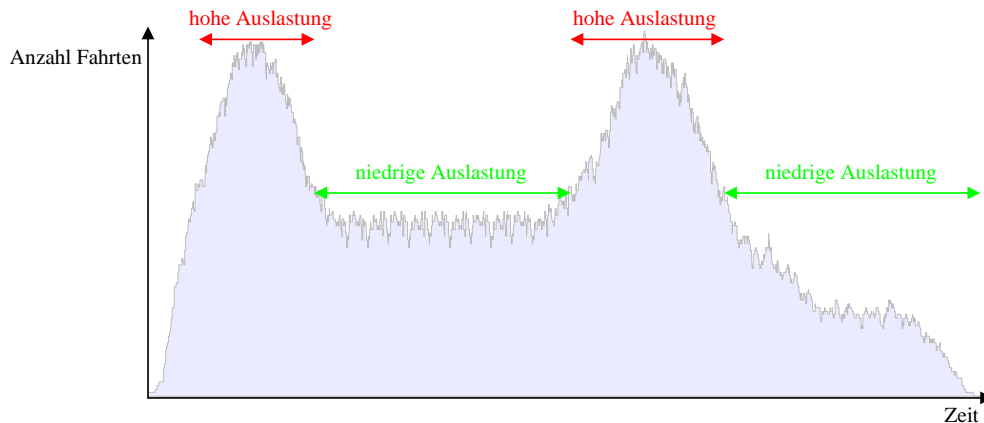


Abbildung 6.15: Fahrtenauslastung eines Fahrplans

überhaupt lohnen könnten. UP' dient als Ausgangslösung zur Bestimmung von kritischen Fahrten.

Um kritische Fahrten zu ermitteln, wird bei jeder Fahrtenverknüpfung überprüft, ob sie ohne Verkürzung ausführbar wäre. Falls nicht - tritt ein Konflikt auf und die beiden Fahrten, die diesen Konflikt auslösen sollen verschiebbar gemacht werden.

Für alle auf diese Weise ermittelten kritischen Fahrten werden die Verschiebeintervalle entsprechend der zeitlichen Überschneidung des Konflikts gewählt. Somit wird die Ausnutzung der gefundenen Potentiale in der Lösung des Fahrplans mit Zeitfenstern gewährleistet. Mit der so bestimmten kritischen Fahrtenmenge wird ein neuer Fahrplan erstellt, der die entsprechenden Zeitfensterangaben beinhaltet und erneut mit den expliziten Zeitfenstern gelöst werden soll.

Schnitt-Heuristik

Die Schnitt-Heuristik zur Bestimmung verschiebbarer Fahrten basiert auf der Beobachtung, dass die Fahrpläne in der Regel deutliche Auslastungsspitzen aufweisen. Das Ziel dieses Ansatzes ist es, solche Spitzen „abzutragen“ und die Auslastung auf umliegende Gebiete zu verteilen. Grafisch äußert es sich in einer Glättung der Auslastungskurve. Im Gegensatz zu der oben beschriebenen Heuristik ist dafür kein vorgelagerter Optimierungsvorgang notwendig. Die Informationen über die kritische Menge von Fahrten werden ausschließlich aus dem gegebenen Fahrplan bestimmt, was keine zusätzliche Laufzeit verursacht.

In der Abbildung 6.15 ist die Auslastung eines typischen Fahrplans grafisch dargestellt. Die horizontale Achse beschreibt die Zeit und die vertikale stellt die

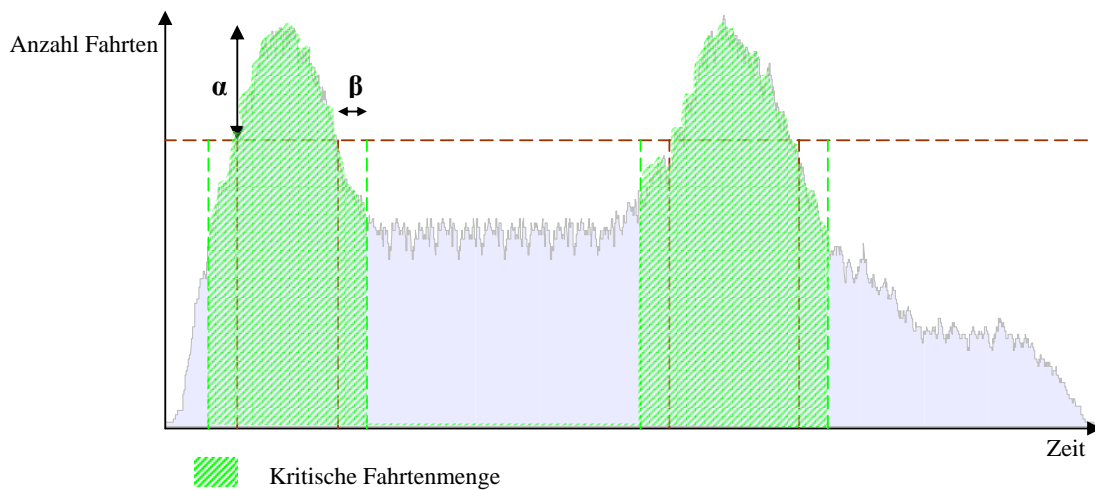


Abbildung 6.16: Ermittlung kritischer Fahrten in der Schnitt-Heuristik

Anzahl der zu einem Zeitpunkt gleichzeitig stattfindenden Servicefahrten dar. Das Maximum dieser Linie stellt eine untere Grenze für die Anzahl der benötigten Fahrzeuge dar. Die Fahrten, die sich in den Spitzen befinden, sollen verschiebbar gemacht werden, so dass die Auslastungskurve geglättet werden kann. Für diese Art der Bestimmung kritischer Fahrten wird ein horizontaler Schnitt in der Fahrtenmenge vorgenommen, der um einen Wert α von der Spitze entfernt ist. Die kritische Menge von Fahrten befindet sich in der Säule, die sich aus der Projektion des Schnittes auf die Zeitachse ergibt.

Zusätzlich zu diesem Schnitt kann eine β -Umgebung angegeben werden. Diese erweitert die Menge der kritischen Fahrten über die Grenzpunkte des durch α definierten Schnittes hinaus. Beispielhaft ist eine Ermittlung der Schnittmenge in der Abbildung 6.16 aufgezeigt. Die Höhe des Schnittes wird vom Parameter α festgelegt. Zusätzlich kommen Fahrten aus der durch den Parameter β festgelegten Umgebung hinzu. Der kritische Bereich ist durch die schraffierte Fläche kenntlich gemacht. Verschiebungen von kritischen Fahrten sollen die Spitzen der Kurve auf die Umgebung verteilen, so dass die Kurve wie in der Abbildung 6.17 geglättet wird und niedrigere Maxima aufweist.

6.2.4 Testergebnisse

Die Einbeziehung der zeitlichen Verschiebeintervalle für Fahrplanfahrten in das Time-Space-Netzwerkmodell wurde auf Probleminstanzen aus der Realität getestet. Insbesondere wurden die angesprochenen Möglichkeiten der Ermittlung von verschiebbaren Fahrten getestet und im Bezug auf die Laufzeiten und auf die Lösungsqualität gegenübergestellt. Nachfolgend werden die Testergebnisse dargestellt

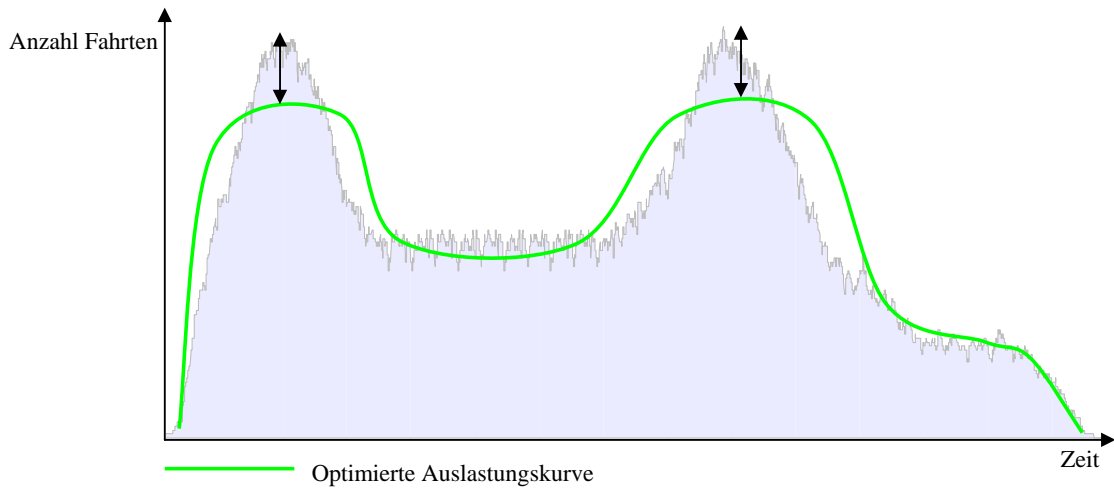


Abbildung 6.17: Glättung der Auslastungskurve durch die Schnitt-Heuristik

Instanz	#Servicefahrten	#Haltestellen	⊙ Gruppengröße
Monheim	424	34	1
Saarbrücken	1296	89	2
Halle	2047	21	3,2
München-28	1808	76	13

Tabelle 6.4: Daten der verwendeten Instanzen

und besprochen.

Zur Lösung der gemischt-ganzzahligen Probleme wurde ILOG CPLEX 9.0 (vgl. [ILOG, 2004]) verwendet. Die LP-Relaxationen wurden für die ersten drei Instanzen mit dem dualen Simplex-Algorithmus gelöst. Für die Instanz *München* wurde der Interior-Point-Algorithmus verwendet. Eine ganzzahlige Lösung wurde als optimal angenommen, falls ihr Kostenwert höchstens um ein Prozent vom Optimum der LP-Relaxation entfernt ist.

Die Tests wurden an verschiedenen Instanzen aus der Praxis durchgeführt, deren Charakteristika der Tabelle 6.4 entnommen werden können. Die Fahrtenmengen der vier untersuchten Fahrpläne über einen Betriebstag werden in der Abbildung 6.18 visualisiert. Die Diagramme zeigen in Abhängigkeit von der Zeit die Anzahl der Fahrten, die zu einem bestimmten Zeitpunkt laut jeweiligem Fahrplan gefahren werden müssen.

Die Instanz *Monheim* stellt die kleinste Instanz im Testfeld dar. Da sie nur ein Depot und einen Fahrzeugtyp enthält und relativ wenig Servicefahrten beinhaltet, ist sie in sehr kurzer Zeit lösbar. Der Graph der Fahrtenmenge in der Abbildung

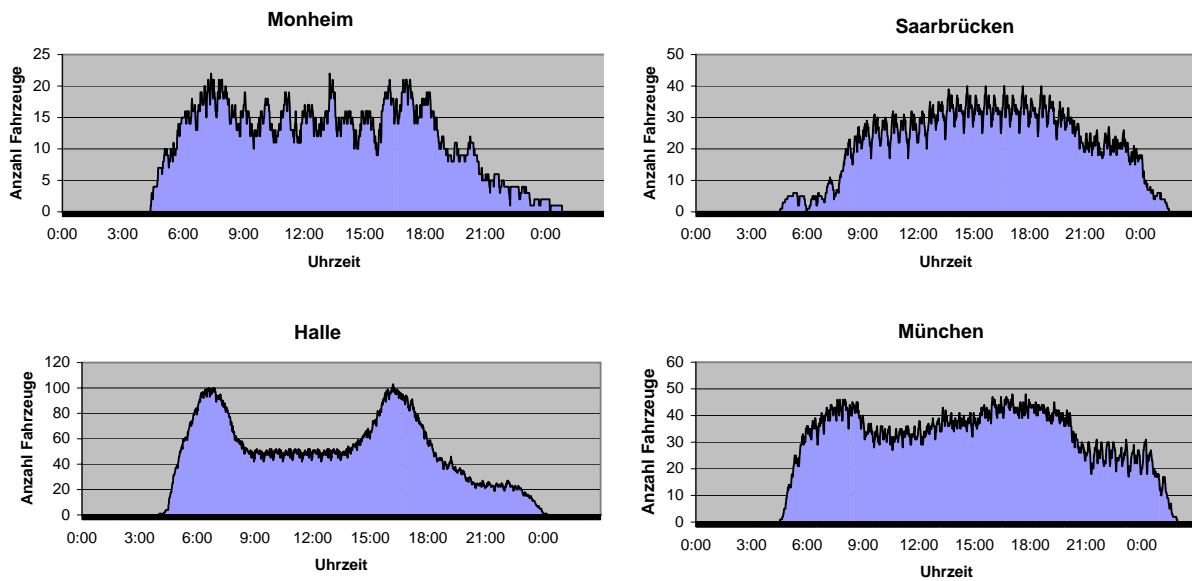


Abbildung 6.18: Fahrtenmengen der Instanzen über die Dauer eines Betriebstages

6.18 ist durch starke Fluktuationen geprägt, was auf die relativ geringe Anzahl an Servicefahrten im Fahrplan zurückzuführen ist. Eine leicht erhöhte Auslastung ist in den Zeiten des Berufsverkehrs gegen 8:00 Uhr und 17:00 Uhr zu erkennen. Die Instanz *Saarbrücken* stellt mit zwei Depots, drei Fahrzeugtypen und 1296 Servicefahrten eine Instanz mittlerer Größe dar. Die Darstellung der Fahrtenmenge ist relativ glockenförmig. Die Instanz *Halle* hat bezüglich der Komplexität ähnliche Eigenschaften wie die Instanz *Saarbrücken*. Mit einer gleichen Anzahl von Depots und Fahrzeugtypen sowie einer nur geringfügig größeren Menge an Servicefahrten stellt sie auch eine Instanz mittlerer Größe dar. Die entsprechende Fahrtenmengenkurve in der Abbildung 6.18 hat die markante Eigenschaft, dass sie nicht wie bei *Saarbrücken* eine Glockenform aufweist, sondern in den Morgenstunden und am Nachmittag deutliche Spitzen hat. Die Teilinstanz aus München *München-28* hat eine geringfügig kleinere Fahrtenmenge als die zwei vorhergehenden Instanzen, zählt aber mit ihren 19 Depots zu den komplexeren Probleminstanzen.

Um eine Vergleichsbasis für die Bewertung der Zeitfenster-Modalitäten zu bekommen, sind alle Instanzen zunächst ohne Zeitfenster gelöst worden. Die Ergebnisse dieser Tests liefern eine Grundlage für die Bewertung der Verschiebeintervalle-Varianten hinsichtlich der Fahrzeuganzahl und der Laufzeit. Weitere Tests wurden für die oben beschriebenen drei Varianten der Menge von verschiebbaren Fahrten durchgeführt.

Für die globale Zeitfenster und für die Verkürzungs-Heuristik wurden unter-

schiedlich große Verschiebeintervalle getestet. Bei der dritten Variante - der Schnitt-Heuristik - müssen neben der Größe des Verschiebeintervalls die beiden Parameter α und β angegeben werden. Daher waren hier umfangreichere Tests notwendig. Um einen Einblick in das Verhalten bei verschiedenen Parametereinstellungen zu bekommen, wurden für die ersten drei beschriebenen Instanzen alle Kombinationen von α und β mit Werten von Null bis Zehn durchgeführt worden. Diese Kombinationen wurden ebenfalls für verschiedene Intervallgrößen getestet.

Die Testläufe für die vier Instanzen sind die in den Tabellen 6.5, 6.6 und 6.7 protokolliert. Die erste Zeile in jedem Segment dieser Tabellen gibt die Anzahl Umläufe und die Laufzeit für normale Optimierung ohne Verschiebeintervalle. Die restlichen Zeilen beinhalten die Angaben zu einer Optimierung mit Verschiebeintervallen verschiedener Größe. Neben den Beurteilungskriterien Umlaufanzahl und Laufzeit sind zusätzlich die Anzahl der in der Lösung verschobenen Fahrten sowie die Gesamtverschiebung über alle Fahrten in Minuten angegeben.

Soweit nicht anders angegeben, wurden die Kosten für eine verschobene Fahrt auf das Dreifache der normalen Kosten gesetzt. Dadurch soll sichergestellt werden, dass die Fahrten nur dann verschoben werden, wenn dadurch die Anzahl der Fahrzeuge reduziert wird. Als Messgröße zur Bestimmung der Qualität einer Lösung wird die Anzahl der Umläufe (bzw. Fahrzeuge) herangezogen. Für alle Tests wurde ein Zeitlimit von fünf Stunden gesetzt.

Die Tests mit globalen Zeitfenstern (s. Tabelle 6.5) konnten für die meisten Instanzen nur für relativ kleine Zeitfenster durchgeführt werden, da mit größeren Angaben die Grenze von fünf Stunden Laufzeit überschritten wurde. Eine Ausnahme war die Instanz *Monheim*, bei der aufgrund ihrer geringen Komplexität auch Lösungen für große Zeitfenster bis hin zu einer Stunde berechnet wurden. Bei einem Verschiebeintervall von 70 Minuten für alle Fahrten wurde die Optimierung ebenfalls nach 5 Stunden abgebrochen.

Die Ergebnisse für globale Verschiebeintervalle (mit einem Γ -Wert zwischen einer und fünf Minuten) sind in der oberen Grafik von Bild 6.19 dargestellt. Hier kann man gut erkennen, dass in den berechneten Instanzen bei einer Verschiebemöglichkeit von fünf Minuten eine Fahrzeugeinsparung zwischen 4% und 12% möglich ist. Dies ist eine beträchtliche Einsparung, wenn man bedenkt, dass jeder Bus im ÖPNV sehr hohe Fixkosten mit sich bringt. Bei der Instanz *Halle*, deren Fahrtenmenge eine auffällige Struktur hat (siehe Abbildung 6.18), ist bereits bei einem Zeitfenster von nur einer Minute eine Fahrzeugeinsparung von 4% möglich. Dagegen sind bei den anderen Instanzen für diesen Grad der Einsparung größere Verschiebemöglichkeiten notwendig. Die untere Grafik der Abbildung 6.19 zeigt, dass die Berechnung dieser Lösung einen enormen Anstieg der Laufzeit verursacht. In der Grafik ist der Laufzeitanstieg in Abhängigkeit von der Größe des berücksich-

Instanz	Zeitfenster Γ (in Min.)	#Umläufe	#verschobene Fahrten	Verschiebung (in Min.)	Laufzeit (in Min.)
Monheim	ohne	49	-	-	0,01
	± 1	49	1	1	0,00
	± 2	49	1	1	0,02
	± 3	48	3	7	0,02
	± 4	47	15	50	0,02
	± 5	47	16	62	0,02
	± 10	46	23	179	0,05
	± 20	43	48	757	0,25
	± 30	39	101	2293	1,08
	± 40	38	116	3555	2,95
	± 50	36	181	6576	17,62
	± 60	35	198	8546	265,58
	± 70	-	-	-	> 300
Saarbrücken	ohne	49	-	-	0,04
	± 1	49	0	0	0,20
	± 2	48	9	16	0,78
	± 3	47	35	87	1,63
	± 4	45	98	328	5,65
	± 5	43	210	837	32,97
	± 6	44	196	927	76,15
	± 7	-	-	-	> 300
Halle	ohne	115	-	-	1,35
	± 1	110	50	50	25,31
	± 2	107	70	133	62,13
	± 3	106	93	245	126,75
	± 4	105	106	316	237,56
	± 5	104	122	440	280,71
	± 6	-	-	-	> 300
München-28	ohne	56	-	-	35,91
	± 1	-	-	-	> 300

Tabelle 6.5: Ergebnisse bei globalen Zeitfenstern

Instanz	Zeitfenster (in Min.)	#Umläufe	#verschobene Fahrten	Verschiebung (in Min.)	Laufzeit (in Min.)
Monheim	ohne	49	-	-	0,01
	±10	47	10	80	0,03
	±20	47	10	174	0,05
	±30	43	37	2762	0,33
	±40	39	54	16444	7,12
	±50	40	71	8192	6,92
	±60	40	46	9265	1,17
	±70	35	93	18422	9,2
Saarbrücken	ohne	49	-	-	0,04
	±1	49	0	0	2,95
	±2	49	0	0	2,83
	±3	49	0	0	2,95
	±4	48	5	23	3,05
	±5	48	8	37	3,13
	±6	47	26	148	3,18
	±7	48	12	99	3,33
	±8	45	72	576	3,83
Halle	ohne	115	-	-	1,35
	±1	112	18	18	2,93
	±2	109	37	257	12,52
	±3	108	51	173	9,48
	±4	107	71	324	18,60
	±5	106	86	394	16,38
	±6	104	106	1203	77,82
München-28	ohne	56	-	-	35,91
	±1	-	-	-	> 300

Tabelle 6.6: Ergebnisse der Verkürzungs-Heuristik

Instanz	Zeitfenster (in Min.)	#Umläufe	#verschobene Fahrten	Verschiebung (in Min.)	Laufzeit (in Min.)
Monheim	ohne	49	-	-	0,01
	±10	46	21	163	0,05
	±20	43	52	781	0,27
	±30	41	58	1355	0,62
	±40	39	87	2611	1,27
	±50	38	71	2871	2,22
	±60	37	76	3590	4,1
	±70	38	76	4209	7,08
Saarbrücken	ohne	49	-	-	0,04
	±1	49	0	0	0,17
	±2	48	8	14	0,5
	±3	47	33	81	1,18
	±4	45	95	307	7,18
	±5	43	215	837	19,67
	±6	42	242	1106	48,73
	±7	-	-	-	>300
Halle	ohne	115	-	-	1,35
	±1	111	26	26	3,97
	±2	107	69	126	7,23
	±3	106	93	242	10,93
	±4	105	107	330	17,45
	±5	104	108	383	20,75
	±6	104	112	452	31,55
	±7	103	132	660	43,6
	±8	-	-	-	>300
München-28	ohne	56	-	-	35,91
	±1	55	8	8	127,08
	±2	-	-	-	>300

Tabelle 6.7: Ergebnisse der Schnitt-Heuristik mit $\alpha = 10$ und $\beta = 10$

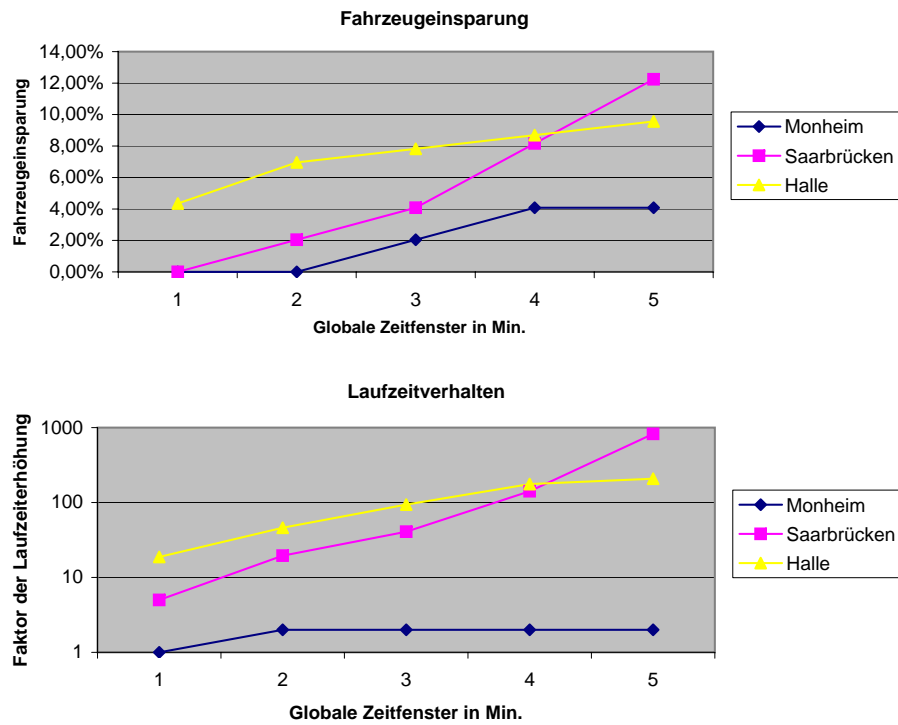


Abbildung 6.19: Einsparung und Laufzeit für globale Zeitfenster bis zu fünf Minuten

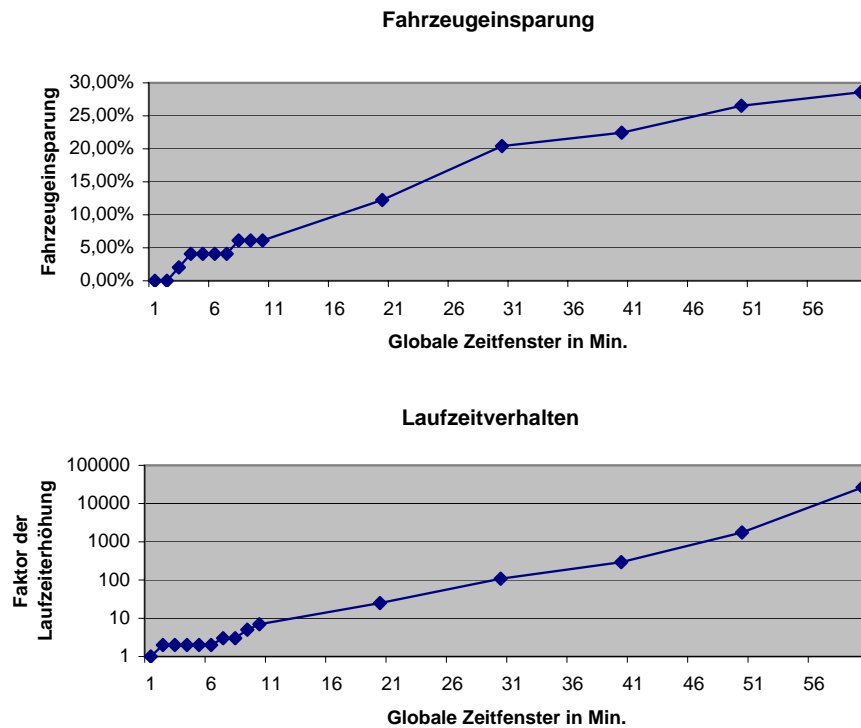


Abbildung 6.20: Einsparung und Laufzeit für große Zeitfenster für Monheim

tigten Zeitfenster als ein *als Faktor im Vergleich zur Laufzeit des Problems ohne Zeitfenster* angegeben. Die Ergebnisse sind in der unteren Grafik auf einer logarithmischen Skala aufgezeichnet, die obere Grafik verwendet hingegen eine lineare Skalierung. Wie man gut erkennen kann, verhalten sich die Kurven einer Instanz in beiden Grafiken sehr ähnlich. Folglich wächst die Laufzeit in dem untersuchten Bereich exponentiell zur Fahrzeugeinsparung.

Auffällig ist die Laufzeitkurve der Instanz *Monheim* mit einem deutlich anderen Verlauf. Eine Darstellung des Verlaufs über einen größeren Bereich in der Abbildung 6.20 zeigt allerdings, dass auch bei dieser Instanz eine lineare Steigerung der Einsparung einem exponentiellen Anstieg der Laufzeit entspricht.

Die beiden heuristischen Ansätze zur Ermittlung der kritischen Fahrtenmenge sollten schon bei kürzeren Laufzeiten gute Ergebnisse berechnen können. Insbesondere sollte dadurch die Instanz *München-28* mit Verschiebeintervallen lösbar werden. Die Ergebnisse der Verkürzungs-Heuristik sind in der Tabelle 6.6 dargestellt. Die Laufzeiten der Verkürzungs-Heuristik sind für die gleichen Verschiebeintervalle deutlich niedriger, als bei der Verwendung globaler Zeitfenster. Dadurch konnten innerhalb des Zeitlimits von fünf Stunden auch Probleme mit größeren Verschiebeintervallen gelöst werden. Bei näherer Betrachtung fällt allerdings auf, dass die Ergebnisqualität zum Teil unter der Lösung der globalen Zeitfenster liegt.

Bei der Benutzung der Schnitt-Heuristik müssen neben der Größe der Zeitfenster auch Werte für α und β gewählt werden. Der Parameter α hat nach Analyse zahlreicher Testläufe den meisten Einfluss auf die Qualität der Lösung. Allerdings führt die Erhöhung des Wertes β in Kombination mit α zu einer Lösung, die mit alleinigem Setzen von α erst für größere Zeitfenster möglich ist. Deshalb scheint eine Kombination der beiden Parameter sinnvoll zu sein. α ist dabei der ausschlaggebende Parameter und β hat eine unterstützende Wirkung. Die nachfolgend vorgestellten Tests sind mit $\alpha = 10$ und $\beta = 10$ durchgeführt worden. Diese Einstellung stellte sich im Laufe der Tests als ein guter Kompromiss zwischen Laufzeit und Lösungsqualität heraus. Die Ergebnisse sind in Tabelle 6.7 aufgeführt.

Für fast alle Instanzen wurde eine Fahrzeugeinsparung erreicht, die der Variante mit globalen Zeitfenstern entspricht. Im Vergleich zu der Verkürzungs-Heuristik sind die Laufzeiten zum Teil erheblich kürzer. Zusätzlich konnten innerhalb des Zeitlimits von fünf Stunden Lösungen berechnet werden, die mit globalen Zeitfenstern nicht erreicht wurden. So wurde für die Instanz *Halle* eine Lösung mit 103 Umläufen in unter 27 Minuten berechnet. Die Schnitt-Heuristik hat auch eine Verbesserung für die Instanz *München-28* gefunden. Somit wurde das MDVSP mit Zeitfenstern mit Hilfe dieser Heuristik auch für komplexere Instanzen lösbar.

Die Abbildungen 6.21 und 6.22 enthalten einen Vergleich der drei Varianten für die Instanz *Halle*. Neben globalen Zeitfenstern, der Verkürzungs-Heuristik und der

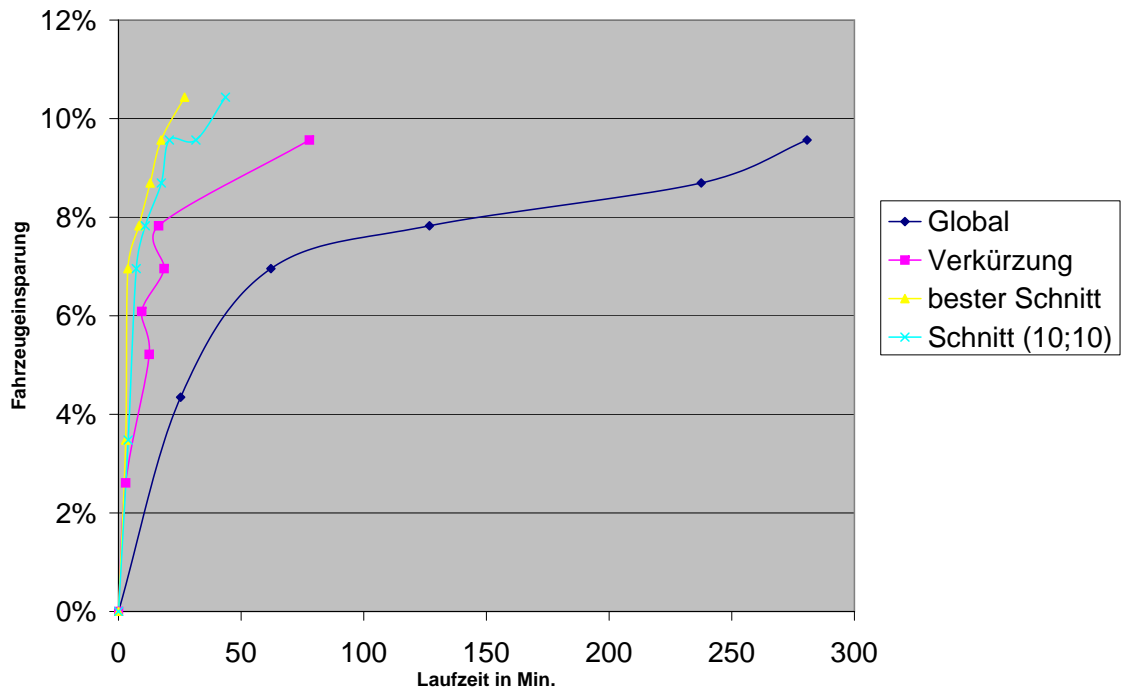


Abbildung 6.21: Vergleich von Lösungsqualität zu Laufzeit für Halle

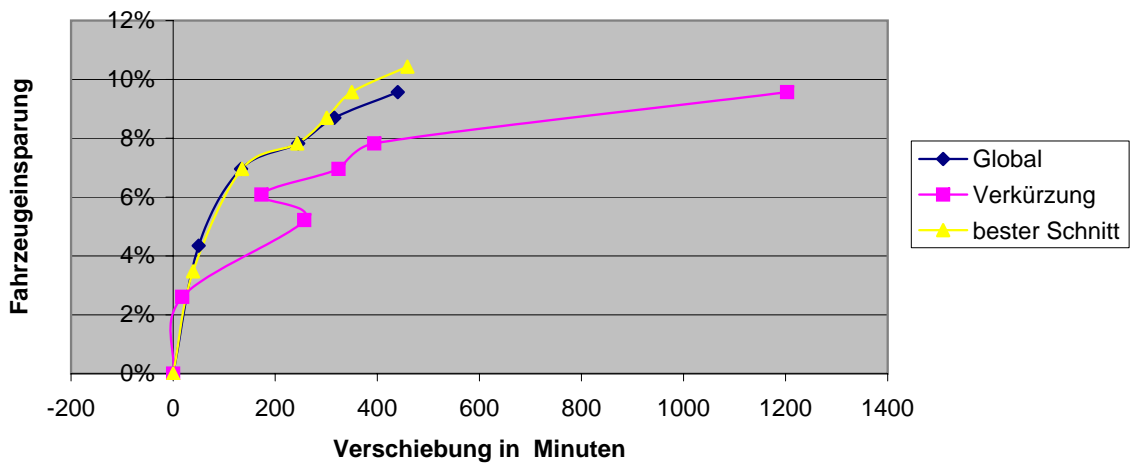


Abbildung 6.22: Vergleich von Lösungsqualität zur Gesamtverschiebung für Halle

Schnitt-Heuristik mit den Parametern $\alpha = 10$ und $\beta = 10$ ist eine weitere Kurve eingezeichnet, die die Ergebnisse einer Schnitt-Heuristik mit optimal gewählten Parametern darstellt. Sie wurde aus der Sensibilitätsanalyse für die Schnitt-Heuristik gewonnen und stellt keine realistische Modalität dar, da die optimalen Parameter für eine Instanz nicht ohne vorheriges Testen aller Kombinationen bestimmt werden können. Die Kurve gibt also lediglich das theoretische Potential der Schnitt-Heuristik an und kann somit als eine obere Schranke angesehen werden.

Die Lösung mit globalen Zeitfenstern benötigt viel Zeit bereits für relativ klein gewählte Zeitfenster (die Kurve ganz rechts in der Grafik). Bessere Performance weist die Verkürzungs-Heuristik auf. Sie findet in einer erheblich kürzeren Zeit die gleichen Einsparungsergebnisse (vgl. die höchsten Punkte der beiden Kurven). Sie braucht allerdings für diese Lösungen größere Verschiebeintervalle - die Datenpunkte in dem Graphen stehen dichter zusammen.

Die Schnitt-Heuristik liefert die beste Performance - die beiden Kurven liegen ganz links in der Grafik. Es werden auch die höchsten Einsparungen erreicht - vgl. die höchsten Punkte der Kurven. Zusätzlich sind nahezu die gleichen Einsparungspotentiale mit dem gleichen Verschiebeintervall wie bei der Anwendung von globalen Zeitfenstern zu erreichen - die entsprechenden Punkte der beiden Graphen liegen fast auf der gleichen Höhe. Wie Abbildung 6.21 zeigt, ist die Einstellung von $\alpha = 10$ und $\beta = 10$ für die Schnitt-Heuristik in dieser Instanz bereits sehr dicht an der aus den besten Ergebnissen zusammengesetzten Schranke. Weitere Tests haben gezeigt, dass die größeren Verschiebeintervalle auch tendenziell größere α und β Parameter für die Schnitt-Heuristik benötigen, um die gleiche Fahrzeugeinsparung wie bei den entsprechenden globalen Zeitfenstern zu erreichen.

Die Verkürzungs-Heuristik liefert gleich gute Einsparungen, wie die globalen Zeitfenster, braucht aber dafür größere Verschiebeintervalle. Daher kommt es bei dieser Methode zu einem für den Einsatz in der Praxis relevanten Nachteil: die Abbildung 6.22 zeigt, dass für gleiche Einsparung von Fahrzeugen erheblich größere Verschiebungen im Fahrplan vorgenommen werden müssen.

Für einen ÖPNV-Betrieb, dessen verschiebbare Fahrten klar definiert sind, bietet sich der Einsatz der benutzerdefinierten Zeitfenster an. Die Heuristiken sind besonders interessant für die Benutzer, die das Einsparungspotential bei einer begrenzten Laufzeit möglichst voll ausschöpfen wollen. Die deutlich besseren Ergebnisse liefert im momentanen Entwicklungsstand die Schnitt-Heuristik, die die oben formulierte Anforderung voll erfüllt.

6.2.5 Überlegungen zu der Anschlussfahrten-Problematik

Bei der Betrachtung der Zeitfenster kann es passieren, dass die ursprünglichen Planungen mit Fahrzeuganschlüssen nicht eingehalten werden können. Zur Erläuterung folgendes Beispiel: Es existiert eine Servicefahrt 1 von A nach B und eine Servicefahrt 2 von B nach C , die als Anschlussfahrt fungiert. Bei der Angabe von Zeitfenstern kann es nun passieren, dass nach der Optimierung die Fahrt 2 beginnt noch bevor die Fahrt 1 beendet wurde. Somit ist die Anschlussmöglichkeit nicht mehr vorhanden. Deshalb könnte man durch eine Schnittstellenerweiterung festlegen, welche Fahrt eine Anschlussfahrt einer anderen Fahrt ist und durch Restriktionen diese Bedingung innerhalb der Optimierung sicherstellen. Denkbar wäre folgende Erweiterung in der Datenschnittstelle, um bestimmte Anschlüsse zu garantieren: `CONNECTION:PreTripID;ConnectionTripID`

Solche Zusammenhänge können ohne größeren Aufwand im TSN-Modell berücksichtigt werden, die Anschlußbeziehungen sind jedoch leider recht kompliziert. In der Regel müssen Anschlüsse während des Fahrtverlaufs (mitten in der Linienfahrt, z.B. am Busbahnhof) gehalten werden und eher selten an den Endhaltestellen. Zudem existieren bei den meisten Kunden keine Anschlußdefinitionen. Daher haben wir es im aktuellen System vorläufig nicht berücksichtigt. Eine spätere Erweiterung diesbezüglich wäre durch zusätzliche Restriktionen im mathematischen Modell zu realisieren. Die aktuelle Lösung besteht in einer entsprechenden Definition der Verschiebeintervalle, so dass die betroffenen Fahrten nur so weit verschoben werden dürfen, wie es die Anschlussbeziehungen erlauben.

6.3 Das Depotgruppen-Konzept

In Kapitel 4 wurde die Modellierung des allgemeinen Mehrdepot-Umlaufplanungsproblems formuliert. In der Praxis werden oft zusätzliche Nebenbedingungen oder Erweiterungen des Modells gewünscht. Eine der möglichen Erweiterungen ist die Regelung, dass das Start- und Enddepot eines Umlaufs bzw. Umlaufstücks zwischen zwei Depotaufenthalten verschieden sein kann. Das kann mit Hilfe des „*Depotgruppen-Konzepts*“ modelliert werden. Durch diese Regelung können die Anzahl der Leerfahrten und „teures“ Warten außerhalb des Depots reduziert werden.

Das Liniennetz mittelgroßer und großer Städte ist dicht und komplex. Der Bedarf an Personenbeförderung ist groß, so dass selten alle Fahrten nur von einem Anbieter des ÖPNV bedient werden können. Für ein Unternehmen kann es sinnvoll sein, in die eigene Planung auch Depots der Partnerunternehmen einzubeziehen, um die Lösung des Umlaufplanungsproblems evtl. zu verbessern. Auch wenn ein Unternehmen den kompletten Bedarf an Personenbeförderung einer Stadt selbst

deckt, kann eine Unterteilung aller Depots in Gruppen (z.B. nach Stadtteilen) zur Kostensenkung führen.

6.3.1 Einsatzmöglichkeit

Bei der allgemeinen Modellierung wird jeder Umlauf eindeutig einem Depot zugeordnet. Diese Zuordnung bedeutet, dass ein entsprechendes Fahrzeug nur in diesem Depot sowohl starten und enden als auch alle Pausen machen darf. In der Regel verursachen sämtliche Pausen im eigenen Depot, im Gegensatz zu Pausen an Haltestellen, keine Personalkosten. Allerdings ist eine Fahrt zum eigenem Depot und zurück zwecks einer Pause in vielen Fällen wegen der großen Entfernung zum Depot nicht rentabel bzw. unmöglich.

Die Regelung, dass das Start- und Enddepot eines Umlaufs bzw. Umlaufstückes zwischen zwei Depotaufhalten verschieden sein kann, kann Kostenersparnisse mit sich bringen. Dadurch können nicht nur lange Aus- bzw. Einsetzfahrten gespart werden, sondern auch viele Leerfahrten, die durch eine Fahrt über eins der vordefinierten Depots ersetzt werden können. Außerdem wird in vielen Fällen anstatt des „teuren“ Wartens an einer Anschlusshaltestelle eine günstigere Alternative mit einer Pause in einem der Depots möglich.

Eine weitere Kostensenkung entsteht, falls einem Fahrzeug die Möglichkeit gegeben wird, seinen Umlauf nicht nur in eigenen und fremden Depots für Pausen unterbrechen zu können, sondern auch an bestimmten Haltestellen, die wegen ihrer speziellen Eigenschaften auch für „tarifliche“ Pausen geeignet sind. Das können große Haltestellen (z.B. Bahnhof) oder Rasthäuser sein, wo es sanitäre Einrichtungen gibt, aber keine Möglichkeit, die Fahrzeuge zu warten oder nachts abzustellen, um sie als Depots definieren zu können.

Zur Verdeutlichung der oben beschriebenen Aspekte wird an dieser Stelle ein Beispiel angeführt. Sei ein Umlauf bei der Umlaufplanung so bestimmt worden, dass er in der Nähe des Depots A startet und in der Nähe des Depots B endet, siehe Abbildung 6.23 (Fahrten f_1, \dots, f_6). Dieser Umlauf wird entweder von einem Fahrzeug des Depots A oder von einem Fahrzeug des Depots B bedient. Im ersten Fall ist keine kostengünstige Pause im Depot zwischen den Fahrten f_2 und f_3 wegen der langen Fahrzeit zum Depot A möglich, und außerdem soll eine lange Aussetzfahrt nach der Bedienung der Fahrt f_6 zurück zum Depot A ausgeführt werden (Abbildung 6.23, links). Ähnlich im zweiten Fall, hier ist keine Pause im Depot zwischen den Fahrten f_4 und f_5 möglich und eine lange Einsetzfahrt vom Depot B zum Startpunkt der Fahrt f_1 notwendig (Abbildung 6.23, rechts).

Fängt der Umlauf im Depot A an und darf außerdem Depot B sowohl für Pausen als auch für die Stationierung genutzt werden, dann kann er günstiger bedient

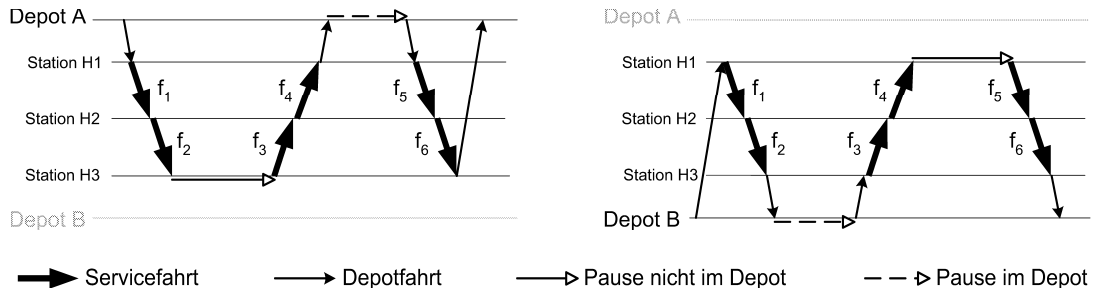


Abbildung 6.23: Nur-Eigen-Depot A und Nur-Eigen-Depot B

werden. Es sind weder lange Ein-/Aussetzfahrten noch langes Warten außerhalb des Depots mehr notwendig, siehe Abbildung 6.24.

6.3.2 Modellierung

Die oben genannten Vorteile können durch eine alternative Modellierung bzw. eine Anpassung der existierenden Modellierung erreicht werden. Diese Anpassung stellt das Depotgruppen-Konzept dar. Bei diesem Konzept werden aus der Menge aller Depots D die Untermengen D_i gebildet, die nicht unbedingt disjunkt sein müssen. Solche Untermengen werden *Depotgruppen* genannt und weisen folgende Eigenschaft auf:

Ein Fahrzeug, das in einem Depot $d \in D_i$ gestartet ist, darf seinen Umlauf in jedem anderen Depot aus D_i beenden bzw. für Pausen unterbrechen.

Das Flussnetzwerk der allgemeinen Modellierung wird ebenfalls entsprechend angepasst. Sei D die Menge aller Depots und G Menge aller Depotgruppen, wobei jede Gruppe mindestens ein Depot beinhaltet. Dann wird bei dem Netzwerkaufbau nicht für jedes Depot-Fahrzeugtyp-Paar, sondern für jedes Depotgruppe-Fahrzeugtyp-

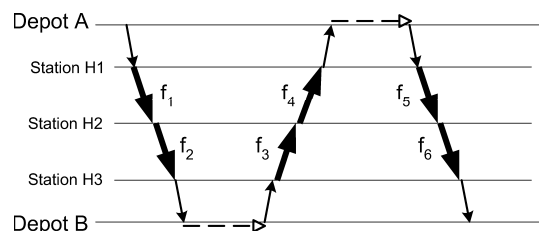


Abbildung 6.24: Depot A mit Pausen in B und Rückkehr zu B

Paar eine Netzwerkschicht angelegt. Der Unterschied zu dem zuvor beschriebenen Flussnetzwerk besteht darin, dass eine Schicht für die Gruppe $g \in G$ mehrere Depots beinhalten kann. Das heißt, dass für jede Fahrt i nicht mehr maximal 2, sondern maximal $2 \cdot |g|$ zusätzliche Depotkanten in das Netzwerk eingefügt werden müssen, und zwar jeweils eine Ausrückfahrt-Kante von jedem Depot $d \in g$ zur Fahrt i , falls die Abfahrtsstation von i ungleich d ist, und eine Einrückfahrt-Kante von i zu jedem Depot $d \in g$, falls Ankunftsstation von i ungleich d ist. Außerdem gibt es in jeder Schicht für die Depotgruppe g nicht eine, sondern $|g|$ Zirkulationsflusskanten.

In der mathematischen Formulierung des Problems aus dem Kapitel 4 wird die Menge K weiterhin nicht als Menge aller Depot-Fahrzeugtyp-Kombinationen, sondern als Menge aller Depotgruppe-Fahrzeugtyp-Kombinationen $K = G \times T$ interpretiert.

Sei $x_{dd}^{g,t}$ Flussvariable für die Zirkulationsflusskante des Depots $d \in g$ in der Netzwerkschicht k , wobei k eine Kombination aus der Depotgruppe $g \in G$ und dem Fahrzeugtyp $t \in T$ repräsentiert. Da ein Depot bei dem Depotgruppen-Konzept in mehreren Schichten vertreten sein kann, ist es nicht mehr ausreichend, die Kapazitätsrestriktionen der Stellplätze in Depots als obere Schranke der Zirkulationsflusskanten zu definieren. Stattdessen muss der Fluss über alle entsprechenden Zirkulationsflusskanten der zugehörigen Schichten aufsummiert und begrenzt werden. Sei $B_{d,t}$ die Kapazität der Stellplätze für den Fahrzeugtyp $t \in T$ im Depot $d \in D$. Dann wird das gemischt-ganzzahlige Programm (4.8)-(4.14) um eine weitere Nebenbedingung ergänzt:

$$\sum_{g \in G, \text{ mit } d \in g} x_{dd}^{g,t} \leq B_{d,t} \quad \forall d \in D, \forall t \in T \quad (6.9)$$

Die Anzahl der Fahrzeuge von jedem Typ am Ende des Tages in jedem Depot $d \in D_i$ bleibt gleich der Anzahl am Anfang des Tages. Somit werden in jedem Depot $d \in D_i$ nicht unbedingt dieselben Fahrzeuge über die Nacht stationiert, allerdings Fahrzeuge desselben Typs, so dass alle Fahrten am nächsten Tag wieder bedient werden können. Außerdem kehrt ein Fahrzeug, das ein Depot d aus der Gruppe D_i verlassen hat, spätestens nach $|D_i|$ Tagen wieder in d zurück. Allerdings erfolgt das nicht automatisch, sondern nur, wenn ein Fahrzeug höchstens einmal in jedem anderen Depot der Gruppe über Nacht stationiert wird, ohne wieder das eigene Depot zu erreichen. Dies kann durch eine bestimmte Zuordnung der Umläufe zu physikalischen Fahrzeugen erreicht werden. Jedem Bus b_i^k ist ein *Homedepot* $d_k \in D_j$ eindeutig zugeordnet. In der Variable U_i^k werden alle anderen Depots der Gruppe D_j gespeichert, wo b_i^k über Nacht schon stationiert wurde, seitdem er d_k verlassen hatte. Beendet der Bus b_i^k seinen Umlauf eines Tages wieder im Homedepot d_k , wird U_i^k neu initialisiert. Ferner sollen in jedem Depot bei der

Zuordnung der Umläufe zu den Fahrzeugen folgende Regeln eingehalten werden:

- die Umläufe, die sowohl im Homedepot anfangen als auch enden, werden zu den Fahrzeugen zugeordnet, die immer nur in diesem Depot stationiert sind (*depot-eigene Fahrzeuge*),
- jedem restlichen *depot-eigenen* oder *depot-fremden* Fahrzeug b_i^k wird ein Umlauf zugeordnet, der als Ziel ein Depot d_m hat, in dem b_i^k noch nicht über Nacht stationiert war, also $d_m \notin U_i^k$ (dabei kann der Umlauf, bei dem das Zieldepot mit dem Homedepot des Fahrzeuges b_i^k übereinstimmt, präferiert werden, um alle Fahrzeuge möglichst schnell wieder in ihre Homedepots zu schicken).

Ein Fahrzeug wird höchstens einmal in allen anderen Depots der Gruppe D_j über Nacht stationiert sein, bevor es wieder sein Homedepot erreicht. Somit kann jedes Fahrzeug spätestens alle $|D_j|$ Tage in seinem „eigenen“ Depot gewartet werden.

Das ursprüngliche Problem mit der Forderung, dass jedes Fahrzeug wieder in sein Depot zurückkehren muss, kann relativ einfach mit dem Depotgruppen-Konzept modelliert werden. Dafür wird jedes Depot einer separaten Depotgruppe zuordnet, die nur es allein beinhaltet.

Welche Depots zu einer Gruppe zusammengefasst werden können, hängt von der Struktur des Liniennetzes, technischen Charakteristiken und Aufgaben des betreffenden Depots sowie Zielen des Unternehmens ab. Ist eine möglichst große Kostenreduzierung von Interesse, dann können auch solche Depots zu einer Gruppe zusammengeführt werden, die geographisch weit auseinander liegen. Allerdings, müssen dann die Fahrer am Ende ihrer Dienstschicht mit anderen öffentlichen Verkehrsmitteln zum dem Ort zurückkehren, an dem sie eventuell ihren privaten PKW am Anfang der Schicht abgestellt hatten (Alternativ kann ein Fahrer seinen Wagen dort abstellen, wo seine Schicht enden wird und den Einsatzort mit anderen öffentlichen Mitteln erreichen). Eine andere Möglichkeit ist, solche Depots zusammenzufassen, die sich geographisch in der Nähe befinden (z.B. ein Stadtteil), die wiederum ihre Vor- und Nachteile hat.

Beispiele

Wie verschiedene Szenarien, die in der Praxis denkbar wären, modelliert werden können, zeigen folgende Beispiele:

Beispiel 1: Sei eine Menge $D = \{A, B, C, D\}$ verschiedener Depots gegeben. In der Vormittags- und Mittagszeit finden viele Fahrten des Fahrplans in der Nähe

vom Depot A statt (z.B. Schule, Büroviertel). Abends gibt es dagegen viele Fahrten, die in der Nähe vom Depot B stattfinden (z.B. Freizeitzentrum, Wohnviertel). Da bei der Umlaufplanung die Anzahl der eingesetzten Fahrzeuge minimiert wird, bedienen viele Fahrzeuge sowohl die Fahrten in der Nähe von A als auch von B. Dabei sind entweder lange Ein- oder lange Aussetzfahrten notwendig, um die Fahrzeug an die Startposition bzw. wieder ins Depot zu bringen.

→ *Aus diesem Grund werden die Depots A und B zu einer Gruppe zusammengefasst, um die Anzahl der Aus- bzw. Einsetzfahrten zu reduzieren und Personalkosten bei der Pausendurchführung zu senken.*

Es werden 3 Depotgruppen $g_1 = \{A, B\}$, $g_2 = \{C\}$ und $g_3 = \{D\}$ definiert.

Beispiel 2. Sei eine Menge $D = \{A, B, C, D\}$ verschiedener Depots und eine Haltestelle $h \in H$ gegeben. Die Haltestelle h liegt in der Stadtmitte am Hauptbahnhof. Der Hauptbahnhof verfügt über einen großen Busparkplatz, der nie voll besetzt wird, aber über Nacht und am Wochenende auch nicht überwacht wird. Er kann daher in dieser Zeit nicht in Anspruch genommen werden. Am Bahnhof sind einige Pausenräume vorhanden, sodass er für tariflich unbezahlte Pausen verwendet werden kann.

→ *Der Hauptbahnhof wird als reines Pausendepot definiert. Die Haltestelle h wird in die Menge aller Depots D aufgenommen, allerdings mit einer verfügbaren Stellplatzkapazität von 0; also $D = \{A, B, C, D, h\}$. Es werden 4 Depotgruppen $g_1 = \{A, h\}$, $g_2 = \{B, h\}$, $g_3 = \{C, h\}$ und $g_4 = \{D, h\}$ definiert. Somit können alle Fahrzeuge aus allen anderen Depots $d \in (D - h)$ ihre Pausen auch im Depot h durchführen.*

6.3.3 Ergebnisse

Testinstanzen

Beim Testen des Depotgruppen-Konzepts wurden drei Fahrpläne verwendet, deren Eigenschaften in der Tabelle 6.8 dargestellt sind. Abhängig von dem durchzuführenden Test sind in den Fahrplänen entsprechende Haltestellen als Depots definiert (siehe Depot-IDs in der Tabelle 6.8) und ggf. zu Depotgruppen zusammengefasst worden.

Zusammenfassung von Depots

Es wurde untersucht, wie die Zusammenfassung von 2 bzw. 3 Depots zu einer Depotgruppe (siehe Tabelle 6.9 bzw. 6.10) sich auf die Problemgröße, Laufzeit

	Servicefahrten	Haltestellen	Fahrzeugtypen	Depot-IDs	
				2 Depots	3 Depots
Fahrplan 1	2047	21	3	7, 47	7, 47, 316
Fahrplan 2	3054	48	1	4, 5	4, 5, 426
Fahrplan 3	7068	125	1	4, 5	4, 5, 426

Tabelle 6.8: Fahrplandaten

	Schichten	Anzahl			Laufzeit (sek.)	Operative Kosten
		Variablen	Restriktionen	Non-Zeros		
Fahrplan 1						
ohne Depotgruppen	6	16.218	60.409	127.398	190,3	1.027,741
mit Depotgruppen	3	10.593	33.314	69.924	9,1	1.025,209
Fahrplan 2						
ohne Depotgruppen	2	16.064	128.001	262.114	47,3	55.424,75
mit Depotgruppen	1	10.438	68.081	139.220	10,7	55.419,26
Fahrplan 3						
ohne Depotgruppen	2	34.076	490.719	995.578	278,7	75.921,68
mit Depotgruppen	1	21.758	241.876	490.806	46,7	75.917,87

Tabelle 6.9: 2-Depots-Instanzen ohne und mit Depotgruppen

und operative Kosten gegenüber der allgemeinen Modellierung ohne Depotgruppen-Konzept auswirkt.

Die Anzahl der Netzwerkschichten bei dem Depotgruppen-Konzept hängt nicht von der Anzahl der Depots, sondern von der Anzahl der Depotgruppen ab, vgl. Tabellen 6.9 und 6.10. Da die Lösungszeit des Problems mit jeder neuen Schicht exponentiell wächst, kann die Laufzeit der Optimierung durch passende Bildung der Depotgruppen „im Rahmen“ gehalten werden. Wird beispielsweise einer Gruppe mit nur einem Depot ein zweites hinzugefügt, wird die Ein-Depot-Schicht um ein zweites Depot erweitert. Dabei werden für jede Fahrt maximal zwei zusätzliche Kanten zum und vom zweiten Depot zugefügt. Entscheidend ist aber, dass die komplexitätsproblematischen Verbindungsfahrten, die bei der zweiten Schicht nochmal eingeführt werden sollten, nur einmal oder gar nicht mehr im Netzwerk vertreten sind (eine Fahrt über das zweite Depot kann evtl. günstiger als eine direkte Verbindungsfahrt sein). Somit verändert sich die Modellgröße und die Laufzeit beim Zufügen eines weiteren Depots zu der existierenden Depotgruppe nur geringfügig, vgl. Tabellen 6.9 und 6.10.

	Anzahl			Laufzeit (sek.)	Operative Kosten	
	Schichten	Variablen	Restriktionen			Non-Zeros
Fahrplan 1						
ohne Depotgruppen	9	23.314	92.105	194.080	695,7	935,403
mit Depotgruppen	3	11.938	37.769	78.840	11,4	929,780
Fahrplan 2						
ohne Depotgruppen	3	22.257	188.015	385.198	271,6	54.705,80
mit Depotgruppen	1	11.028	68.032	139.124	10,5	54.614,28
Fahrplan 3						
ohne Depotgruppen	3	47.194	722.825	1.466.860	10783	75.095,05
mit Depotgruppen	1	22.570	225.167	457.408	42,7	75.073,47

Tabelle 6.10: Drei Depots ohne und mit Depotgruppen

Ist ein Transportunternehmen bereit die „nur-eigenes-Depot“-Forderung abzuschwächen, indem einem Fahrzeug erlaubt wird, seinen Umlauf auch in anderen vordefinierten Depots zu beenden bzw. für Pausen zu unterbrechen, können viele große Probleme mit mehreren Depots schneller gelöst werden. Z.B. wird das 3-Depot Problem für den Fahrplan 3 um 252 Mal schneller gelöst, wenn alle drei Depots zu einer Depotgruppe zusammengefasst werden, siehe Tabelle 6.10.

Durch die Zusammenfassung von Depots wird das gesamte Problem nicht nur schneller gelöst, sondern es wird auch ein besseres Ergebnis bzgl. der Kosten erreicht. Diese Kostenersparnis ist zwar nicht groß, wird aber mit einer steigenden Anzahl an Depots immer deutlicher. So hat sich die Lösung mit drei zusammengefassten gegenüber der mit zwei zusammengefassten Depots beim Fahrplan 1 um 9,3%, Fahrplan 2 um 1,45% und Fahrplan 3 um 1,11% verbessert (vgl. Tabellen 6.9 und 6.10). Allgemein gilt, je mehr Depots zusammengefasst werden, umso günstiger können die Umläufe bedient werden.

Verteilung von „depot-eigenen“ Umläufen

Umläufe, die immer nur ein Depot sowohl für Pausen als auch für die Stationierung nutzen, werden hier „*depot-eigene*“ Umläufe genannt. Alle anderen Umläufe heißen „*depot-gemischte*“. Bei der allgemeinen Modellierung ohne Depotgruppen sind somit alle Umläufe „depot-eigen“.

Die Tabelle 6.11 zeigt für berechnete Umlaufpläne, wie viele Umläufe andere Depots aus der Depotgruppe nutzen und wie viele dagegen „depot-eigene“ sind.

Wie die Tests zeigen, sind die meisten Umläufe sowohl im 2- als auch im 3-

	Fahrplan 1		Fahrplan 2		Fahrplan 3	
	2 Depots	3 Depots	2 Depots	3 Depots	2 Depots	3 Depots
alle Umläufe	115	115	118	118	242	242
davon						
- „depot-eigene“	69	64	84	67	172	142
- „depot-gemischte“	46	51	34	51	70	100

Tabelle 6.11: Verteilung der „depot-eigenen“ und „depot-gemischten“ Umläufe

Depot Fall depot-eigene und weisen somit alle Eigenschaften der Umläufe bei der allgemeinen Modellierung ohne Depotgruppen-Konzept auf, siehe Tabelle 6.11, wobei aber die Laufzeit viel geringer als bei der allgemeinen Modellierung war (siehe Tabellen 6.9 und 6.10). Die Fahrzeuge, die andere, depot-gemischte Umläufe bedienen, kehren maximal nach 2 bzw. 3 Tagen wieder in das Hometown zurück und können entsprechend gewartet und gepflegt werden.

Es sind einige Konzepte vorstellbar, die den Anteil der depot-eigenen Umläufe weiter erhöhen können. Zu einem kann eine alternative Dekompositionsstrategie oder ein exaktes Verfahren bei der Generierung der Umläufe entworfen und eingesetzt werden, das die Anzahl der depot-eigenen Umläufe maximiert. Alternativ ist auch eine Heuristik denkbar, die bei einem depot-gemischtem Umlauf ein solches Umlaufstück findet, das durch ein anderes, aus einem anderen depot-gemischtem Umlauf, ersetzt werden kann, wobei einer der Umläufe dadurch zum depot-eigenen wird.

Einfügen von Pausendepots

Die Tabelle 6.12 repräsentiert die Ergebnisse der Untersuchung von Laufzeit- und Kostenverhalten beim Zufügen reiner Pausendepots beim Fahrplan 2 (vgl. Abschnitt 6.3.1). Die eckigen Klammern repräsentieren Depotgruppen mit Depot-IDs, wobei reine Pausendepots unterstrichen sind. Wird z.B. die Haltestelle mit ID 426 als ein Pausendepot definiert, sodass alle Fahrzeuge sie für ihre Pausen nutzen dürfen, können die operativen Kosten bei relativ gleich bleibender Laufzeit fast um 1% gesenkt werden (vgl. Zeilen 1 und 3 der Tabelle). Werden weitere Haltestellen als Pausendepots definiert, können die operativen Kosten weiter gesenkt werden (vgl. Zeilen 1, 3 und 4 der Tabelle).

Die Pausendepots wurden in diesem Beispiel zufällig ausgewählt, da ihre geographische Lage aus dem Fahrplan nicht hervorgeht. In der Praxis spielt es aber eine Rolle, wo sie sich befinden. Ein Pausendepot in der Mitte eines Liniennetzes kann beispielsweise mehr Kostenersparnisse bringen als ein Pausendepot am

Nr.	Depotgruppen (Pausendepots sind unterstrichen)	Laufzeit (sek.)	Kosten
1	< 4 >; < 5 >	47,3	55.424,75
2	< 4, 5 >	10,7	55.419,26
3	< 4, <u>426</u> >; < 5, <u>426</u> >	57,2	54.961,51
4	< 4, <u>426</u> , <u>26</u> >; < 5, <u>426</u> , <u>26</u> >	38,2	54.817,66
5	< 4 >; < 5 >; < 426 >	271,6	54.705,80
6	< 4, <u>26</u> >; < 5, <u>26</u> >; < 426, <u>26</u> >	354,7	54.577,45
7	< 4 >; < 5 >; < 426 >; < 26 >	575,7	54.302,19

Tabelle 6.12: Hinzufügen von Pausendepots (Fahrplan 2)

Stadtrand.

6.4 Weitere praktische Aspekte

Nachfolgend sollen im Überblick weitere Aspekte aus der Praxis der Umlaufplanung beschrieben werden, die in der Implementierung berücksichtigt wurden.

6.4.1 Kapazitäten

Verschiedene Arten von Kapazitäten können im mathematischen bzw. im Netzwerkmodell berücksichtigt werden:

- Stellplatzkapazität
- Begrenzungen in der Fahrzeugflotte
- Fahrzeugtyp-Depot Kapazitäten bzw. feste Vorgaben

Die Stellplatzkapazitäten werden für jeden Abstellort vorgegeben. Im Modell muss für jedes Depot die gesamte Stellplatzkapazität für alle Fahrzeuge aller Typen berücksichtigt werden. Dabei wird das gemischt-ganzzahlige Programm aus dem Kapitel 4 um folgende Nebenbedingung erweitert:

$$\sum_{t \in T} \sum_{g \in G, \text{ mit } d \in g} x_{dd}^{g,t} \leq B_d \quad \forall d \in D, \quad \text{wobei} \quad (6.10)$$

B_d die gesamte Stellplatzkapazität für Depot $d \in D$ und

$x_{dd}^{g,t}$ die Zirkulationsflusskante für das Depot $d \in g$ in der Netzwerkschicht, die ein Depotgruppe-Fahrzeugtyp-Paar $g \in G$ und $t \in T$ repräsentiert, ist.

Die Nebenbedingung 6.10 bewirkt, dass die Summe der Flüsse über Zirkulationsflusskanten der Netzwerkschichten, die dem Depot d entsprechen, durch B_d - die Anzahl der im Depot d verfügbaren Stellplätze begrenzt wird.

Soll die Anzahl der verfügbaren Fahrzeuge eines Fahrzeugtypen begrenzt werden, wird im Modell für jeden Fahrzeugtyp die Gesamtzahl der Umläufe, die von diesem Fahrzeugtyp gefahren werden, über alle Depots begrenzt. Dabei wird folgende Nebenbedingung im mathematischen Modell hinzugefügt:

$$\sum_{g \in G} \sum_{d \in g} x_{dd}^{g,t} \leq B_t \quad \forall t \in T, \quad \text{wobei} \quad (6.11)$$

B_t die Gesamtzahl aller verfügbaren Fahrzeuge vom Typ $t \in T$ und

$x_{dd}^{g,t}$ wie in (6.10) definiert ist.

Die Nebenbedingung 6.11 bewirkt, dass der Fluss über die Zirkulationsflusskanten in allen Schichten, die den Typ $t \in T$ beinhalten, aufsummiert und durch B_t begrenzt wird.

Die Begrenzungen der Anzahl Fahrzeuge für bestimmte Fahrzeugtyp-Depot Kombinationen werden durch entsprechende Flusschranken auf den Zirkulationsflusskanten in der Netzwerkschicht, die dieser Kombination entspricht, gesetzt. Es werden dabei sowohl obere als auch untere Schranken berücksichtigt, was bei einer Gleichheit der oberen und unteren Schranke feste Vorgaben für die Anzahl Fahrzeuge realisieren lässt.

6.4.2 Mindestaufenthaltszeit im Depot

Zur Berücksichtigung einer vorgegebenen Mindestaufenthaltszeit im Depot wird im Netzwerkmodell eine Modifikation vorgenommen. Dafür werden die Kanten für die Depoteinrückfahrten an einer anderen Stelle in der Depot-Timeline eingefügt. Und zwar werden sie nicht mehr zum Zeitpunkt, zu dem die entsprechende Depotfahrt im Depot ankommt eingefügt, sondern genau um die Mindestaufenthaltszeit später. Die Kostenberechnung bleibt davon unbeeinflusst. Nach dieser Netzwerktransformation sind nur solche Ankünfte und Abfahrten im Depot kompatibel, die den entsprechenden zeitlichen Abstand (die Mindestaufenthaltszeit) erfüllen.

6.4.3 Tageszeitabhängige Verbindungsdauer und -Entfernung

Der Planer hat eine Möglichkeit die Zeitintervalle vorzugeben, in denen die jeweilige Verbindungsdauer und Entfernung gültig ist. Es gibt keine Standardverbindungs-dauer zwischen A und B. Vielmehr gibt es mehrere unterschiedliche Verbindungs-dauern, die zu unterschiedlichen Tageszeiten gültig sind. Zur Verdeutlichung wird nachstehendes Beispiel angeführt. Beträgt die Fahrtzeit von A nach B nachmittags (Abfahrt ab 12 Uhr) 15 Minuten und Vormittags 20 Minuten, so kommt ein um 12:00 in A startendes Fahrzeug, in B früher an, als ein anderes, das um 11:59 - also früher - gestartet ist.

Ein Ausschnitt aus der Relation für Verbindungsfahrten kann wie folgt aussehen:

```

FromStopID, ToStopID; FromTime; ToTime; Distance; RunTime;
1; 104; 000 : 00 : 00; 000 : 06 : 29; 3760; 540
1; 104; 000 : 06 : 30; 001 : 00 : 00; 3760; 720
104; 1; 000 : 00 : 00; 001 : 12 : 00; 2810; 480

```

Dies würde bedeuten, dass eine Fahrt von der Haltestelle 1 zu der Haltestelle 104 vor 6 : 30 Uhr morgens nur 9 Minuten danach aber ganze 12 Minuten in Anspruch nehmen würde. Eine so genaue Berücksichtigung der Verbindungsdauer ist wichtig, um wirklich funktionierende Lösungen zu produzieren. Eine grobe Darstellung würde dazu führen, dass Umsetzfahrten eingeplant werden, die so gar nicht zustande kommen dürfen oder aber mehr Fahrzeuge als nötig eingeplant werden, falls die Verbindungszeit größer angenommen wird, als tatsächlich zu der jeweiligen Tageszeit gegeben.

Variable (tageszeitabhängige) Fahrtzeiten für die Verbindungsfahrten können aber das exakte Modell zerstören, da es dadurch im Modell zu einer (für die Wirklichkeit) etwas absurden Situation kommen kann. Und zwar könnte ein Fahrzeug, das in einer Haltestelle früher gestartet ist, später an einer anderen Haltestelle ankommen, als ein anderes, das später an derselben Haltestelle gestartet ist. Beträgt zum Beispiel die Fahrtzeit von A nach B nachmittags (Abfahrt ab 12 Uhr) 15 Minuten und Vormittags 20 Minuten, so kommt ein um 12:00 in A startendes Fahrzeug, früher in B an, als ein anderes, das um 11: 59 - also früher - gestartet ist. Für solche Fälle ist eine Sonderbehandlung der Verbindungskanten im Netzwerkmodell notwendig. Dafür sind entsprechende Datenstrukturen und Kontrollmechanismen vorgesehen und implementiert.

6.4.4 Vergabe von Fahrten

Verkehrsbetriebe sind in der Regel bestrebt, die für den eigenen Betrieb ungeeignete Fahrten durch private (Sub-)Unternehmer durchführen zu lassen. Die Auswahl solcher Fahrten geschieht in der heutigen Praxis eher unsystematisch. Für die Auswahl solcher Fahrten sind zwei Verfahren denkbar:

Einbeziehung von privaten Unternehmern in die Multi-Depot-Planung

Der Vorteil hierbei ist, dass auch für den privaten Unternehmer kostengünstige Umläufe geplant werden, die dieser preiswert für den Verkehrsbetrieb anbieten kann. Zusätzliches Optimierungsziel muss dann sein, möglichst wenige Fahrzeuge von bestimmten Depots zu verwenden und dafür möglichst kostengünstige Umläufe zu planen.

Gewichtung der Linienfahrten Eine Gewichtung der Linienfahrten abhängig davon, mit welcher Priorität sie ausgeführt werden sollen. D.h. der Optimierungsalgorithmus erhält Fahrten, die unbedingt ausgeführt werden müssen, und solche die u.U. auch ausgelassen werden können. Weiterhin soll in diesem Fall festgelegt werden wie viele Fahrten, unter welchen Umständen ausgelassen werden dürfen.

Die erste Variante wird in der aktuellen Version der Software durch eine Einbeziehung der Depots externer Unternehmen als zusätzliche Depots mit angepassten Kostensätzen verwendet. Die Modellierung der zweiten Variante erfolgt durch entsprechenden Änderungen im Netzwerk und zusätzliche Restriktionen im mathematischen Modell.

6.4.5 Vorbereitungs- und Wendezeiten

Unter Vorbereitungs- und Wendezeiten werden Pufferzeiten verstanden, die unter anderem für eine Umrüstung des Fahrzeugs, für den Fahrerwechsel und zum Abfangen von infolge von Störungen auftretender Fahrzeugverspätungen eingeplant werden. Das Modell wurde diesen Anforderungen angepasst dadurch, dass die Vorbereitungszeit von der Abfahrtszeit der entsprechenden Fahrt abgezogen und die Wendezeit auf die Ankunftszeit addiert wird. Diese Zeiten können entweder für alle Fahrten des Fahrplans gleich oder individuell für jede Fahrt angegeben werden und werden beachtet, falls der Planer den entsprechenden Schalter gesetzt hat.

Kapitel 7

Maßnahmen zur Kostensenkung in der Umlaufplanung

In diesem Kapitel werden unterschiedliche Möglichkeiten der Kosteneinsparung durch effiziente Ressourceneinsatzplanung aufgezeigt. Jede Maßnahme wird zunächst durch ein einfach konstruiertes Beispiel erklärt und dann anhand realer Datensätze in Bezug auf Einsparungsmöglichkeiten quantitativ untersucht.

7.1 Untersuchte Instanzen

Die Kennzahlen der untersuchten Instanzen sind in der Tabelle 7.1 enthalten. Für die ersten Analysen wurde ein Fahrplan der Halleschen Verkehrs AG (HAVAG) verwendet. Dieser Fahrplan ist im Vergleich zu den anderen untersuchten Plänen (Münchener Teilprobleme) relativ klein, so dass er sich sehr gut für umfangreiche Untersuchungen eignet.

Für die weiteren Untersuchungen wurden modifizierte Fahrpläne des Fahrplans der Münchener Verkehrsgesellschaft mbH (MVG) verwendet. Diese sind unterschiedlich komplex. Die Parameter der Basisversion sind unten angegeben. Die fixen Kosten für Fahrzeuge sind bei dieser Instanz extrem hoch gesetzt, sie entsprechen etwa dem 1000-fachen der fixen Fahrzeugkosten der Halle-Instanz. Ursprünglich wurden diese Kosten künstlich erhöht, um zu gewährleisten, dass der Optimierer primär die Anzahl der Fahrzeuge reduziert und erst dann die operativen Kosten minimiert. Experimentelle Untersuchungen zeigen aber, dass diese beiden Ziele, nämlich die Reduzierung der fixen Kosten und die der variablen Kosten, keine konkurrierenden Ziele sind. Eine Minimierung der operativen Kosten führte in allen Fällen zu den Lösungen mit minimaler Fahrzeugflotte.

Als letztes wurden noch Instanzen der BVO Busverkehr Ostwestfalen GmbH

(BVO) untersucht. Diese Instanz unterscheidet sich von den beiden vorherigen vor allem dadurch, dass es sich bei der BVO um ein Unternehmen handelt, das für den regionalen Nahverkehr zuständig ist. Der regionale Nahverkehr ist hierbei als die Personenbeförderung zwischen benachbarten Städten zu verstehen. Bei den beiden anderen Unternehmen handelt es sich um Betreiber für die innerstädtische Personenbeförderung. Folglich ist der Fahrplan der BVO nicht so dicht wie der der beiden anderen Unternehmen und die Strecken einer Linie sind teilweise erheblich länger (s. Tabelle 7.1). Die Instanz der BVO enthielt keine direkten Angaben über die möglichen Leerfahrten. Allen Haltestellen sind jedoch Koordinaten zugewiesen, die eine approximative Bestimmung dieser Leerfahrten ermöglichen.

	#Fahrten	#Depots	#Haltestellen	#Fzg.-typen	∅Linie
Halle	2047	2	21	3	5 km
München	11063	18	160	12	6,5 km
BVO	1219	5	168	3	23 km

Tabelle 7.1: Eigenschaften untersuchter Testfälle

7.2 Linienbündelung

Die linienreine Planung ist ein gängiges Vorgehen in der Planung der Fahrzeugumläufe. Linienreine Planung bedeutet in diesem Fall, dass ein Fahrzeug nur Fahrten genau einer Linie durchführen kann, bevor es zum Depot zurückkehrt. Für diese Linie kann es dann aber beliebig viele Fahrten durchführen. Im Gegensatz zu der linienreinen Planung lässt es die liniengemischte Planung zu, dass ein Fahrzeug Fahrten mehrerer Linien bedient.

Wird eine liniengemischte Planung durchgeführt, so kann ein Fahrzeug nach Beendigung einer Fahrt von der Ankunftshaltestelle eine Leerfahrt zu einer beliebigen anderen Haltestelle machen und von dort aus eine neue Fahrt beginnen. Die liniengemischte Planung sollte sowohl Einsparungen bei den operativen Kosten als auch bei den Fixkosten ermöglichen. Durch die liniengemischte Planung können Fahrzeuge z.B. in großen Pausen, die sie zwischen zwei Fahrten einer Linie haben, Fahrten einer anderen Linie durchführen. Dadurch verringern sich die Standzeiten der Fahrzeuge und damit die Kosten für das Personal. In der Regel können durch liniengemischte Planung Fahrzeuge eingespart werden.

Das folgende Beispiel soll die beiden Möglichkeiten verdeutlichen. Betrachtet werden Fahrten zweier Linien (Linie 1 und Linie 2). Beide Linien sowie die Fahrzeiten und Haltestellen sind aus den Instanzen der BVO entnommen. Linie 1 ver-

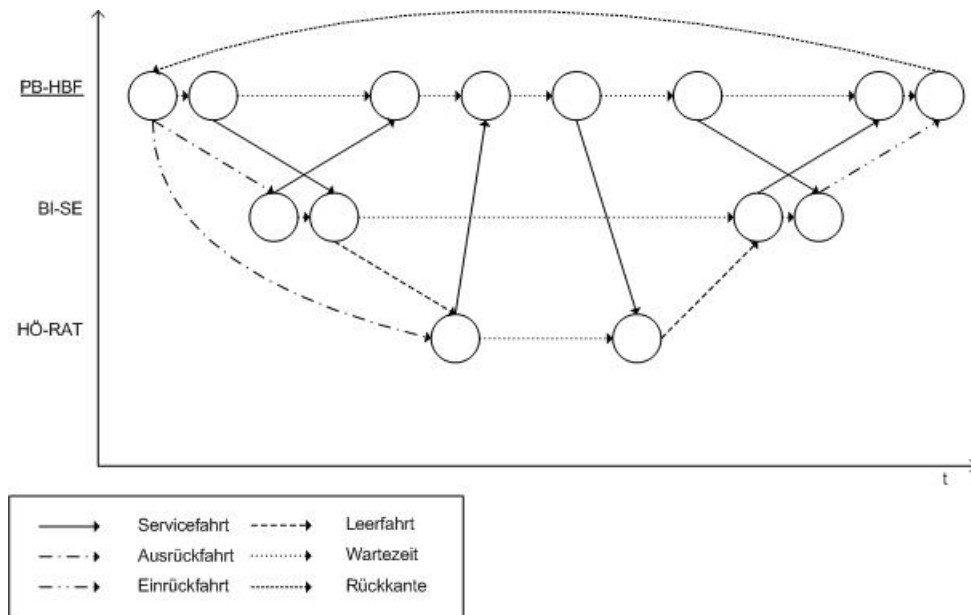


Abbildung 7.1: Time-Space Netzwerk für das Beispiel

bindet die Haltestelle Paderborn, Bus-/Hauptbahnhof (PB-HBF) mit der Haltestelle Bi-Senne (Stadtbahn) (BI-SE). Linie 2 führt vom PB-HBF zur Haltestelle Hövelhof, Rathaus (HÖ-RAT). Die Tabelle 7.2 enthält den daraus entstehenden Fahrplan. Das entsprechende Time-Space Netzwerkmodell gestaltet sich entsprechend der Abbildung 7.1.

Linie	Von	Nach	Abfahrt	Ankunft
Linie 1	PB-HBF	BI-SE	07:00	08:08
Linie 1	BI-SE	PB-HBF	07:30	08:38
Linie 2	HÖ-RAT	PB-HBF	09:00	09:31
Linie 2	PB-HBF	HÖ-RAT	10:00	10:31
Linie 1	PB-HBF	BI-SE	11:00	12:08
Linie 1	BI-SE	PB-HBF	11:30	12:38

Tabelle 7.2: Beispielfahrplan

Eine Lösung dieses Umlaufplanungsproblems sowohl mit linienreiner als auch mit liniengemischter Planung führt zu den Ergebnissen in der Tabelle 7.3. Bei liniengemischter Planung kann ein Fahrzeug die Fahrten der Linie 1 bedienen und in den Standzeiten Fahrten der Linie 2 durchführen. Bei dieser Lösung wird ein Fahrzeug eingespart, was zu geringeren Fixkosten führt. Ebenfalls findet eine Reduzierung der operativem Kosten statt.

Im aktuellen Fall ist diese Reduzierung jedoch sehr gering. Dies liegt daran,

	Kosten gesamt	Kosten variabel	Anzahl Fahrzeuge	Fahrzeug- kosten
Linienrein	5.361	2.361	3	3.000
Liniengemischt	4.323	2.323	2	2.000
% gespart:	19,38	1,64	33,33	33,33
Abs. gespart:	1.039	39	1	1.000

Tabelle 7.3: Beispiel - linienrein vs. liniengemischt

	Kosten gesamt	Kosten variabel	Anzahl Fahrzeuge	Fahrzeug- kosten
Halle linienrein	337.006	134.006	117	203.000
Halle liniengemischt	332.138	131.138	115	201.000
% gespart:	1,44	2,14	1,71	0,99
Abs. gespart:	4.868	2.868	2	2.000

Tabelle 7.4: Halle linienrein vs. liniengemischt

dass sich die operativen Kosten nicht in dem Maße reduzieren, wie es im ersten Moment den Anschein hat. Soll ein Fahrzeug Fahrten mehrerer Linien bedienen, so müssen die Kosten für die Leerfahrt ebenfalls mit in Betracht gezogen werden. Diese können die Einsparungen bei den operativen Kosten machmal erheblich reduzieren. Ein Beispiel aus der Praxis hierfür tritt im folgenden Abschnitt bei der linienreinen Planung für Halle auf.

7.2.1 Tests an ausgewählten Instanzen

Die Untersuchung der linienreinen Planung wurde mit den Instanzen aus Halle, München und der BVO durchgeführt.

	Ausrück- fahrten	Einrück- fahrten	Service- fahrten	Leer- fahrten	Warte- zeit
Halle linienrein	6.637	7.727	102.488	593	16.560
Halle liniengemischt	5.936	6.156	102.488	2.217	14.342
% gespart:	10,57	20,33	0	-273,69	13,4
Abs. gespart:	701	1.571	0	-1.624	2.218

Tabelle 7.5: Halle linienrein vs. liniengemischt, detailliert

	Kosten gesamt	Kosten variabel	Anzahl Fahrzeuge	Fahrzeug- kosten
M. linienrein	2.066.582.114	19.782.114	553	2.046.800.000
M. liniengemischt	1.553.241.447	17.811.447	417	1.535.430.000
% gespart:	24,84	9,96	24,59	24,98
Abs. gespart:	513.340.667	1.970.667	136	511.370.000

Tabelle 7.6: München linienrein vs. liniengemischt

	Ausrück- fahrten	Einrück- fahrten	Leer- fahrten	Service- fahrten	Warte- zeit
M. linienrein	1.872.754	1.784.870	50.487	13.176.113	2.897.890
M. liniengemischt	837.839	943.857	395.758	13.176.113	2.457.880
% gespart:	55,26	47,12	-683,88	0,00	15,18
Abs. gespart:	1.034.915	841.013	-345.271	0	440.010

Tabelle 7.7: München linienrein vs. liniengemischt, detailliert

Halle

In den Tabellen 7.4 und 7.5 werden die Ergebnisse der Haller Instanzen dargestellt. Auffällig ist, dass es keine großen Einsparungen gibt, wenn von linienreiner zu liniengemischter Planung gewechselt wird. Beim Fahrplan von Halle können zwar zwei Fahrzeuge eingespart werden, dies macht jedoch nur eine Reduzierung der Fixkosten von 1% aus. Bei den operativen Kosten hingegen werden über 2% eingespart. Diese Einsparungen entstehen hauptsächlich durch die verminderten Wartezeiten und die dadurch bessere Auslastung der Fahrzeuge. Die Kosten für Ein- und Ausrückfahrten sinken ebenfalls. Die Kosten für Leerfahrten steigen jedoch - von 593 zu 2217, so dass hierdurch der positive Effekt der hohen Auslastung wieder reduziert wird.

	Anzahl Fahrten	Kosten für Ein-/Ausrücken		
		gesamt	normal	nicht definiert
München linienrein	2.016	3.657.624	1.705.426	1.952.198
München liniengemischt	1.414	1.781.037	1.107.458	673.579
% gespart:	29,86	51,31	35,06	65,50

Tabelle 7.8: München linienrein vs. liniengemischt, Ein-/Ausrückfahrten

München

Wie der Tabelle 7.6 zu entnehmen ist, zeigt die linienreine bzw. liniengemischte Planung bei der Münchener Instanz ein gänzlich anderes Bild. Die Einsparungspotentiale sind bei den variablen Kosten um ein vielfaches höher und bei der Anzahl eingesetzter Fahrzeuge lassen sich sogar fast 25 Prozent einsparen. Die ca. 10 Prozent der eingesparten variablen Kosten resultieren aus den reduzierten Kosten für Ein- und Ausrückfahrten. Die 136 eingesparten Fahrzeuge verursachen diese Kosten nicht. Trotzdem werden gerade bei den Ein- und Ausrückfahrten um die 50 Prozent der Kosten eingespart, was ca. dem doppelten der eingesparten Fahrzeuge entspricht. Folglich muss die Verteilung der Fahrzeuge auf die Depots bei zugelasenem Linienwechsel wesentlich effizienter sein.

Bei genauerer Betrachtung der beiden Fahrpläne kann jedoch festgestellt werden, dass es unter den Ein- und Ausrückfahrten einige wenige gibt, die unverhältnismäßig hohe Kosten verursachen. Es handelt sich hierbei um Depotausrück- und -eintrückfahrten, die durch die Verbindungsmatrix nicht explizit definiert sind, die aber für einen gültigen Fahrplan unbedingt nötig sind und daher als sehr lange Leerfahrten im Modell angenommen werden. Bei der liniengemischten Planung sind dies 10 Fahrten, bei linienreiner Planung 30. Bereinigt man nun den Fahrplan um diese 10 bzw. 30 Fahrten, so ergibt sich das Bild aus der Tabelle 7.8. Es ist deutlich zu erkennen, dass das Einsparungspotential gesunken ist. Es beträgt nun ca. 35 %. Die Anzahl der Fahrten ist jedoch auch um ca. 30 Prozent gesunken, von 2016 auf 1414, d.h. dass die Ein-/Ausrückfahrten an sich nicht wesentlich kostengünstiger ausgeführt werden, sondern verhältnismäßig gleich bleiben. Im Gegensatz zu den Kosten für Ein- und Ausrückfahrten steigen die Kosten für Leerfahrten um fast das siebenfache, ohne jedoch den positiven Effekt der Ein-/Ausrückfahrten aufheben zu können (siehe Tabelle 7.7). Insgesamt bedeutet dies, dass für die Erfüllung eines Fahrplan bei liniengemischter Planung weniger Fahrzeuge zum Einsatz kommen, diese aber wesentlich besser ausgelastet sind.

BVO

Das größte Einsparungspotential ergibt sich bei den Instanzen der BVO. Wie den Tabellen 7.9 und 7.10 zu entnehmen ist, können hier ca. 20% der variablen Kosten eingespart werden und über 40% der Fahrzeuge. Diese wirklich hohen Werte, die sich deutlich von denen der anderen beiden Instanzen abheben, sind maßgeblich auf die Struktur des Liniennetzes der BVO - relativ weite Entfernungen zwischen den Depots und den Haltestellen und damit lange Ein- und Ausrückfahrten - zurückzuführen. Außerdem haben diese Auswirkungen mit einer Besonderheit des Fahrplans zu tun. Der Fahrplan der BVO enthält viele Linien, die nur sogenannte

Unterstützungsfahrten ausführen. Diese Linien haben dann nur eine einzige Fahrt, was im linienreinen Modell zu Umläufen mit einer Ausrückfahrt, einer Servicefahrt und einer Einrückfahrt führt. Die dafür benötigten Fahrzeuge sind dann natürlich überhaupt nicht ausgelastet und alles andere als effizient eingesetzt.

	Kosten gesamt	Kosten variabel	Anzahl Fahrzeuge	Fahrzeug- kosten
BVO linienrein	789.773	214.773	243	575.000
BVO liniengemischt	510.955	172.955	142	338.000
% gespart:	35,30	19,47	41,56	41,22
Abs. gespart:	217.818	41.818	101	237.000

Tabelle 7.9: BVO linienrein vs. liniengemischt

Der Fahrplan der BVO enthält insgesamt 56 Linien, von denen 20 Linien 5 oder weniger Fahrten haben. Im Schnitt hat eine Linie der BVO ca. 22 Fahrten, ein relativ geringer Wert im Vergleich zu Halle mit ca. 108 und München mit ca. 147 Fahrten. Die hohen Einsparungspotentiale hängen also mit der Beschaffenheit des Liniennetzes und den Linien mit extrem wenigen Fahrten zusammen.

Um diesen Ansatz weiter zu prüfen, kann ein reduzierter Fahrplan untersucht werden. In diesen Fahrplan werden nur Linien aufgenommen, denen mindestens 20 Fahrten zugeordnet sind. Das Ergebnis der bereinigten Optimierung sieht dann wie in der Tabelle 7.11 aus. Es ist deutlich zu erkennen, dass sich die Einsparungspotentiale reduzieren, wenn nur Linien mit relativ vielen Fahrten in die Optimierung einbezogen werden. In der gewählten Modellierung sinkt das Einsparungspotenzial von 35,3 auf 26,5 Prozent. Bei einer höheren Anzahl von Fahrten pro Linie reduziert sich also das Einsparungspotential.

7.3 Fahrzeugtypgruppen

Das Konzept der Fahrzeugtypgruppen besteht darin, dass eine Fahrplanfahrt von Fahrzeugen verschiedener Typen bedient werden kann. Damit entstehen für jede

	Ausrück- fahrten	Einrück- fahrten	Service- fahrten	Leer- fahrten	Warte- zeit
BVO linienrein	40.582	44.932	95.523	13.927	19.807
BVO liniengemischt	21.154	20.526	95.523	22.615	13.136
% gespart:	47,87	54,32	0	-62,38	33,68
Abs. gespart:	19.428	24.406	0	-8.688	6.672

Tabelle 7.10: BVO linienrein vs. liniengemischt, detailliert

	Kosten gesamt	Kosten variabel	Anzahl Fahrzeuge	Fahrzeug- kosten
BVO linienrein	529.928	160.928	150	369.000
BVO liniengemischt	389.372	136.372	104	253.000
% gespart:	26,52	15,26	30,67	31,44
Abs. gespart:	140.556	24.556	46	116.000

Tabelle 7.11: BVO linienrein vs. liniengemischt, dicht

Servicefahrt mehr Zuordnungsmöglichkeiten als bei einer eindeutigen Fahrzeugtypvorgabe. Eine Fahrzeugtypgruppe beschreibt also eine Menge von Fahrzeugtypen, mit denen eine bestimmte Fahrplanfahrt/Linie bedient werden kann. Bei der Bildung der Fahrzeugumläufe ist demzufolge zu beachten, dass die Schnittmenge der Fahrzeugtypgruppen aller in einem Umlauf enthaltenen Fahrten nicht leer ist.

Wendet man dieses Konzept auf den schon modellierten Beispielfahrplan (vgl. Abschnitt 7.2), so könnte man z.B. alle Fahrten der Linie 1 einer Fahrzeugtypgruppe Gelenkbusse zuordnen. Für die Fahrten der Linie 2 sind nur Fahrzeuge der Gruppe Klimabusse zugelassen. Ändert man den o.g. Fahrplan entsprechend ab, so werden nun wieder 3 Fahrzeuge benötigt. Interessant wird die Untersuchung der Fahrzeugtypgruppen aber erst auf größeren Instanzen, wo ein Fahrzeug mehreren unterschiedlichen Fahrzeugtypgruppen angehören kann, so dass die Gruppen sich überschneiden. Würden im obigen Beispiel Fahrzeuge eines Typs Klima-Gelenkbus eingesetzt, so könnten wieder alle Fahrten von 2 Fahrzeugen bedient werden, da der klimatisierte Gelenkbus sowohl in die Fahrzeugtypgruppe Gelenkbusse als auch in die Fahrzeugtypgruppe Klimabusse einzuordnen wäre. Als Ergebnis würde der Fahrplan von einem Gelenkbus und einem Klimagelenkbus, dessen operative und fixe Kosten etwas höher liegen, bedient.

Im Bezug auf die Vorgabe der zulässigen Fahrzeugtypen für jede Fahrt hat der Planer grundsätzlich drei Möglichkeiten (s. auch Abbildung 7.2):

feste Vorgabe (1:1) - eine Fahrt darf nur mit einem fest vorgegebenen Fahrzeugtyp bedient werden,

Fahrzeugtypgruppen-Bildung - eine Fahrt darf mit Fahrzeugen einer bestimmten Teilmenge der Fahrzeugtypen (=Fahrzeugtypgruppe) bedient werden,

alle Fahrzeugtypen in einer Gruppe - eine Fahrt darf mit einem beliebigen Fahrzeugtyp bedient werden.

Durch die Zuordnung von Fahrzeugen zu Fahrzeugtypgruppen, kann der Planer im Vergleich zu der ersten Variante (feste Vorgabe) die Fahrzeuge stärker auslasten, die die erforderlichen Eigenschaften (z.B. Klimaanlage) haben und dabei die

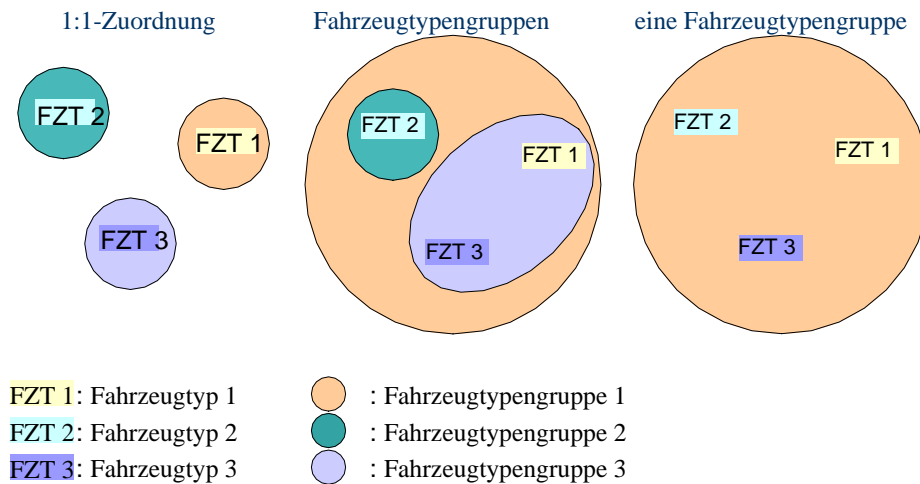


Abbildung 7.2: Modellierung der Fahrzeugtypgruppen als Mengen

geringsten Kosten verursachen. In der Variante „alle Fahrzeugtypen in einer Gruppe“ werden immer die günstigsten Fahrzeuge für die Fahrten gewählt. Dies kann aber zu unzulässigen Umlaufplänen führen.

Die beiden Extremfälle - „feste Vorgabe“ und „alle Fahrzeugtypen in einer Gruppe“ werden nachfolgend an denselben Instanzen aus der Praxis untersucht, für die in diesem Kapitel schon der Vergleich zwischen linienreinen und liniengemischten Umaluferstellung durchgeführt wurde. Im Vergleich zum Normalfall (mit Fahrzeugtypgruppen) ist zu erwarten, dass der Bedarf an Fahrzeugen bei der dritten Variante sinkt und bei der ersten Variante steigt.

7.3.1 Testergebnisse

Alle drei Fahrpläne können jeweils mit unterschiedlichen Freiheitsgraden bei den Fahrzeugtypgruppen gelöst werden. Am striktesten ist die 1:1-Zuordnung der Fahrzeugtypen zu den Gruppen. Werden Gruppen verwendet, die mehrere Fahrzeugtypen enthalten, so sind die Freiheitsgrade höher. Bei der letzten Variante wird ganz auf Typen und Gruppen verzichtet bzw. es gibt nur eine Gruppe mit allen Typen.

Halle

Bei dem Modell mit einer 1:1-Zuordnung entstehen im Vergleich zum Originalmodell, das mit sich überschneidenden Gruppen arbeitet, nur höhere Kosten (s. Tabelle 7.12). Die Anzahl der Fahrzeuge bleibt konstant. Ein weiterer Vergleich wird in der Tabelle 7.13 zum Modell mit nur einer Gruppe gemacht.

	Kosten gesamt	Kosten variabel	Fahrzeuge	Fahrzeugkosten
Halle 1 zu 1	376.629	145.629	115	231.000
Halle original	332.138	131.138	115	201.000
% gespart:	11,81	9,95	0	12,99
Abs. gespart:	44.491	14.491	0	30.000

Tabelle 7.12: Halle: feste Vorgabe

	Kosten gesamt	Kosten variabel	Fahrzeuge	Fahrzeugkosten
Halle original	332.138	131.138	115	201.000
Halle 1 FTG	469.658	127.658	114	342.000
% gespart:	-41,40	2,65	0,87	-70,15
Abs. gespart:	-137.520	3.480	1	-141.000

Tabelle 7.13: Halle: alle Fahrzeugtypen in einer Gruppe

Der wichtigste Wert in der obigen Tabelle ist die Anzahl der Fahrzeuge. Diese reduziert sich nur um ein Fahrzeug von 115 auf 114, was nicht einmal einem Prozent entspricht. Die Frage, die sich nun stellt, ist, warum nur eine so geringe Reduktion der Fahrzeuge erreicht wurde. Eine Möglichkeit, hier Informationen zu gewinnen, ist es, die Auslastung der Fahrzeuge zu betrachten. Eine Antwort kann unter anderem aus den Wartezeiten hergeleitet werden. Bei den unterschiedlichen Modellen des Haller Fahrplans schwankt der Anteil der Kosten für Wartezeiten zwischen 9,77% - Modell mit einer Fahrzeugtypgruppe - und 10,94% - Modell mit 1:1-Zuordnung. Der Einfluss der Fahrzeugtypgruppen hängt also anscheinend stark von der Auslastung der Fahrzeuge im originalen Fahrplan zusammen. Ist diese hier schon recht hoch oder lässt sich nicht entscheidend erhöhen, so ist das Optimierungspotential gering.

BVO

Die Instanz der BVO enthält 4 Fahrzeugtypgruppen, die sich im Wesentlichen an den Fahrzeugtypen orientieren. Für jede der drei Fahrzeugtypen existiert eine Fahrzeugtypgruppe, die nur diesen Fahrzeugtyp enthält. Zusätzlich gibt es eine Gruppe, die alle Typen enthält. Die Fahrten verteilen sich wie in der Tabelle rechts auf die unterschiedlichen Gruppen. Diese Zuordnung entspricht ca. einer 1:1-Zuordnung wie schon beschrieben. Sie wird deshalb auch bei der Analyse nicht verändert und das Modell unverändert untersucht. Als Modell mit mehr Freiheitsgraden werden die Gruppen nur anders zugeordnet. Fahrten, für die ein Minibus ausreicht, könnten auch von einem größeren Fahrzeug bedient werden, z.B. einem Midibus oder einem Gelenkbus. Für Midibusse können Standardgelenkbusse eingesetzt werden. Dies führt dann zu

Gruppe	Fahrten
Gelenkbus	582
Minibus	598
Midibus	36
Alle	3

genau drei Fahrzeugtypgruppen, wie in der Tabelle und Bild 7.3. Normalerweise - siehe Modellierung mit 4 Fahrzeugtypgruppen - werden für die Fahrten der BVO 142 Fahrzeuge benötigt, bei einer Modellierung mit 3 Fahrzeugtypgruppen, wie oben beschrieben, werden 135 Fahrzeuge benötigt. Wird der Umlaufplan nach der Variante drei (alle Fahrzeugtypen in einer Gruppe) gebildet, reduziert sich diese Anzahl ebenfalls um 7 auf 135 Fahrzeuge.

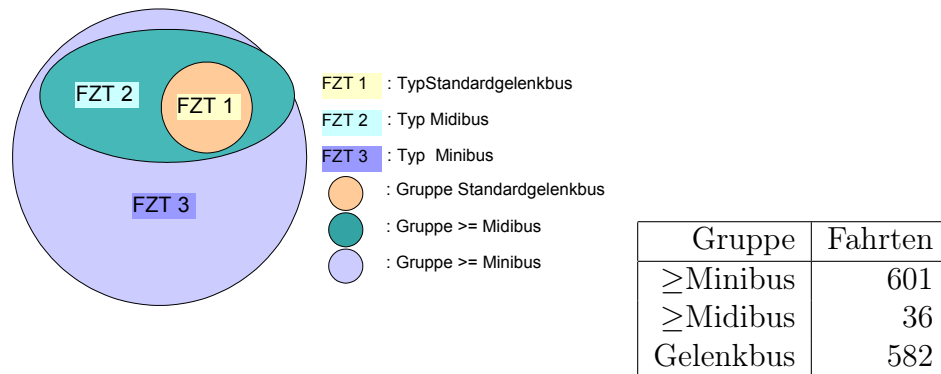


Abbildung 7.3: Gruppenbildung für BVO

München

Modelliert man die originalen Münchner Instanzen ohne Fahrzeugtypgruppen, also mit der durchschnittlichen Gruppengröße gleich der Anzahl Schichten ($12 \times 18 = 216$), so lässt sich der entstehende Fahrplan nicht mehr exakt optimal lösen. Die Optimierung bricht reproduzierbar mit einer Out-Of-Memory-Exception ab. Um dennoch die Münchener Instanz untersuchen zu können, musste die Komplexität des Modells reduziert werden. Dies kann z.B. durch einen Wegfall der meisten Depots geschehen und durch eine Reduktion der durchzuführenden Fahrten. Wird die Anzahl der Depots auf eins reduziert und die Anzahl der Fahrten auf 10.000, so lässt sich die Untersuchung problemlos durchführen. Die Münchener Instanzen enthalten 12 Fahrzeugtypgruppen und 12 Fahrzeugtypen. Die Zuordnung ist bijektiv. Wird das wie oben beschrieben reduzierte Modell gelöst, werden für den verbleibenden Fahrplan 403 Fahrzeuge benötigt. Wird hingegen nur eine Fahrzeugtypgruppe verwendet, so reduziert sich die Anzahl der benötigten Fahrzeuge auf 365. Dies entspricht einem Wert von 11 Prozent.

Um die Auswirkungen der Fahrzeugtypgruppen weiter zu untersuchen, können die Anzahl der Fahrzeugtypgruppen reduziert und die Fahrzeuge diesen Gruppen neu zugeordnet werden. Die drei modellierten Gruppen sind die

Gruppe	Fahrten
Kleinbus	1232
Niederflur	2495
Behindertengerecht	7382

Gruppe Kleinbus, die alle Fahrzeuge enthält, die Gruppe Niederflur, die alle Fahrzeugtypen mit dem Merkmal Niederflur enthält und die Gruppe Behindertengerecht, die alle Fahrzeugtypen enthält, die als behindertengerecht gekennzeichnet sind. Die Gruppen Niederflur und Behinderten gerecht sind disjunkt (s. Tabelle rechts). Bei einer Optimierung dieses Modells sinkt die Anzahl der benötigten Fahrzeuge entgegen der Modellierung mit 12 Fahrzeugtypgruppen auf 367, was 10 Prozent entspricht.

7.3.2 Zusammenfassung

Die oben gemachten Untersuchungen der Fahrzeugtypgruppen weichen zum Teil stark voneinander ab. Trotzdem macht eine Untersuchung ohne Fahrzeugtypgruppen durchaus Sinn: Durch einen Vergleich der beiden Modelle - mit Fahrzeugtypgruppen und ohne - kann z.B. sehr einfach festgestellt werden, wie stark die Anforderungen an die Fahrzeuge den Bedarf an Fahrzeugen erhöhen. So können ggf. weitere Optimierungspotentiale - z.B. der Einsatz flexibler einsetzbarer Fahrzeuge - untersucht werden. Die Anzahl Fahrzeuge des Modells ohne Fahrzeugtypgruppen ist der optimal erreichbare Wert. Alle anderen Modelle lassen sich hiermit vergleichen. Es kann auch als Basis für sog. "What-If Analysen" dienen. Im Allgemeinen reduziert sich die Anzahl der Fahrzeuge mit steigenden Freiheitsgraden.

7.4 Mehrdepotplanung

Auch wenn ein Unternehmen über mehrere Abstellorte zum Parken von Fahrzeugen verfügt, müssen am Ende eines Arbeitstages an einem Abstellort so viele Fahrzeuge eines Types abgestellt werden, wie morgens aus diesem Depot ausgerückt sind. Ohne diese Regel könnten die Umläufe am nächsten Tag u.U. nicht bedient werden, da die Fahrzeuge sich an einem falschen Ort befinden.

Es gibt jedoch unterschiedliche Modalitäten in der Rückkehr der Fahrzeuge, die nachfolgend im Bezug auf Kosteneinsparungen untersucht werden sollen. Die Rückkehr eines Fahrzeugs ist entweder nur zum eigenen Depot oder auch zu einem Fremddepot möglich:

- nur Eigendepot - d.h. am Ende des Tages stehen in einem Depot *dieselben Fahrzeuge*, die im Laufe des Tages aus diesem Depot ausgerückt sind,
- Depotgruppe - nicht nur Eigendepot, d.h. am Ende des Tages stehen in einem Depot *die gleichen Fahrzeuge* (vom gleichen Fahrzeugtyp) wie die, die im Laufe des Tages aus diesem Depot ausgerückt sind.

	Kosten gesamt	Kosten variabel	Fahrzeuge	Fahrzeugkosten
1. Depot	4.323	2.323	2	2.000
2. Depots	3.495	1.495	2	2.000
% gespart:	19,14	35,65	0	0
Abs. gespart:	828	828	0	0

Tabelle 7.14: Beispiel: zwei Depots

7.4.1 Klassische Mehrdepotplanung

Die Umlaufplanungsaufgabe eines ÖPNV-Unternehmens kann als ein Eindepot- oder eine Mehrdepot-Umlaufplanungsproblem formuliert werden. Besitzt das Unternehmen nur ein Abstellort für die Fahrzeuge oder ist für jede Fahrt fest vorgegeben, von welchem Depot aus sie ausgeführt werden darf, so handelt es sich um die Eindepot-Umlaufplanung. Kommen für einen Fahrzeugumlauf mehrere Depots in Frage, handelt es sich um eine Mehrdepot-Problemstellung. Das folgende Beispiel zeigt, wie sich Kosten in einem Netzwerk durch Einführen eines weiteren Depots verringern lassen. Dabei werden die Kosten, die durch das zusätzliche Depot für den Verkehrsbetrieb verursacht werden, nicht mit in das Modell aufgenommen. Sie können jedoch leicht in die späteren Berechnungen als zusätzliche Fixkosten mit einbezogen werden. Im bisher verwendeten Beispiel werden alle Fahrten des Fahrplans von einem Depot (PB-HBF) aus bedient. Dabei können alle Fahrten durch zwei parallel eingesetzte Fahrzeuge bedient werden. Diese Fahrzeuge starten beide aus dem einzigen Depot. Ein Fahrzeug muss aber erst eine Leerfahrt durchführen, um zu der Starthaltestelle (BI-SE) der ersten Servicefahrt zu gelangen. Diese Fahrt ist ca. 50 km lang. Könnte dieses Fahrzeug hingegen direkt aus BI-SE starten, so würde die Leerfahrt entfallen. Durch mehrere Depots können also operative Kosten eingespart werden. Definiert man nun die Haltestelle BI-SE ebenfalls als Depot, so reduzieren sich die operativen Kosten wie in der Tabelle 7.14 angegeben.

Die Frage, ob auch die Anzahl der einzusetzenden Fahrzeuge verringert werden kann und damit die Fixkosten, ist hingegen etwas schwieriger zu beantworten. In den untersuchten Beispielen werden nur bei den Münchener Instanzen Fahrzeuge eingespart. Bei den anderen bleibt die Fahrzeuganzahl konstant.

Um den Einfluss der Mehrdepotplanung zu verdeutlichen, wird angenommen, dass die Depotzuordnung von einer Fahrplanfahrt eindeutig ist, also handelt es sich bei der Modellierung und Lösung um ein Eindepot-Umlaufplanungsproblem. Für die untersuchten Instanzen wurde es dadurch realisiert, dass nur ein Depot für die Umläufe zugelassen wurde. Die Ergebnisse werden mit denen der vollständigen Instanz mit mehreren Depots verglichen.

	Kosten gesamt	Kosten variabel	Fahrzeuge	Fahrzeugkosten
Halle Depot A	336.577	135.577	115	201.000
Halle Depot B	344.498	143.498	115	201.000
Halle original	332.138	131.138	115	201.000
% gespart A:	1,32	3,27	0	0
Abs. gespart A:	4.439	4.439	0	0
% gespart B:	3,59	8,61	0	0
Abs. gespart B:	12.360	12.360	0	0

Tabelle 7.15: Halle: Eindepot- versus Mehrdepot-Umlaufplanung

	Ausrück- fahrten	Einrück- fahrten	Service- fahrten	Leer- fahrten	Warte- zeit
Halle Depot A	7.444	7.825	102.488	1.628	16.192
Halle Depot B	10.859	10.616	102.488	2.109	17.425
Halle original	5.936	6.156	102.488	2.217	14.342
% gespart A:	20,26	21,33	0	-36,13	11,43
Abs. gespart A:	1.508	1.669	0	-588	1.850
% gespart B:	45,34	42,01	0	-5,10	17,69
Abs. gespart B:	4.923	4.460	0	-108	3.083

Tabelle 7.16: Halle: Eindepot- versus Mehrdepot-Umlaufplanung detailliert

Halle

Die originale Haller Instanz enthält zwei Depots. Sie ist als *Halle normal* in den Tabellen unten gekennzeichnet. Verglichen wird der normale Umlaufplan (mit zwei Depots) mit Umlaufplänen, die jeweils auf Modellen mit nur einem Depot basieren. In der Tabelle 7.15 ist deutlich zu erkennen, dass sich durch die Verwendung von zwei Depots signifikante Einsparungen ergeben. Diese weichen bei den beiden untersuchten Möglichkeiten jedoch stark voneinander ab. Geht man von Depot A aus und vergleicht dann mit der Lösung für beide Depots, so ergibt sich ein Einsparungspotenzial von 1,32 %. Dies ist nicht einmal die Hälfte des Einsparungspotenzials, das sich ergeben würde, wenn Depot B betrachtet wird. Die Wahl des Depots spielt also eine wichtige Rolle beim Optimierungspotential der Mehrdepot-Planung. Ebenfalls interessant ist, dass durch die Verdopplung der Depots von eins auf zwei keine Fahrzeuge eingespart werden, sondern nur die operativen Kosten.

Die detaillierte Kostentabelle 7.16 zeigt, wo genau die variablen Kosten eingespart werden. Es sind maßgeblich die Ein- und Ausrückfahrten, die die Kosten reduzieren. Dies entspricht genau der Beschreibung aus dem einleitenden Beispiel. Es fällt auf, dass sich die Kosten für Leerfahrten bei beiden Analysen erhöht haben und die Kosten für Wartezeiten gesunken sind.

Depot Nr	Kosten gesamt	Kosten variabel	Fahrzeuge	Fahrzeugkosten
7000	553.708	215.708	142	338.000
7177	611.453	273.453	142	338.000
7500	576.041	238.041	142	338.000
8042	599.009	261.009	142	338.000
8130	664.134	326.134	142	338.000
Alle	510.956	172.955	142	338.000

Tabelle 7.17: BVO: Eindepot-Umlaufplanung

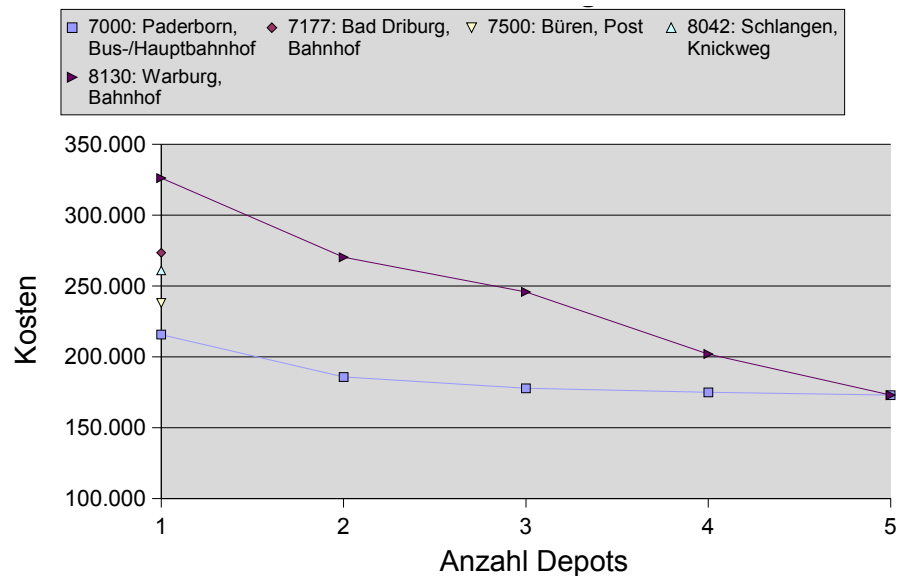


Abbildung 7.4: BVO: Kostenentwicklung

BVO

Da sich Instanzen, die auf dem Fahrplan der BVO basieren, relativ schnell lösen lassen, bietet sich dieser Fahrplan für eine weitere Untersuchung an. Untersucht werden zuerst die Kosten, die bei Verwendung eines Depots entstehen im Vergleich zu den Kosten, die bei Verwendung aller fünf Depots anfallen. Im Anschluss wird untersucht, wie sich die Kosten verändern, wenn man die Anzahl der Depots Stück für Stück erhöht. Dazu werden 2 Fälle betrachtet. Im ersten Fall wird immer von den minimalen Kosten der aktuellen Depotkombination ausgegangen und dann das Depot hinzugenommen, dass zur größten Reduktion der Kosten führt, im 2. Fall wird genau entgegengesetzt vorgegangen.

Für die oben beschriebenen weiteren Betrachtungen werden dann die Depots 7000 - als das mit den geringsten Kosten - und das Depot 8130 - das mit den höchsten Kosten ausgewählt. Zum Depot 7000 können dann nacheinander die Depots

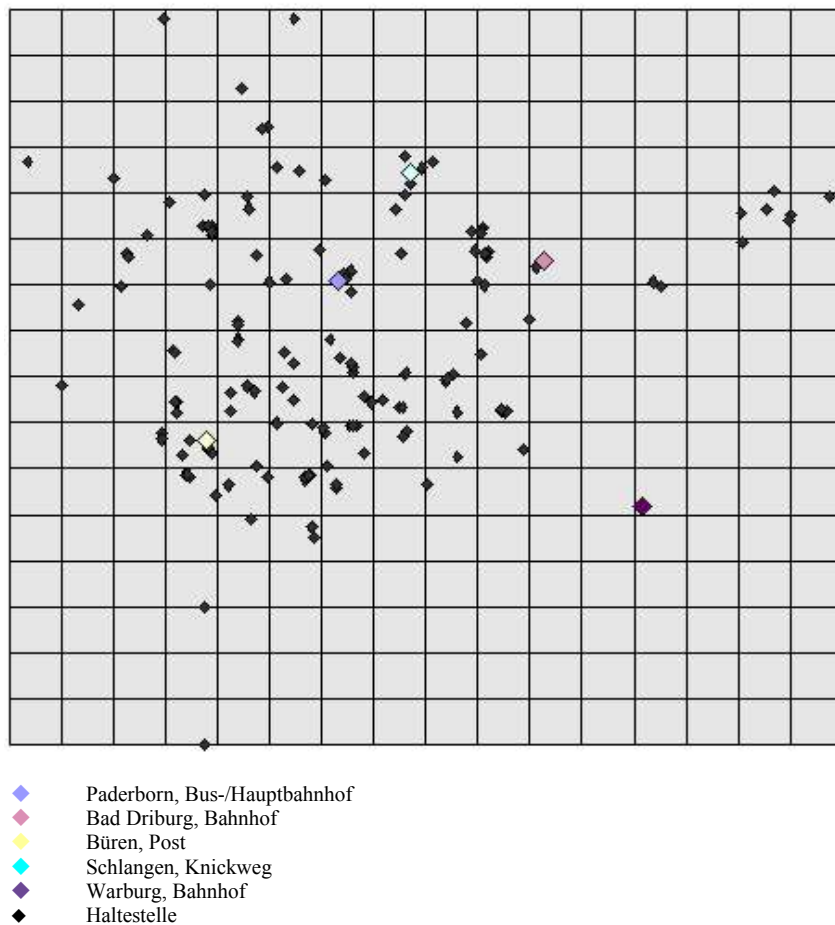


Abbildung 7.5: BVO: Verteilung Depots

7500, 7177, 8042 und 8130 hinzugefügt und der Fahrplan dann optimiert werden. Zum Depot 8130 werden die restlichen Depots in der Reihenfolge 7177, 8042, 7000, 7500 hinzugefügt. Die Ergebnisse dieser Untersuchung sind in der Tabelle 7.17 und in der Abbildung 7.4 zusammengefasst.

Interessant ist an dieser Stelle eine Betrachtung der Lage der Depots, die zu den guten - kostengünstigen - Ergebnissen führen. Die Grafik in der Abbildung 7.5 verdeutlicht, dass die Depots, die relativ zentral liegen, die besten Ergebnisse liefern. Zusätzlich hat natürlich auch die Anzahl der Fahrplanfahrten, die in der Nähe des Depots beginnen bzw. enden einen Einfluss auf die Vorteilhaftigkeit eines bestimmten Standortes für das Depot.

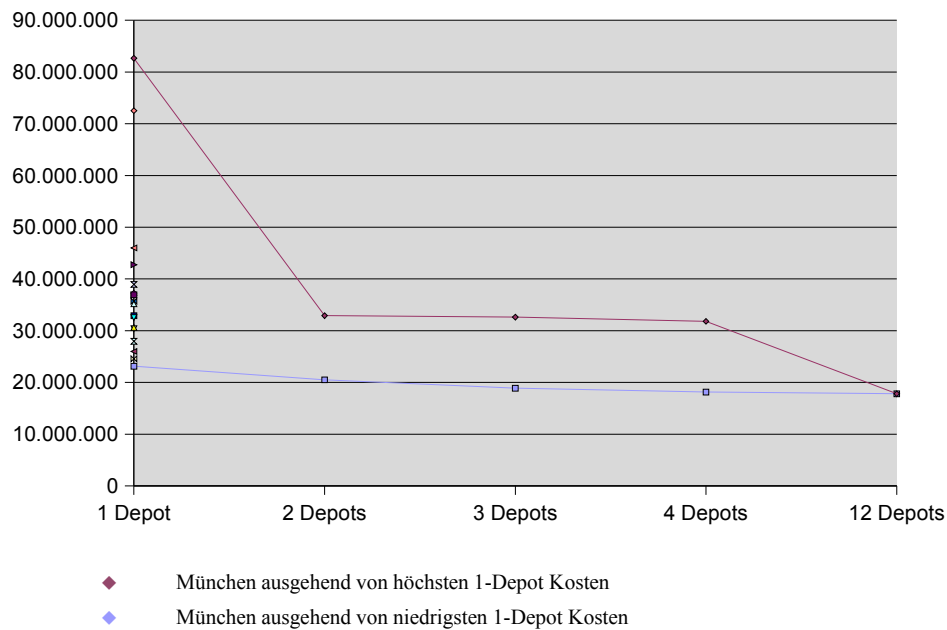


Abbildung 7.6: München: Variable Kosten bei unterschiedlicher Depotanzahl

München

Eine ähnliche Untersuchung lässt sich auch für den Münchener Fahrplan durchführen. Eine Untersuchung mit jeweils einem Depot führt hier zu stark unterschiedlichen Ergebnissen. Basierend auf dieser ersten Untersuchung kann dann die Anzahl der Depots schrittweise erhöht werden. Wie bei der Untersuchung der BVO ausgehend von 2 Fällen: Dem Depot mit dem geringsten Kosten und dem mit den höchsten. Diese Untersuchung ist in den Abbildungen 7.6 und 7.7 bis zu einem Wert von 12 Depots dargestellt.

7.4.2 Depotgruppen für gleitende Fahrzeug-Depot-Zuordnung

Gleitende Depotzuordnung wird möglich durch die Regelung, dass das Start- und Enddepot eines Umlaufs bzw. Umlaufstücks zwischen zwei Depotaufenthalten verschieden sein dürfen. Das wurde im Time-Space-Netzwerk Modell mit Hilfe des „*Depotgruppen-Konzepts*“ aus dem Kapitel 6 abgebildet. Durch diese Regelung können die Anzahl der Leerfahrten und „teures“ Warten außerhalb des Depots reduziert werden. Das Fahrzeug ist schneller wieder in einem Depot, verbraucht weniger Kraftstoff, und der Fahrer kann früher seinen Arbeitstag beenden, was zu geringeren Lohnkosten durch weniger Arbeitsstunden führt.

Damit aber am nächsten Tag alle Fahrten wieder so bedient werden können wie

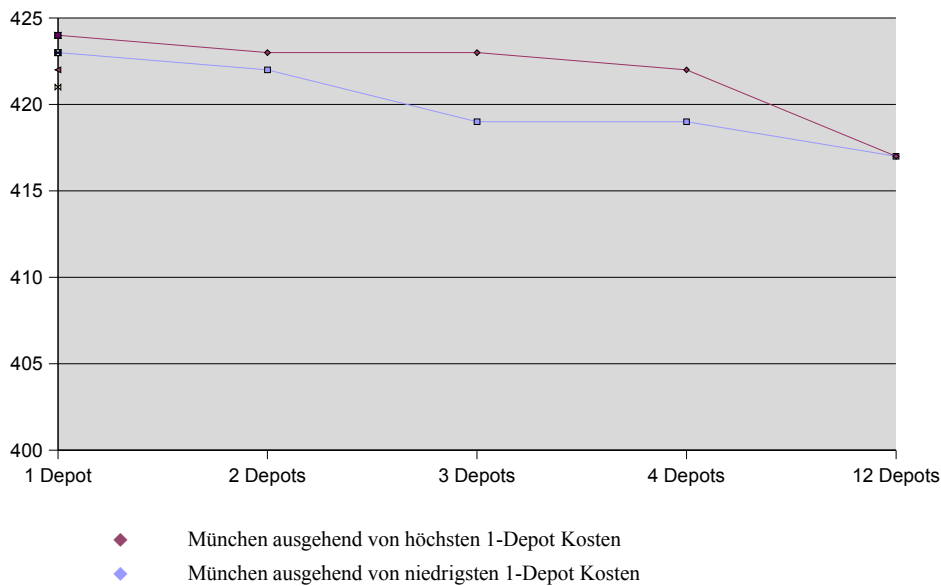


Abbildung 7.7: München: Anzahl Fahrzeuge bei unterschiedlicher Depotanzahl

am Vortag, muss eine weitere Restriktion bei der Modellierung beachtet werden: Nachdem alle Fahrzeuge in ein Depot zurückgekehrt sind, müssen sich in jedem Depot so viele Fahrzeuge eines Typs befinden, wie am Morgen desselben Tages. Nur so kann gewährleistet werden, dass der Fahrplan am nächsten Tag entsprechend abgearbeitet werden kann. Irgendwann wird ein Fahrzeug zu dem Depot zurückkehren, aus dem es bei der ersten Ausrückfahrt gestartet ist. Nach einer endlichen Anzahl von Tagen werden alle Fahrzeuge wieder in ihr ursprüngliches Startdepot zurückgekehrt sein.

Im bisher verwendeten Beispiel führt die Verwendung der gleitenden Fahrzeug-Depot-Zuordnung zu einer Reduzierung der operativen Kosten von 1.495 auf 1.235, was 17,4 % entspricht. Dieser relativ hohe Wert ist in der Praxis jedoch nicht zu erwarten, da dort das Verhältnis von Ein- und Ausrückfahrten zu den Service-, Leerfahrten und Wartezeiten anders ist als im Beispiel, wo 2 Linien nur 6 Fahrten zugeordnet sind.

Beim Optimieren der drei Fahrpläne sowohl mit fester als auch mit gleitender Depotzuordnung wurde bei allen Instanzen eine Reduzierung der Kosten festgestellt. Diese Reduzierung ist jedoch immer in einem relativ geringen Rahmen im Vergleich zu den anderen, vorher untersuchten Modellierungsmöglichkeiten. Die geringsten Einsparungen ergeben sich bei der Instanz aus Halle (s. Tabelle 7.18). Hier werden nur 0,15% der gesamten Kosten eingespart. Diese 0,15% resultieren aus den 0,38% Einsparungen bei den variablen Fahrzeugkosten. Es wurden keine Fahrzeuge eingespart.

	Kosten gesamt	Kosten variabel	Fahrzeuge	Fahrzeugkosten
Halle gl. Zuordn.	331.638	130.638	115	201.000
Halle feste Zuordn.	332.138	131.138	115	201.000
% gespart:	0,15	0,38	0	0
Abs. gespart:	500	500	0	0

Tabelle 7.18: Halle: gleitende Depotzuordnung

	Ausrück- fahrten	Einrück- fahrten	Service- fahrten	Leer- fahrten	Warte- zeit
Halle gl. Zuordn.	6.094	6.531	102.488	1.525	14.000
Halle feste Zuordn.	5.936	6.156	102.488	2.217	14.342
% gespart:	-2,66	-6,09	0	31,19	2,38
Abs. gespart:	-158	-375	0	691	342

Tabelle 7.19: Halle: gleitende Depotzuordnung, detailliert

Die detaillierten Ergebnisse in der Tabelle 7.19 zeigen, dass nicht, wie vermutet, bei den Kosten für Ein- und Ausrückfahrten gespart wird - diese Kosten sind sogar noch gestiegen, sondern bei den Kosten für Leerfahrten und Standzeiten.

Bei der Instanz der BVO ist das Einsparungspotenzial schon wesentlich höher (s. Tabelle 7.20). Es kann zwar kein Fahrzeug gespart werden, die variablen Kosten für den Betrieb der eingesetzten Fahrzeuge sinken aber um beachtliche 2,44%. Es ist zu vermuten, dass diese Einsparungen mit den weit auseinander liegenden Depots und den langen Distanzen der Servicefahrten zusammenhängen. Wie schon bei der Haller Instanz sinken hier die Kosten für Leerfahrten und Standzeiten, die Kosten für Ein- und Ausrückfahrten steigen dagegen verhältnismäßig gering an (s. Tabelle 7.21).

Bei der Münchener Instanz fällt das Einsparungspotenzial höher aus als bei der Haller Instanz, aber niedriger als bei der BVO. Wie die Tabelle 7.22 zeigt, können hier ca. 1% der variablen Kosten eingespart werden. Im Gegensatz zu den anderen beiden Fahrplänen kann ein Fahrzeug eingespart werden. Die in der Tabelle unten angegebene Einsparung von 0,29% für die Gesamtkosten kann natürlich nicht so gesehen werden. Die künstlich erhöhten Fahrzeugkosten machen die Einsparungen bei der Fahrzeuganzahl dominierend. Bereinigt um diese künstlichen Kosten ergibt

	Kosten gesamt	Kosten variabel	Fahrzeuge	Fahrzeugkosten
BVO gl. Zuordn.	506.731	168.731	142	338.000
BVO feste Zuordn.	510.955	172.955	142	338.000
% gespart:	0,83	2,44	0	0
Abs. gespart:	4.223	4.223	0	0

Tabelle 7.20: BVO: gleitende Depotzuordnung

	Ausrück- fahrten	Einrück- fahrten	Service- fahrten	Leer- fahrten	Warte- zeit
BVO gl. Zuordn.	22.452	24.704	95.523	16.083	9.969
BVO feste Zuordn.	21.154	20.526	95.523	22.615	13.136
% gespart:	-6	-20	0	29	24
Abs. gespart:	-1.297	-4.178	0	6.532	3.167

Tabelle 7.21: BVO: gleitende Depotzuordnung, detailliert

	Kosten gesamt	Kosten variabel	Anzahl Fahrzeuge	Fahrzeug- kosten
München gl. Zuordn.	1.548.791.723	17.621.723	416	1.531.170.000
München feste Zuordn.	1.553.241.447	17.811.447	417	1.535.430.000
% gespart:	0,29	1,07	0,24	0,28
Abs. gespart:	4.449.724	189.724	1	4.260.000

Tabelle 7.22: München: gleitende Depotzuordnung

	Ausrück- fahrten	Einrück- fahrten	Service- fahrten	Leer- fahrten	Warte- zeit
München gl. Zuordn.	881.811	984.974	13.176.113	263.428	2.315.397
München feste Zuordn.	837.839	943.857	13.176.113	395.758	2.457.880
% gespart:	-5,25	-4,36	0	33,44	5,80
Abs. gespart:	-43.972	-41.118	0	132.331	142.483

Tabelle 7.23: München: gleitende Depotzuordnung, detailliert

sich für München ein Einsparungspotenzial von ca. 1% der operativen Kosten. Bei den detaillierten variablen Kosten ergibt sich in der Tabelle 7.23 ein ähnliches Bild wie bei den Haller Instanzen: Die Kosten für Ein- und Ausrückfahrten steigen, die für Leerfahrten und Standzeiten sinken.

Das Liniennetz mittelgroßer und großer Städte ist dicht und komplex. Der Bedarf an Personenbeförderung ist groß, so dass selten alle Fahrten nur von einem Anbieter des ÖPNV bedient werden können.

Für ein Unternehmen kann es sinnvoll sein, in die eigene Planung auch fremde Depots anderer Partnerunternehmen einzubeziehen, um die Lösung des Umlaufplanungsproblems evtl. zu verbessern. Auch wenn ein Unternehmen den kompletten Bedarf an Personenbeförderung einer Stadt selbst deckt, kann eine Unterteilung aller Depots in Gruppen (z.B. nach Stadtteilen) zur Kostensenkung führen.

Durch die oben beschriebenen Untersuchungen kann festgestellt werden, dass sich eine gleitende Depotzuordnung durchaus positiv auf die Kosten eines Fahrzeugsatzplans auswirken kann. Der Umfang des Einsparungspotenzials schwankt bei den untersuchten Instanzen jedoch von 0,15% - also einem relativ kleinen Wert - bis zu ca. 1%. Beim Betrachten der Kostenstruktur fällt dann auf, dass die Kosten für Ein- und Ausrückfahrten zunehmen, die für Leerfahrten und Standzeiten sinken. Die geringen Einsparungen bei der gleitenden Depotzuordnung müssen natürlich auch den zusätzlichen Aufwänden, die für das Unternehmen und seine Mitarbeiter entstehen, entgegengestellt werden. So kann die Verwaltung eines Fuhrparks kostspieliger sein, wenn die Fahrzeuge jeden Abend in einem anderen Depot abgestellt werden. Ein anderer Aspekt bei der gleitenden Depotzuordnung ist auch die Frage, wie die Mitarbeiter, die die Fahrzeuge fahren, nach Beendigung der Fahrt wieder nach Hause kommen. PKW, die z.B. für die Fahrt zur Arbeit genutzt worden sind, sind noch in dem Depot, von dem aus der Fahrzeugumlauf gestartet wurde. Sind diese Probleme jedoch zu handhaben, bietet es sich an, die Fahrzeugumläufe mit gleitender Depotzuordnung zu modellieren, da sich Kosteneinsparungen ergeben, die keine besonders großen Umstellungsmaßnahmen im Unternehmen erfordern.

7.5 Zusammenfassung

Die durchgeführten Untersuchungen zeigen die unterschiedlichen Optimierungspotenziale, die sich durch verschiedene Maßnahmen bei der Ressourceneinsatzplanung ergeben. Eine wichtige Erkenntnis ist, dass der Erfolg der unterschiedlichen Maßnahmen maßgeblich von der Beschaffenheit des zu optimierenden Fahrzeugsatzplans abhängt. Wichtige Aspekte eines Fahrplans, die sich bei den Untersuchungen herausstellten, sind die Dichte des Fahrplans, was direkten Einfluss auf die Aus-

lastung der Fahrzeuge hat, die Anzahl der Fahrten pro Linie bzw. der Anteil an Linien mit sehr wenigen Fahrten, die durchschnittliche Streckenlänge und natürlich die Anzahl und die Lage der Depots.

Die in diesem Kapitel untersuchten Möglichkeiten einer Kostenreduktion lassen sich durch den Einsatz der in Kapitel 6 vorgestellten Modellierungsansätze realisieren. Allerdings führt eine solche Berücksichtigung fast immer zu einer weiteren Komplexitätserhöhung der ohnehin schwer zu lösenden Probleme. Die Testfälle aus der Praxis der Umlaufplanung erfordern oft eine Anwendung weiterführender Verfahren, die optimale oder fast optimale Lösungen für erweiterte Problemstellungen in annehmbaren Zeit berechnen. Solche Verfahren wurden ebenfalls im Rahmen dieser Arbeit entwickelt und werden im nächsten Kapitel beschrieben.

Kapitel 8

Lösungsansätze für sehr große Probleminstanzen

Die Lösungszeit für Mehrdepot-Umlaufplanungsprobleme wächst stark mit der Problemgröße, die durch die Anzahl Depots und Fahrzeugtypen, Anzahl Fahrten, Anzahl Haltestellen und Anzahl möglicher Verbindungsfahrten beeinflusst wird. Diese Aufzählung lässt erahnen, dass es mehrere Möglichkeiten geben kann, die Problemstellungen zu vereinfachen - natürlich abhängig davon, ob die Kostenoptimalität garantiert werden muss.

In diesem Kapitel werden Methoden beschrieben, die die Modellgröße reduzieren und dabei immer noch sehr gute bis optimale Lösungen des gleichen Problems in vertretbarer Zeit ermöglichen. Nachfolgend wird die Dekomposition der Aufgabenstellung in mehrere unabhängige Teilprobleme sowie einige Heuristiken wie eine Mustererkennung oder eine Kantenbewertung zur Reduktion der Modellgröße beschrieben.

8.1 Problemdekomposition im Preprocessing

Durch eine intelligente Analyse der Daten kann die Problemkomplexität reduziert werden. Dafür wurde im Rahmen dieser Arbeit ein Verfahren konzipiert und in der MDVS-Software integriert, das die vorliegenden Datenstrukturen hinsichtlich der möglichen Aufteilung in unabhängige Teilprobleme analysiert und die Daten gegebenenfalls entsprechend transformiert. Die Vorteile dieser Methode bestehen darin, dass weiterhin exakt optimale Lösungen berechnet werden und dass die Methode vollkommen transparent für den Benutzer ist. Dabei werden in einer Analyse der Fahrzeugtypen und -gruppenzuordnungen Teile des Fahrplans ermittelt, die unabhängig vom Rest verplant werden können, ohne dabei auf die Optimalität der

Lösung zu verzichten.

Fahrzeugtypen, die identisch in Bezug auf Kosten und Gruppenzuordnungen sind, dürfen und sollten zusammengefasst werden, da ihre Differenzierung keine zusätzlichen Freiheitsgrade für die Optimierung bietet und so nur zu einer unnötigen Komplexität in dem zu lösenden Modell führt. Lässt sich erkennen, dass einige Fahrzeugtypen in allen Kostenkomponenten teurer sind, als die anderen, deren Anzahl zudem auch noch unbegrenzt ist, empfiehlt es sich, die teureren Fahrzeugtypen von vornherein aus der Planung auszuschließen. Falls mehrere Fahrzeugtypen zusammengefasst werden, soll darauf geachtet werden, dass die Vorgaben für die minimal und maximal einzusetzende Anzahl von Fahrzeugen eines Fahrzeugtyps angepasst werden.

Die Methode startet mit einem Fahrplan, der durch Datenrelationen für Fahrten, Entfernungen, Depots, Fahrzeugtypen, Fahrzeugtypgruppen und Kapazitäten gegeben ist. Das Vorgehen der Methode sieht wie folgt aus:

1. Unabhängige Teilprobleme werden durch eine Analyse der Fahrzeugtypgruppen und entsprechende Aufteilung der Fahrten gebildet.
2. Unter Beachtung fahrzeugtypspezifischer Kostensätze und Kapazitäten werden Fahrzeugtypen ggf. wie folgt zusammengefasst:
 - (a) Feste Kombinationen von Fahrzeugtypen ermitteln, die immer zusammen als Bestandteil von Fahrzeugtypgruppen auftreten,
 - (b) kostendominierte Typen innerhalb solcher fester Kombinationen bestimmen,
 - (c) Kapazitäten analysieren, Fahrzeugtypen entsprechend zusammenfassen.
3. Nach der Optimierung unabhängiger Teilprobleme werden die Ergebnisse in einer Datei zusammengefasst. Dabei werden die Umläufe einheitlich nummeriert und die zusammengefassten Fahrzeugtypen den Kapazitäten entsprechend wieder aufgeteilt.

Durch das Lösen mehrerer kleiner Probleme statt eines großen, sinken insbesondere die Speicheranforderungen des Systems, so dass nunmehr Probleme gelöst werden können, die ohne Anwendung der Problemdekomposition aufgrund ihrer Größe nicht in den verfügbaren Hauptspeicher passten.

Diese Methode vereinfacht die größeren Problemstellungen, die aus mehreren unabhängigen Teilproblemen bestehen und liefert trotz der Vereinfachung exakt optimale Lösungen für das Gesamtproblem. Die im Weiteren vorgestellten Verfahren vereinfachen die Problemstellung ebenfalls, aber nicht mehr durch eine zulässige

Dekomposition, sondern durch das Verkleinern des Lösungsraumes. Dies kann dazu führen dass Lösungen berechnet werden, die nicht exakt optimal für die ursprüngliche Problemstellung sind. Die Verfahren sind aber so konzipiert, dass nur die Teile des Lösungsraums abgeschnitten werden, die mit hoher Wahrscheinlichkeit nicht optimal sind. Diese Verfahren werden allgemein heuristische Verfahren genannt.

8.2 Lambda-gesteuerte Reduktion der Verbindungskanten

Die Methode *Arc Reduction* besteht darin, dass nicht alle Leerfahrtsverknüpfungen betrachtet werden. Dies führt zu einer Verringerung der Modellgröße, dadurch, dass das mathematische Modell entsprechend weniger Entscheidungsvariablen besitzt. Dabei werden einige potenzielle Leerfahrtsverknüpfungen aus dem Netzwerkmodell ausgeschlossen, falls eine Über-Depot-Verbindung zwischen den betreffenden Fahrten möglich ist und falls die zusätzlichen Kosten einer solchen indirekten Verbindung nicht zu hoch sind. Dies wird durch einen Parameter λ gesteuert, der angibt, um wie viel teurer eine solche Fahrt über Depot sein darf, um sie anstelle einer direkten Fahrt nutzen zu dürfen. Die Idee basiert auf dem Gedanken, dass ein Aufenthalt eines Fahrzeugs im Depot für ein Unternehmen von Vorteil sein kann, ungeachtet geringfügig höherer operativer Kosten. Was unter „geringfügig“ verstanden wird, kann der Planer durch die entsprechende Parametereinstellung steuern.

Der Parameter λ bietet also eine Möglichkeit an, eine Rückkehr zum Depot zwecks Pause gegenüber einer direkten Verbindungsfahrt zwischen zwei kompatiblen Fahrgastfahrten mehr oder minder zu präferieren. Seien i und j zwei kompatible Fahrten. Seien K_{depot} die Kosten der Rückkehr zum Depot zwecks Pause und K_{dh} die Kosten einer direkten Verbindungsfahrt von i nach j . Eine direkte Verbindungsfahrt von i nach j wird nur dann in das Modell eingeführt, wenn

$$K_{depot} > (1 + \lambda) \cdot K_{dh}$$

Ist also $\lambda = 0$, werden die Kosten direkt verglichen. Bei $\lambda > 0$ wird die Rückkehr zum Depot zwecks Pause gegenüber der direkten Verbindungsfahrt präferiert. Somit werden nicht alle günstigeren direkten Verbindungsfahrten in das Netzwerk eingeführt (abhängig davon, wie groß λ ist), was die Problemgröße und die Laufzeit der Optimierung reduzieren kann. Allerdings kann in diesem Fall die optimale Lösung nicht garantiert werden. Die Tabelle 8.1 zeigt, wie sich die Problemgröße, Laufzeit und Kosten mit steigendem λ entwickeln (Fahrplan mit 1 Depot, 1 Fahrzeugtyp, 124 Haltestellen und 7068 Fahrgastfahrten).

λ	Anzahl			Laufzeit(Sek.)	Kosten
	Restriktionen	Variablen	Non-Zeros		
0	20.577	248.239	503.548	41,51	77.659,559
0,5	20.577	242.439	491.948	39,64	77.659,559
1	20.577	242.437	491.944	40,04	77.659,559
5	20.577	242.437	491.944	39,89	77.659,559

Tabelle 8.1: Änderung von λ

Es ist zu erkennen, dass bei dem vorgestellten Problem die Erhöhung von λ die Problemgröße reduziert. Dabei wurde aber in diesem Fall stets dieselbe optimale Lösung gefunden. Auf die Laufzeiten schien es in diesem Beispiel keine Auswirkungen zu haben. Trotzdem bietet es sich an, besonders bei den größeren Instanzen den Parameter λ höher zu setzen, da dadurch die Speicheranforderungen bei der Lösung des Problems reduziert werden.

8.3 Sequenzielles Scheduling und Clustering

Clustering First - Scheduling Second Es ist denkbar, die Linien oder die einzelnen Fahrten des Fahrplans zunächst auf die Depots des Unternehmens zu verteilen und erst dann die eigentlichen Fahrzeugumläufe zu bilden. Analog kann das Problem vereinfacht werden, indem konkrete Zuweisungen von Fahrzeugtypen vor der Bildung der Umläufe vorgenommen werden. Dies führt aber zu einer zu starken Einschränkung des Lösungsraums und somit zum „Abschneiden“ von guten und insbesondere von optimalen Lösungen. Derartiges Vorgehen garantiert nicht, dass die auf beschriebene Weise produzierten Lösungen mit einer minimalen Anzahl von Fahrzeugen auskommen, geschweige denn mit minimalen operativen Kosten.

Scheduling First - Clustering Second Das umgekehrte Vorgehen ist auch realisierbar. Dabei werden erst *anonyme* Umläufe gebildet, die in einem zweiten Durchlauf auf die Depots verteilt werden und/oder Fahrzeugtypen zugewiesen bekommen, wie beispielsweise in [Daduna and Mojsilovic, 1988] beschrieben. In seiner reinen Form ist dieses Vorgehen auch von einem Optimalen weit entfernt. Basierend auf der „Scheduling First - Clustering Second“-Idee wurde in dieser Arbeit ein heuristisches Verfahren entwickelt, das diesen Nachteil beseitigt.

8.4 Fahrtenketten-Heuristik: Fix-and-Optimize Methode

In diesem Abschnitt wird ein heuristisches Vorgehen für sehr große Probleminstanzen erarbeitet, das die Modellgröße durch das Fixieren der Fahrtenverknüpfungen verringert, und das auf diese Weise reduzierte Modell exakt optimal lösbar macht.

Das heuristische Vorgehen basiert auf der Idee, zuerst mehrere vereinfachte Probleme zu lösen, die beispielsweise als Eindepot-Probleme formuliert werden, und dann gemeinsame Ketten von Fahrten in allen unterschiedlichen Lösungen zu suchen. Eine Kette von Fahrten, die in allen Lösungen enthalten ist, wird als *stabile Kette* bezeichnet. Es wird angenommen, dass diese Kette mit hoher Wahrscheinlichkeit auch in der optimalen Lösung vorkommt. Eine stabile Kette wird in der nachfolgenden Optimierung des Gesamtproblems als eine Fahrt behandelt, so dass bei großer Anzahl von solchen Ketten die Modellgröße wesentlich reduziert werden kann. Da hierbei einige Verbindungen in der ersten Phase fixiert werden und dann eine exakte Optimierung durchgeführt wird, bezeichnen wir diesen Ansatz im Weiteren *fix-and-optimize* oder die *Zwei-Phasen-Heuristik*. Die Abbildung 8.1 illustriert das Vorgehen der Zwei-Phasen-Heuristik.

8.4.1 Phase 1

Das Ziel der Phase 1 ist es, mehrere unterschiedliche Umlaufpläne zu erhalten. Dafür werden zunächst, basierend auf dem ursprünglichen MDVSP, eine oder mehrere vereinfachte Problemstellungen konstruiert, die beispielsweise darin bestehen, dass nur ein Abstellort für die Fahrzeuge zugelassen wird. Dies können Ein-Depot-Probleme oder Probleme mit weniger Depot-Fahrzeugtyp-Kombinationen, als im ursprünglichen Problem sein. Das Lösen dieser einfacheren Probleme liefert mehrere Umlaufpläne. Dabei können für jede vereinfachte Instanz mehrere Umlaufpläne nach verschiedenen Dekompositionsstrategien, z.B. *LiFo* und *FiFo*, erzeugt werden, wie in Kapitel 5 beschrieben wurde. Damit können bei der Fahrtenketten-Erkennung noch mehr Umlaufpläne berücksichtigt werden, was bei gleicher Optimierungszeit in der ersten Phase der Heuristik zu einer besseren Qualität der Lösung in der zweiten Phase führt.

8.4.2 Phase 2

Das Ziel der Phase 2 ist es, aus den in der Phase 1 konstruierten Umlaufplänen Erkenntnisse über bestimmte Eigenschaften der optimalen Lösung des schwierigen ursprünglichen Problems zu gewinnen und diese für die Lösung der ursprünglichen

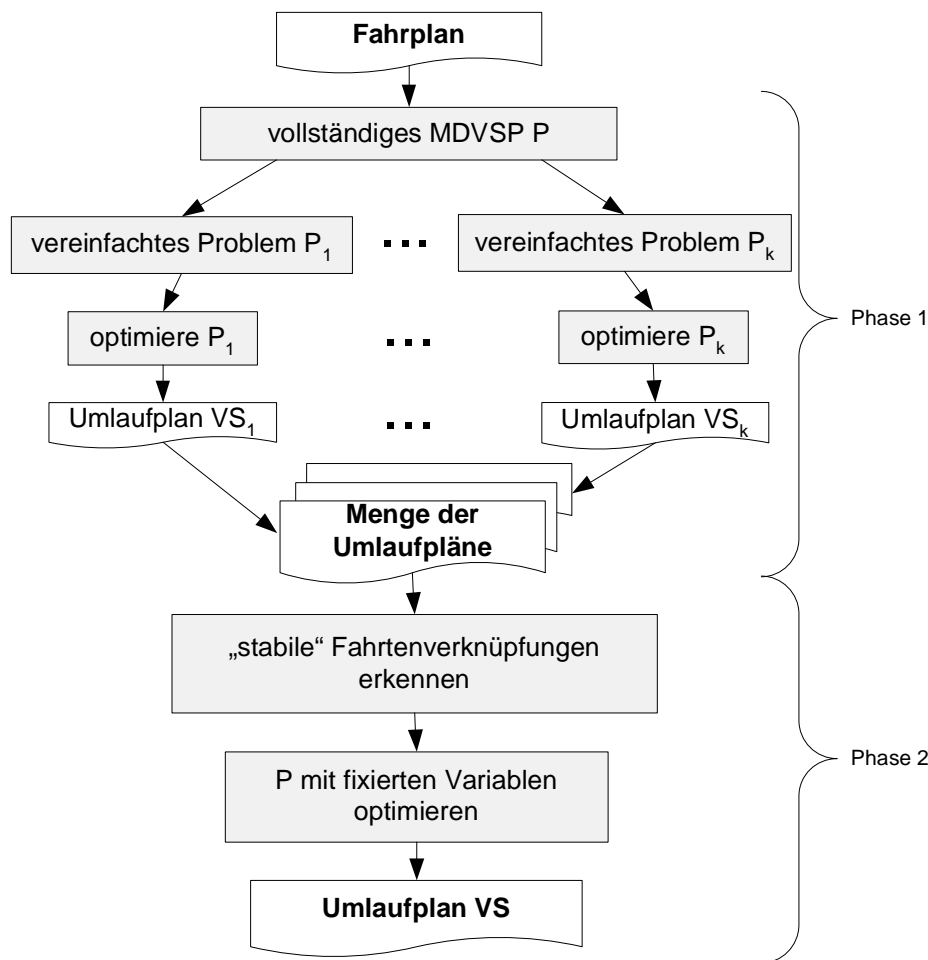


Abbildung 8.1: Arbeitsweise der „fix-and-optimize“ Zwei-Phasen Heuristik

Problemstellung zu benutzen.

Dafür werden die in Phase 1 produzierten Umlaufpläne auf das Vorhandensein der stabilen Ketten (gleichbleibende Fahrtenverknüpfungen) untersucht. Nach einer Fixierung solcher Verknüpfungen kann das verbleibende kleinere MDVSP mit Standard-Optimierungssoftware gelöst werden. Diese Fixierung geschieht durch das bilden fiktiver Fahrten, die eine solche Kette von Fahrten beinhalten. Selbstverständlich hängt es von den Daten ab, ob das resultierende Modell exakt optimal gelöst werden kann, wie jedoch später in diesem Abschnitt gezeigt wird, sind unsere Tests mit realistischen Testfällen sehr vielversprechend.

Die Voraussetzung zur Erkennung stabiler Fahrtenketten sind verschiedene Umlaufpläne, die vom Planungsexperten ausgewählt werden.

Servicefahrt-ID	Nachfolger-Servicefahrt-ID			
	Umlaufplan 1	Umlaufplan 2	Umlaufplan 3	Umlaufplan 4
1	4	4	5	4
→ 2	3	3	3	3
3	5	5	4	5
→ 4	6	6	6	6
5	8	10	9	9
→ 6	7	7	7	7
7	9	11	8	8
...

Tabelle 8.2: Zuordnungsfunktion für Nachfolger-Servicefahrten

Umlaufplan als Fahrtenzuordnungsfunktion

Jeder Servicefahrt kann eindeutig eine Nachfolger-Servicefahrt zugeordnet werden, außer sie ist die letzte Fahrt im Umlauf. Seien T die Menge aller Servicefahrten, und M die Menge aller Umlaufpläne, die diese Servicefahrten beinhalten. Außerdem sei für jeden zur Auswertung herangezogenen Umlaufplan $m \in M$ eine Zuordnungsfunktion $F_m : T \rightarrow T$ definiert, die für eine Servicefahrt ihre Nachfolger-Servicefahrt liefert. D.h., wenn $F_m(i) = j$, dann hat die Fahrt i im Umlaufplan m als Nachfolger die Fahrt j .

Damit kann folgender Prozess zur Erkennung stabiler Fahrtenketten initiiert werden. Falls $F_m(i) = j, \forall m \in M, i, j \in T$, dann bilden die Fahrten i und j eine stabile Fahrtenkette. Wird die nachfolgende Suche mit Fahrt j rekursiv fortgesetzt, können zur entdeckten Fahrtenkette neue Servicefahrten hinzugenommen werden. Sind alle Servicefahrten aus T untersucht, ist der Erkennungsprozess abgeschlossen.

In der Tabelle 8.2 ist die Zuordnungsfunktion F_m anhand eines Beispiels mit vier Umlaufplänen dargestellt. Bei der ersten Servicefahrt mit ID 1 unterscheidet sich ein Nachfolger von den anderen. So bildet sie allein eine triviale Fahrtenkette. Nachfolger der nächsten Servicefahrt mit ID 2 in allen 4 Umlaufplänen ist die Fahrt 3. Aus 2 und 3 wird eine Fahrtenkette „2-3“ erzeugt. Wären für die Servicefahrt mit ID 3 auch alle Nachfolger gleich, würde die Kette um diesen Nachfolger ergänzt. Anschließend wird aus den Servicefahrten mit ID 4, 6 und 7 die letzte Fahrtenkette „4-6-7“ gebildet.

Jede Fahrtenkette wird für die nachfolgende Modellierung als eine einzelne Fahrt dargestellt. Die Anfangszeit und die Abfahrtshaltestelle werden von der ersten Fahrt der Kette und die Ankunftszeit und die Endhaltestelle entsprechend von der letzten Fahrt der Kette übernommen, mehr dazu unten in Abschnitt Fahrplan-

```

Procedure Algorithmus zum Erkennen stabiler Fahrtenketten
 $S = \emptyset$ 
 $S_{all} = \emptyset$ 
for  $k = 1$  to  $K^1 = \text{Anzahl Umläufe im Umlaufplan 1}$ 
  for  $i = 1$  to  $I_k^1 = \text{Anzahl Fahrten im Umlauf } k \text{ des Umlaufplans 1}$ 
     $S = S \cup t_{k,i}^1$  //  $i$ -te Fahrt des  $k$ -ten Umlaufs des Umlaufplans 1
    zur aktuellen Fahrtenkette  $S$  hinzufügen
    for  $m = 2$  to  $M = \text{Anzahl der Umlaufpläne}$ 
      if  $F_m(t_{k,i}^1) \neq F_1(t_{k,i}^1)$  or  $i = I_k^1$  then
         $S_{all} := S_{all} \cup S$  // Fahrtenkette  $S$  ist abgeschlossen
        und wird gespeichert
         $S = \emptyset$  // neue Fahrtenkette anfangen
      break for
    endif
  endfor
endfor
endfor

```

Abbildung 8.2: Basis-Algorithmus zur Fahrtenkettenerkennung

konstruktion. Im nächsten Abschnitt wird ein Algorithmus präsentiert, der das vorgestellte Konzept realisiert.

Algorithmus zur Fahrtenkettenerkennung

Die Abbildung 8.2 zeigt den Algorithmus zur Erkennung stabiler Fahrtenketten in M unterschiedlichen Umlaufplänen, die durch das Lösen der einfachen Umlaufplanungsprobleme in der Phase 1 entstanden sind. Sei T die Menge aller Servicefahrten und S_{all} die Menge aller stabilen Fahrtenketten, die erkannt wurden. Weiterhin, sei $F_m : T \rightarrow T$ eine Zuordnungsfunktion, die für jede Servicefahrt $t_i \in T$ die Nachfolgerfahrt von t_i innerhalb des Umlaufplans m als $F_m(t_i)$ definiert. S ist die aktuell zu bildende Fahrtenkette, die im Trivialfall aus einer Servicefahrt besteht.

Die Fahrtenketten werden also entweder aus einer oder aus mehreren Servicefahrten gebildet. Die Entscheidung, wann eine Fahrtenkette zu Ende ist, erfolgt durch die Prüfung der Nachfolger. Wenn sich mindestens ein Nachfolger von allen anderen unterscheidet oder wenn die aktuelle Fahrt letzte Fahrt eines Umlaufs ist, wird die laufende Fahrtenkette abgeschlossen und in der Menge S_{all} gespeichert. Der Algorithmus terminiert, wenn alle Servicefahrten abgearbeitet sind. Danach erfolgt die Bildung eines neuen Fahrplans.

Fahrplankonstruktion

Ausgangsbasis für die Bildung eines neuen Fahrplans bietet die Fahrtenkettenliste S_{all} , die nach dem Durchlaufen des Algorithmus alle erkannten Fahrtenketten beinhaltet. In diesem neuen Fahrplan repräsentieren die gebildeten Fahrtenketten die neuen Servicefahrten. Die entsprechenden Fahrplanparameter für eine Fahrtenkette, wie Abfahrts-, Ankunftszeit und Anfangs- und Endhaltestelle, werden von der Anfangsfahrt bzw. der Endfahrt der Fahrtenkette übernommen.

Darüber hinaus muss noch über die Angehörigkeit zu den Fahrzeugtypgruppen für die einzelnen Fahrtenketten entschieden werden. Eine Fahrzeugtypgruppe besteht aus einem oder mehreren Fahrzeugtypen, die eine Servicefahrt oder in diesem Fall eine Fahrtenkette ausführen dürfen. Diese Entscheidung wird im Laufe der Fahrplanerstellung getroffen. Die Fahrzeugtypgruppen gewährleisten eine zulässige Verteilung der Fahrzeugtypen auf die Servicefahrten.

Bei der Zuordnung der Fahrzeugtypgruppen zu den Fahrtenketten wird wie folgt vorgegangen. Als erstes werden Fahrzeugtypgruppen aller Servicefahrten einer Fahrtenkette untersucht. Stellt sich dabei heraus, dass es sich um verschiedene Fahrzeugtypgruppen handelt, wird dann nach einer Fahrzeugtypgruppe gesucht, die nur Fahrzeugtypen beinhaltet, die in allen in der Fahrtenkette vorkommenden Fahrzeugtypgruppen existieren. Existiert so eine Fahrzeugtypgruppe nicht, wird eine neue Fahrzeugtypgruppe als Schnittmenge der in der Fahrtenkette vorkommenden Fahrzeugtypgruppen konstruiert. Die Schnittmenge der Fahrzeugtypgruppen aller in einer Fahrtenkette vorkommenden Fahrten ist nicht leer, da sonst die Kette nicht in den Umlaufplänen enthalten wäre. Diese Fahrzeugtypgruppe wird anschließend der Fahrt zugeordnet, die die Fahrtenkette im modifizierten Fahrplan darstellt.

Gegenüber dem Originalfahrplan wird der neue Fahrplan kleiner. Der Ausmaß der Komprimierung richtet sich nach der Anzahl der ausgewerteten Umlaufpläne. Nach dem Lösen des Umlaufplanungsproblem mit dem neuen Fahrplan aus Fahrtenketten basiert das Ergebnis (Umlaufplan) auf Fahrtenketten. Um zu einem zulässigen Umlaufplan zu kommen, der als Lösung für die ursprüngliche Problemstellung gelten soll, müssen die Fahrtenketten wieder aufgelöst und als einzelne Fahrten im endgültigen Umlaufplan ausgewiesen werden.

8.4.3 Numerische Ergebnisse

Der in diesem Abschnitt vorgestellte heuristische Ansatz wurde implementiert und in der Umlaufplanungssoftware integriert. Nachfolgend sollen die Ergebnisse der Testläufe vorgestellt werden.

Probleminstanzen und exakte Optimierung

Die Testläufe wurden auf zahlreichen unterschiedlich komplexen Testinstanzen durchgeführt, die basierend auf fünf realen Fahrplänen erzeugt sind.

city 1 Fahrplan eines regionalen ÖPNV-Unternehmens mit drei unterschiedlichen Fahrzeugtypen und 21 Endhaltestellen,

city 2 Fahrplan einer mittelgroßen deutschen Stadt mit homogener Fahrzeugflotte und 48 Endhaltestellen,

city 3 Fahrplan einer deutschen Großstadt mit homogener Fahrzeugflotte und 124 Endhaltestellen,

mun 1 Fahrplan von München mit 11062 Fahrten, 12 Fahrzeugtypen und 160 Endhaltestellen,

mun 2 ein weiterer Fahrplan von München mit 10710 Fahrten, 12 Fahrzeugtypen und 140 Endhaltestellen.

	# Fahrten	# Haltestellen	# Fahrzeugtypen
city 1	2047	21	3
city 2	3054	48	1
city 3	7068	124	1
mun 1	11062	160	12
mun 2	10710	140	12

Tabelle 8.3: Eigenschaften untersuchter Testfälle

Die wichtigsten Eigenschaften der Basis-Fahrpläne sind in der Tabelle 8.3 zusammengefasst. Die Komplexität der erzeugten Testinstanzen wurde durch das Hinzufügen oder Weglassen von Depots gesteuert. So wurden für city 1 Instanzen mit drei Fahrzeugtypen und bis zu neun Depots erzeugt, für city 2 - mit einem Fahrzeugtyp und bis zu acht Depots, für city 3 - mit wiederum einem Fahrzeugtyp und bis zu vier Depots. Auf diese Weise sind 18 Testinstanzen entstanden. Dazu kommen zwei Instanzen der Münchener Verkehrsgesellschaft - mun1 und mun2.

Um eine spätere Analyse der heuristischen Ergebnissen bezüglich der Laufzeit und der Lösungsqualität zu ermöglichen, wurden diese Instanzen zunächst mit dem exakten Verfahren gelöst. Wie der Tabelle 8.4 zu entnehmen ist, lag die Laufzeit des exakten Verfahrens für alle auf den Fahrplänen city 1 bis city 3 basierten MDVSP-Instanzen unter 10 Stunden, für mun 1 - etwas darüber.

	#Depots	#Schichten	$\varnothing G$	Zeilen	Spalten	Nonzeros	CPU(sec)
city 1	1	3	1,6	8799	31788	66866	8
	2	6	3,2	16218	67997	142574	114
	3	9	4,8	23314	102366	214602	695
	4	12	6,4	30260	136459	286078	1380
	5	15	8	36463	167875	352200	3396
	6	18	9,6	43661	204009	427758	8568
	7	21	11,2	50822	238227	499484	9228
	8	24	12,8	58020	253861	534042	14040
	9	27	14,4	65541	311422	652454	29040
city 2	1	1	1	9599	64283	131622	11
	2	2	2	16064	129384	264880	53
	3	3	3	22257	190342	389852	270
	4	4	4	28813	253663	519550	676
	5	5	5	35369	319311	653902	1338
	6	6	6	41888	393095	804526	5332
	7	7	7	48406	461587	944566	22200
	8	8	8	54959	515607	1055662	27720
city 3	1	1	1	20577	248239	503548	42
	2	2	2	34076	502373	1018886	289
	3	3	3	47586	749607	1520424	16026
	4	4	4	61097	941658	1911596	32400
mun 1	18	55	6,3	162858	1479848	3031173	39020
mun 2	3	15	7,1	159659	2032883	4141716	>100000

Tabelle 8.4: Ergebnisse der exakten Optimierung

In der Tabelle 8.4 ist jeweils die Anzahl der Depots sowie die Anzahl der Depot-Fahrzeugtypkombinationen angegeben, die im Modell berücksichtigt wurden. Die zweite Zahl entspricht auch der Anzahl der nichtleeren Schichten im Netzwerkmodell. Für city 1, beispielsweise, entspricht die Anzahl der Schichten dem Dreifachen der Anzahl von Depots. Dies bedeutet, dass in diesem Fall beliebige Kombinationen von Depots und Fahrzeugtypen erlaubt sind. Im Allgemeinen können einige Depot-Fahrzeugtypkombinationen gesperrt sein, je nach benutzerdefinierten Angaben. Dies ist in den Datensätzen mun 1 und mun 2 zu erkennen: von den $12 \times 18 = 216$ möglichen Depot-Fahrzeugtyp-Kombinationen sind im Fall mun 1 55 erlaubt, bzw. sind von $12 \times 3 = 36$ möglichen im Fall mun 2 nur 15 zugelassen.

Eine weitere Kennzahl für die Komplexität der jeweiligen Instanz ist die durchschnittliche Anzahl der erlaubten Depot-Fahrzeugtyp-Zuordnungen pro Fahrplanfahrt. Diese Kennzahl wird die *durchschnittliche Gruppengröße* bezeichnet und ist in der Spalte $\oslash G$ für jede Instanz angegeben und gibt an, in wie vielen Schichten eine Fahrplanfahrt durchschnittlich abgebildet wird.

Die Tabelle 8.4 beinhaltet außerdem die Angaben zu der Größe der MIP-Koeffizientenmatrix sowie die CPU-Zeit für die Berechnung der optimalen Lösung auf einem Intel Pentium mit 2,2 GHz. Die mathematischen Modelle wurden mit ILOG CPLEX 8.0 gelöst (s. [ILOG, 2004]).

Der Testfall mun 1 beinhaltet 18 Depots und 55 Schichten, wobei die durchschnittliche Gruppengröße 6,3 beträgt. Der Testfall mun 2 beinhaltet 3 Depots und 15 Schichten, mit einer durchschnittlichen Gruppengröße von 7,1 Schichten pro Fahrt. Wie in der Tabelle 8.4 gezeigt, konnte mun 1 nach ungefähr 11 Stunden exakt optimal gelöst werden. Die Optimierung für den Fall mun 2 wurde dagegen nach mehr als drei Tagen abgebrochen.

Anmerkung: Die Laufzeiten hier und in den folgenden Tabellen beinhalten das Einlesen der Daten, die Konstruktion des Netzwerks und des mathematischen Modells, das Lösen des letzteren, sowie Auswertung und Ausgabe der Umlaufpläne. Die darin enthaltenen Optimierungszeiten sind für die Optimierungsläufe ohne „Fine-Tuning“ der Optimierungsbibliothek, also mit Default-Einstellungen entstanden. Durch eine intelligentere Wahl des LP-Optimierungsverfahrens können die Laufzeiten in einigen Testfällen stark reduziert werden. Da der Gegenstand dieses Kapitels nicht das Optimierungsverfahren an sich ist, sondern vielmehr die Vergleiche zwischen exakten und heuristischen Verfahren, reichen an dieser Stelle die Default-Einstellungen aus.

An dieser Stelle möchten wir unterstreichen, dass die exakt optimale Lösung von Testfällen dieser Größe erst durch das Verwenden der Modellierung aus Kapitel 4 ermöglicht wurde.

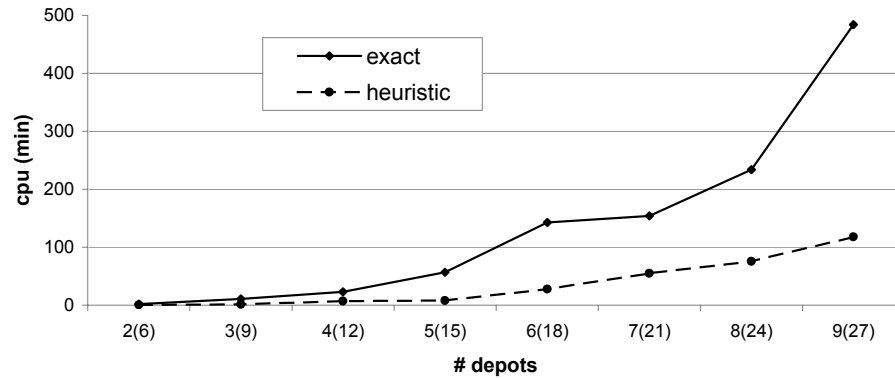


Abbildung 8.3: Laufzeiten versus Anzahl der Depots: city 1

Heuristische Ergebnisse: Laufzeitverhalten und Lösungsqualität

Als nächstes wurden dieselben Probleminstanzen mit Hilfe der Zwei-Phasen-Heuristik gelöst. Dabei wurde wie folgt vorgegangen: für eine n -Depot Problemstellung wurden zunächst n vereinfachte Probleme gelöst. Falls die ursprüngliche Problemstellung nur ein Fahrzeugtyp beinhaltet, handelt es sich bei den vereinfachten Problemen um Eindepot-Umlaufplanungsprobleme. Sind mehrere Fahrzeugtypen vorhanden, so werden die vereinfachten Probleme als Mehrdepot-Umlaufplanungsprobleme gebildet, beinhalten aber weniger Schichten als das ursprüngliche Problem und sind demzufolge viel schneller zu lösen.

Durch diese Art der Lösung war zu erwarten, dass die vereinfachte MIP-Formulierung viel weniger Entscheidungsvariablen enthält, was zu wesentlich reduzierten Laufzeiten führen sollte. Dieses Verhalten, sowie die möglichen heuristischen Verluste durch „Kettenbildung“, konnten in zahlreichen Testläufen untersucht werden. Die Testläufe haben eine signifikante Verbesserung der Laufzeit gegenüber dem exakten Verfahren gezeigt. Dabei weisen die berechneten Umlaufpläne eine sehr hohe Lösungsqualität auf. In den Fällen, für die der optimale Zielfunktionswert aus den Ergebnissen der exakten Läufe bekannt ist (alle außer mun2), liegt die Abweichung von dem optimalen Wert (nur operative Kosten) weit unter einem Prozent.

In Bezug auf die Anzahl der Fahrzeuge, bzw. fixe Fahrzeugkosten, sind die heuristischen Lösungen allesamt optimal. Aus praktischer Sicht sind Lösungen mit solchen Kostenwerten von den optimalen nicht zu unterscheiden. Für den regionalen Fahrplan city 1 liegt die Abweichung in den operativen Kosten zwischen 0,07 und 0,25 % für city 2 - zwischen 0,01 und 0,10 % und für city 3 - zwischen 0,001 and 0,0001 %. Als Optimierungsbibliothek wurde ILOG CPLEX 8.0 ([[ILOG, 2004](#)]) verwendet. Akzeptable Laufzeiten wurden auch bei der Benutzung des Optimierers MOPS ([[Suhl, 2000](#)]) für alle Instanzen, außer der größten gemessen.

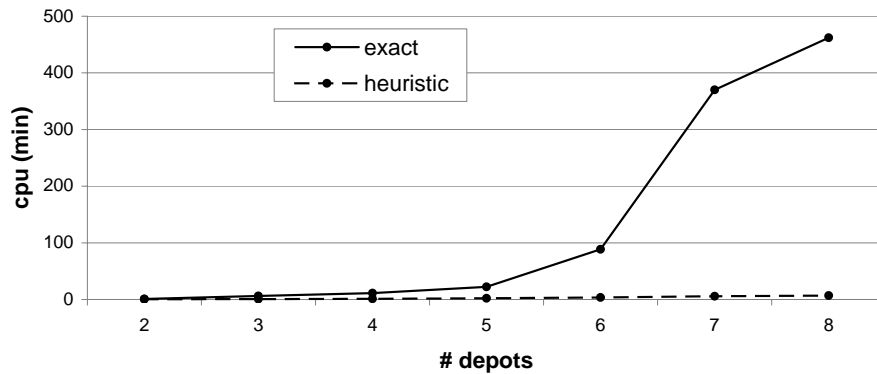


Abbildung 8.4: Laufzeiten versus Anzahl der Depots: city 2

Die Tabelle 8.5 und die Abbildung 8.3 vergleichen die Ergebnisse und die Laufzeiten der exakten und der heuristischen Testläufe für city 1. Die Zeile „#Depots“ enthält die Anzahl der Depots, die zwischen 2 und 9 variiert, und, in Klammern, die Anzahl der Schichten, die zwischen 6 und 27 variiert. Mit steigender Anzahl der Schichten wächst auch die Anzahl der nicht-fixierter Fahrtenverknüpfungen. Und diese Anzahl impliziert einen höheren Freiheitsgrad, was mit einer höheren Komplexität des Modells einhergeht. Die Ergebnisse zeigen, dass die Qualität der heuristischen Ergebnisse sehr hoch ist, während die Laufzeit zwischen 15 und 30% der exakten Optimierungszeit beträgt. Wie die Abbildung 8.3 zeigt, vergrößert sich die Einsparung in der Laufzeit mit wachsender Problemgröße.

#Depots	2(6)	3(9)	4(12)	5(15)	6(18)	7(21)	8(24)	9(27)
#Unfixierte Fahrten	967	1135	1232	1390	1408	1482	1498	1538
% von 2047	47,2	55,4	60,2	67,9	68,8	72,4	73,2	75,1
operative Kosten								
exakt	1027,7	935,4	925,9	789,0	747,8	685,1	647,2	643,3
heur.	1028,5	937,8	928,2	789,7	748,3	686,3	648,6	644,1
GAP(%)	0,07	0,25	0,25	0,09	0,07	0,19	0,22	0,14
CPU (min)								
exakt	1,9	10,9	23,0	56,6	142,8	153,0	234,0	484,0
heur.	0,5	1,5	6,9	8,2	27,5	54,8	75,8	118,0

Tabelle 8.5: Ergebnisse der Zwei-Phasen-Heuristik: city 1

Das Verhalten der Fix-and-Optimize Heuristik für den Testfall city2 ist in der Tabelle 8.6 und in der Abbildung 8.4 illustriert. In diesem Fall ist nur ein Fahrzeugtyp vorhanden, daher ist die Anzahl der Depot-Fahrzeugtyp-Kombinationen identisch mit der Anzahl der Depots und variiert zwischen 2 und 8. Selbst in dem Fall mit 8 Depots wurden etwa 2/3 der Fahrtenverknüpfungen fixiert, was eine

#Depots	2	3	4	5	6	7	8
#Unfixierte Fahrten	745	891	939	939	971	992	1004
% von 3054	24,4	29,2	30,7	30,7	31,8	32,5	32,9
operative Kosten							
exakt	55424,8	54705,8	54486,2	54486,2	54301,4	54325,1	54138,1
heur.	54428,6	54752,2	54535,1	54535,1	54357,7	54351,9	54184,6
GAP(%)	0,01	0,08	0,09	0,09	0,10	0,05	0,09
CPU (min)							
exakt	0,9	6,2	11,3	22,3	88,7	370,0	462,0
heur.	0,4	0,7	1,3	2,2	3,6	5,7	6,9

Tabelle 8.6: Ergebnisse der Zwei-Phasen-Heuristik: city 2

starke Verkürzung der Laufzeit bei der heuristischen Lösung implizieren soll. Die Laufzeit für die größeren Probleme beträgt beim heuristischen Vorgehen in der Tat zwischen einem und zwei Prozent der exakten Optimierungszeit.

Wie der Tabelle 8.7 und der Abbildung 8.5 zu entnehmen ist, zeigt der Testfall city 3 ein ähnliches Verhalten. Im Fall mit fünf Depots konnte das Problem nicht exakt optimal gelöst werden, wurde aber durch die Anwendung der Fix-and-Optimize Heuristik in etwas über eine Stunde gelöst.

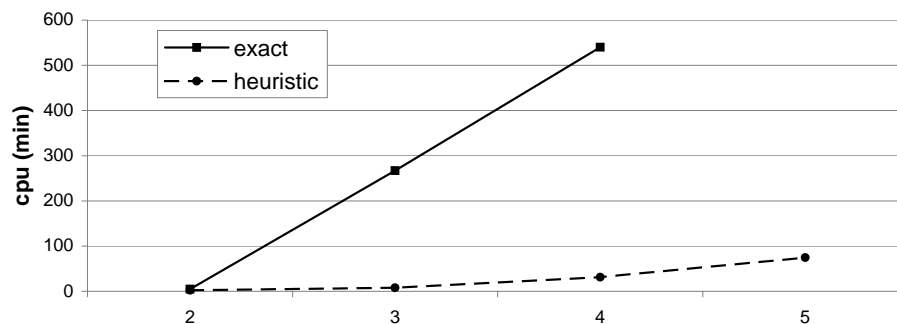


Abbildung 8.5: Laufzeiten versus Anzahl der Depots: city 3

Aus der praktischen Sicht ermöglicht die Fixierung der Fahrtenverknüpfungen durch die Bildung der Fahrtenketten beinahe optimale Lösungen für die extrem großen Problemstellungen, die mit der heutigen Technik und mit den aktuellen Versionen der Standard-Optimierungssoftware nicht in annehmbarer Zeit exakt optimal lösbar sind. Die in diesem Kapitel vorgeschlagene Heuristik basiert, genauso wie auch das in Kapitel 4 vorgeschlagene exakte Verfahren, auf den Methoden der mathematischen Optimierung, die in den Standard-Optimierungsbibliotheken enthalten sind. Daraus können wir schließen, dass sowohl das exakte Verfahren, als auch die Heuristik durch die zukünftig absehbaren Fortschritte in der Rechentech-

#Depots	2	3	4	5
#Unfixierte Fahrten	2269	2443	2630	2653
% von 7068	32,1	34,6	37,2	37,5
operative Kosten				
exakt	75922,0	75480,0	74930,3	—
heuristisch	75928,8	75481,8	74931,4	74839,2
GAP(%)	0,01	0,003	0,001	—
CPU (min)				
exakt	4,9	267,1	540,0	—
heuristisch	2,4	7,8	31,4	74,6

Tabelle 8.7: Ergebnisse der Zwei-Phasen-Heuristik: city 3

nik und in den Optimierungsbibliotheken schneller und leistungsfähiger werden. Das bedeutet, dass der Abstand zwischen den exakten und heuristischen Laufzeiten auch weiterhin erhalten bleibt, auch wenn das exakte Verfahren durch die absehbaren Entwicklungen schneller wird.

Um die Komplexität der beiden größten Testfälle mun 1 und mun 2 zu bewältigen, wurden beim Bilden der vereinfachten Modelle zusätzlich zum Weglassen einiger Depots (wie in den bisher betrachteten Fällen) auch einige Fahrzeugtypen zusammengefasst. Für mun 1 konstruierten wir zwei vereinfachte Modelle, jedes mit 52 Schichten (statt 55 im Original-Problem). Da mehrere unterschiedliche Dekompositionsstrategien auf einen optimalen Fluss angewandt werden können, war es möglich, für jedes von den beiden vereinfachten Problemen mehrere Umlaufpläne zu bilden, die sich durch die jeweilige Dekompositionsstrategie (LI-FO/FIFO/MinAlternation/eXtended MinAlternation I und II/SPP) ergeben haben. Basierend auf einer Teilmenge dieser Umlaufpläne wurde dann die zweite Phase der Heuristik ausgeführt. Die Laufzeiten und die Qualität der Ergebnisse unterscheiden sich in Abhängigkeit von der Größe dieser Teilmenge. Die Tabelle 8.8 enthält einen Vergleich der heuristischen Läufe, für die 2, 4, 6 bzw. 8 Umlaufpläne in der Phase 2 benutzt wurden. Wie die Zeile „#Unfixierte Fahrten“ zeigt, konnten die meisten Fahrten verknüpft werden, so dass die Lösungszeit in der Phase 2 fast vernachlässigbar klein wurde. In der Zeile „heuristisch“ (unter „CPU(min)“) sind jeweils die Zeiten angegeben, die für die Lösung der Hilfsprobleme in der Phase I bzw. des in der Phase II verbleibenden exakten Modells mit fixierten Fahrtenketten benötigt wurden. Hier stellt man fest, dass der Großteil der Zeit in der Phase I verbraucht wurde (301 Minuten für zwei vereinfachte Probleme) und die Zeit für die Optimierung in der Phase II in Vergleich dazu praktisch vernachlässigbar ist (z.B. 1,5 Minuten bei Kettenerkennung mit zwei Umlaufplänen).

Für mun 2 wurde in der Phase I ein vereinfachtes Problem mit 3 Schichten

#ausgewerteter Umlaufpläne	2	4	6	8
#Unfixierte Fahrten	1438	2396	2457	2528
% von 11062	13.0	21.7	22.2	22.9
operative Kosten				
exakt	17.831.010	17.831.010	17.831.010	17.831.010
heuristisch	17.946.839	17.831.739	17.831.73	17.831.739
GAP(%)	0,65	0,004	0,004	0,004
CPU (min)				
exakt	651	651	651	651
heuristisch	301 + 1,5	301 + 1,8	301 + 2	301 + 2

Tabelle 8.8: Ergebnisse der Zwei-Phasen-Heuristik: mun 1

von den ursprünglichen 15 gelöst. Bei der Reduktion der Anzahl berücksichtigter Fahrzeugtyp-Depot-Kombinationen durch eine Streichung der Depots und Zusammenlegung von Fahrzeugtypen wurden die Kapazitäten entsprechend angepasst. Aus der optimalen Flusslösung dieses vereinfachten Problems wurden anschließend zwei Umlaufpläne nach unterschiedlichen Strategien extrahiert. Das vereinfachte Modell in der Phase I wurde in 41 Sekunden gelöst. Nach einem Vergleich beider Umlaufpläne in der Kettenerkennung in der Phase II blieben nur 1646 (15,4%) von ursprünglich 10710 Fahrten unfixiert. Das resultierende Modell mit fixierten Fahrtenverknüpfungen wurde in 11 Minuten gelöst. Insgesamt brauchte die heuristische Lösung etwa 12 Minuten, wobei für das exakte Modell auch nach 1667 Minuten keine Lösung gefunden werden konnte. Da die exakt optimale Lösung für dieses Problem nicht bekannt ist, kann an dieser Stelle keine genaue Aussage über die Qualität der heuristischen Lösung erfolgen.

Numerische Ergebnisse der Fahrtenkettenerkennung

Zum Finden stabiler Fahrtenketten werden die Umlaufpläne verwendet, die bei der Lösung der vereinfachten Probleme entstanden sind und nach unterschiedlichen Dekompositionsstrategien (z.B. LiFo oder FiFo) gebildet wurden. Auf diese Weise wurden in der Phase I für die Probleme city 1 bis city 3 zu jedem vereinfachten Problem zwei Umlaufpläne erzeugt: LiFo- und FiFo-Umlaufplan. Sie konnten dann einzeln oder zusammen bei der Erkennung stabiler Fahrtenketten in der Phase II der des heuristischen Verfahrens eingesetzt werden. Die Tabelle 8.9 fasst die Ergebnisse der Fahrtenketten-Erkennung für alle drei Möglichkeiten (nur LiFo, nur FiFo und beides) für die Fahrpläne city 1 bis city 3 und mun 1 zusammen. Die erste Spalte zeigt, wie viele vereinfachte Probleme in der Phase I gelöst wurden, bzw. für mun1 - wie viele Umlaufpläne der Phase II übergeben wurden. Für city 1 bis 3 zeigen die Spalten 2,3 und 4 die Anzahl der in der Phase II erkannten Fahrtenketten, für den Fall dass nur LiFo-, nur FiFo bzw. sowohl LiFo- als auch FiFo-Umlaufpläne

aus der Phase I ausgewertet wurden. Für die vierte Spalte ist also die Anzahl der ausgewerteten Umlaufpläne doppelt so hoch wie für die beiden ersten Ansätze, da LiFo- und FiFo-Umlaufpläne zusammen zur Fahrtenkettenerkennung herangezogen werden.

Die letzte Spalte gibt den Komprimierungsgrad des Problems in der Phase II der Heuristik im Vergleich zum ursprünglichen Problem an. Für den Fahrplan city 1, beispielsweise, besagt die letzte Spalte, um welchen Anteil der Fahrplan mit Fahrtenketten gegenüber dem initialen Fahrplan mit 2047 Fahrten kleiner geworden ist, wobei bei der Erkennung der Fahrtenketten hier Umlaufpläne verwendet wurden, die durch die Dekompositionen nach LiFo und nach FiFo entstanden sind.

In der Abbildung 8.6 sind die Ergebnisse der Tabelle 8.9 grafisch dargestellt. Die obere Gerade repräsentiert die Anzahl der Fahrten in dem Fahrplan und dient zur Beurteilung des Komprimierungsgrades. Es ist eindeutig zu erkennen, dass ab einer gewissen Anzahl der Umlaufpläne bei der Hinzunahme eines weiteren Umlaufplanes nur wenig neue Fahrtenketten erkannt werden. So werden immer weniger der Fahrtenverknüpfungen zur Erzeugung der neuen Ketten aufgelöst. Außerdem erzeugen die LiFo-Umlaufpläne längere Fahrtenketten als FiFo-Umlaufpläne. Es ist deutlich zu sehen, dass nicht nur der durchschnittliche Komprimierungsgrad bei den Fahrplänen 2 und 3 viel größer als der beim Fahrplan 1 ist, sondern dass auch die „Sättigungsphase“ bei den beiden früher einsetzt.

Generell kann man sagen, dass je mehr verschiedene Umlaufpläne zur Auswertung herangezogen werden, desto näher das erreichte Ergebnis am Optimum liegt. Damit steigt aber die gesamte Lösungszeit des Fahrtenketten-Verfahrens in worst-case an. Die gesamte Lösungszeit besteht aus allen Lösungszeiten der einfacheren Instanzen, die optimal gelöst werden, und der Lösungszeit der konkreten komplexeren Problemstellung, die anhand des Fahrplans aus Fahrtenketten optimal gelöst wird. Man kann hier klar erkennen, dass die Anzahl fixierter Verbindungen dann am größten ist, wenn nur ein Umlaufplan in die Erkennung von Fahrtenketten einbezogen wird. Das heißt, je mehr Umlaufpläne der Phase II übergeben wurden, desto mehr unfixierte Fahrtenketten bleiben für die anschließende Optimierung freigegeben.

Die Anzahl der unfixierten Fahrtenketten ist nach unten durch die Anzahl der Umläufe in jedem der berechneten Umlaufpläne begrenzt. Diese Schranke wird erreicht, falls nur ein vereinfachtes Problem gelöst wird und nur ein Umlaufplan der zweiten Phase übergeben wird. In diesem Fall wird der Umlaufplan mit sich selbst verglichen, so dass alle Fahrtenverknüpfungen durch die vorgegebene Umläufe erhalten bleiben und in der anschließenden Optimierung lediglich auf die Depots und Fahrzeugtypen des ursprünglichen Problems kostenminimal und den Kapazitäten entsprechend aufgeteilt werden müssen.

Anzahl der Unterprobleme	LiFo	FiFo	LiFo/FiFo	Komprimierung bei LiFo/FiFo
city 1				
1	115	115	795	61,16%
2	524	765	967	52,76%
3	629	954	1135	44,55%
4	740	1047	1232	39,81%
5	964	1306	1390	32,10%
6	977	1325	1408	31,22%
7	1097	1436	1482	27,60%
8	1108	1453	1498	26,82%
9	1144	1491	1538	24,87%
city 2				
1	118	118	658	78,45%
2	257	266	745	75,61%
3	410	467	891	70,83%
4	459	529	939	69,25%
5	459	529	939	69,25%
6	495	552	971	68,21%
7	521	582	992	67,52%
8	535	596	1004	67,13%
9	555	673	1082	64,57%
city 3				
1	242	242	1853	73,78%
2	928	959	2269	67,90%
3	1169	1228	2443	65,44%
4	1411	1492	2630	62,79%
5	1435	1522	2653	62,46%

Anz. Umlaufpläne	Fahrtenketten	Komprimierung
mun 1		
1	417	96,23%
2	1438	87,00%
3	2326	78,97%
4	2396	78,34%
5	2452	77,83%
6	2457	77,79%
7	2521	77,21%
8	2528	77,15%

Tabelle 8.9: Anzahl der erkannten Fahrtenketten

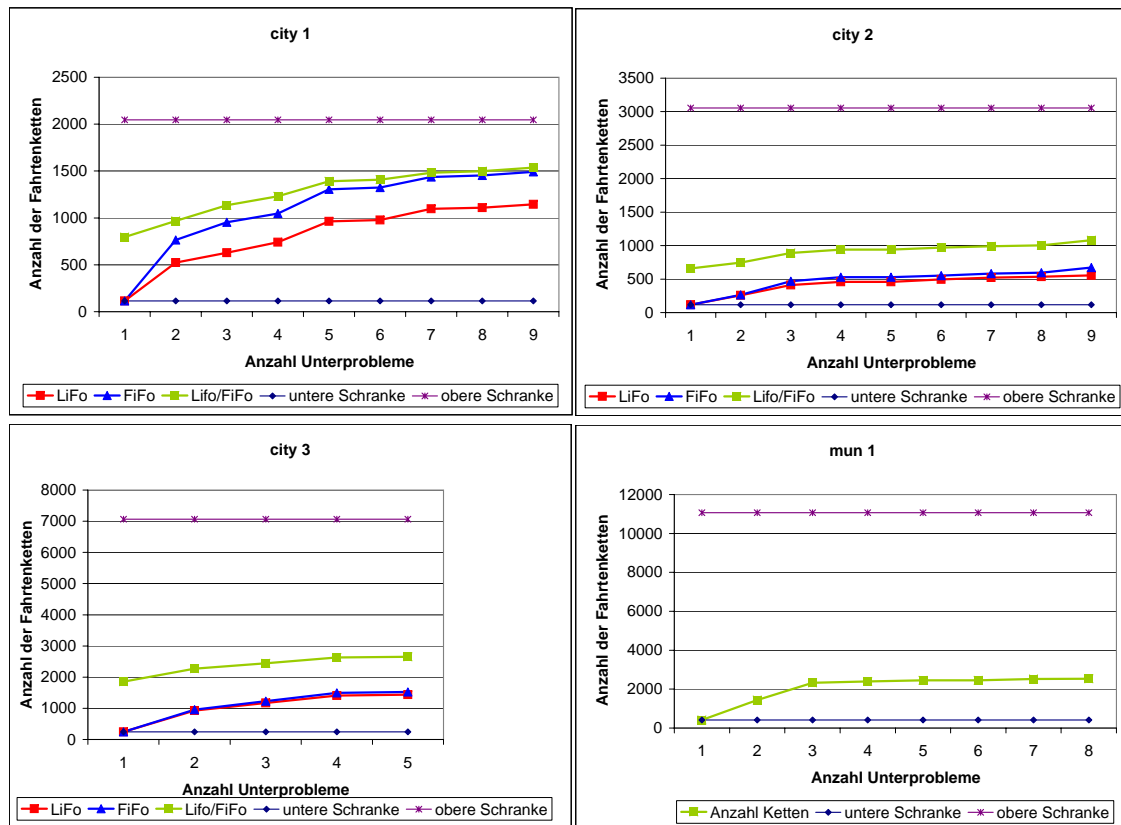


Abbildung 8.6: Ergebnisse der Fahrtenketten-Erkennung für Fahrpläne city 1 bis 3 und mun 1

Eine obere Schranke für die Anzahl unfixierter Fahrtenketten stellt die Anzahl der Fahrten im ursprünglichen Fahrplan dar. Diese Grenze kann nur theoretisch erreicht werden - dafür müssten bei allen untersuchten Umlaufplänen keine Übereinstimmungen in der Zuordnungsrelation der Fahrten (Fahrt zu Nachfolgerfahrt) vorkommen, was für die realistische Fahrpläne äußerst unwahrscheinlich ist. Tatsächlich zeigen die Sättigungskurve in der Abbildung 8.6 dass die Anzahl der unfixierten Fahrtenketten nur einen bestimmten Anteil von der Gesamtzahl der Fahrten des ursprünglichen Fahrplans erreicht. Je mehr unterschiedliche Umlaufpläne ausgewertet werden, desto flacher verläuft die Linie.

Die Experimente zeigen, dass ab einer gewissen Anzahl der Hilfsprobleme bei der Hinzunahme eines weiteren Umlaufplans kaum mehr neue Fahrtenketten erkannt werden - es wird ein Sättigungszustand erreicht, die Größe des komprimierten Fahrplans pendelt sich bei höchstens 40% von der Größe des Originalfahrplans ein.

Zusammenfassend kann man sagen, dass das beschriebene heuristische Vorgehen

vor allem bei den Probleminstanzen aus der Praxis mit sehr vielen Fahrzeugtypen und/oder Depots eingesetzt werden sollte. Das sind die Instanzen mit hoher durchschnittlichen Gruppengröße pro Fahrplanfahrt, für die keine exakt optimale Lösung in annehmbarer Zeit berechnet werden kann. In diesen Fällen liefert das vorgestellte heuristische Verfahren sehr gute zulässige Lösungen in akzeptabler Zeit.

8.4.4 Automatische Erstellung vereinfachter Probleme

Die Bildung der vereinfachten Probleme kann auch automatisch erfolgen. Bei der automatischen Erstellung vereinfachter Probleme für die Phase I der Fahrtenketten-Heuristik kann der Benutzer vorgeben, wie viele solche Probleme bzw. wie viele Umlaufpläne für die Phase II gebildet werden sollen. Diese werden dann durch das Weglassen einiger Depots und/oder Fahrzeugtypen mit Beachtung der Kapazitäten-Relationen gebildet. Je mehr solche Umlaufpläne gebildet werden, desto bessere Ergebnisqualität wird in der Heuristik erreicht. So kann das Trade-Off zwischen der Lösungsqualität und der Laufzeit dadurch gesteuert werden, dass unterschiedlich viele vereinfachte Probleme in der Phase I gebildet werden. Das Verfahren ist so gestaltet, dass die Bildung und Lösung der vereinfachten Hilfsprobleme für den Benutzer völlig transparent verläuft, er muss nur einen Parameter, der ein Maß für die gewünschte Lösungsqualität darstellt, vorgeben.

Kapitel 9

Zusammenfassung und Ausblick

Im Rahmen der vorliegenden Arbeit wurde die Aufgabe der optimalen Planung des Fahrzeugeinsatzes in ÖPNV-Unternehmen behandelt. Dabei bestand die Zielsetzung der Arbeit in der Erarbeitung und Implementierung eines neuen Modells, das die Schwachstellen der existierenden Ansätze vermeidet und dadurch die Lücke zwischen den Anforderungen aus der Praxis und den Möglichkeiten der theoretischen Arbeiten schließt. Dafür mussten insbesondere zwei Ziele erreicht werden. Zum Einen sollten Modelle und Methoden entwickelt werden, die sehr große Instanzen aus der Praxis lösbar machen bzw. die Lösungszeit reduzieren und so einen praktischen Einsatz ermöglichen. Zum Anderen sollten zahlreiche praktische Anforderungen und Zielsetzungen in den Modellen abgebildet bzw. von den Verfahren berücksichtigt werden können. Die Ergebnisse dieser Arbeit werden im Hinblick auf die Zielsetzung im Folgenden zusammengefasst.

Nach der Einführung in die Problemstellung in Kapitel 1 wurden in Kapitel 2 die Rahmenbedingungen und die Zielsetzungen der Mehrdepot-Umlaufplanung herausgearbeitet und diskutiert. Eine Analyse der in Kapitel 3 erwähnten Arbeiten aus der Literatur, die Modelle und Lösungsansätze für das Mehrdepot-Umlaufplanungsproblem bieten, ergab die Motivation zur Erarbeitung eines neuen Ansatzes.

Daher wurde in Kapitel 4 eine Modellierung und Lösung mit Hilfe eines Time-Space-Netzwerkes vorgeschlagen. Als Ergebnis dieser Arbeit wurden reale Problemstellungen mit zusätzlichen praktischen Anforderungen und Rahmenbedingungen in akzeptabler Zeit ohne Einsatz spezialisierter Algorithmen lösbar gemacht. Wie die Testergebnisse belegen, ermöglicht diese Modellierungstechnik eine exakt optimale Lösung auch für große reale Umlaufplanungsprobleme mit mehreren Depots und vielen Tausend Fahrten. Ein positiver Nebeneffekt dieser Modellierung ist, dass eine optimale Lösung des Netzwerkflussmodells nicht nur einen, sondern mehrere, allesamt optimale Umlaufpläne beschreibt. Wie in Kapitel 5 gezeigt wurde, können

diese durch unterschiedliche Dekompositionsstrategien aus dem optimalen Fluss im Netzwerk extrahiert werden. Dies ermöglicht eine Generierung von Umlaufplänen, die neben den optimalen Kosten weitere erwünschte Eigenschaften aufweisen.

Wichtig ist, dass die Probleminstanzen unter Verwendung der in dieser Arbeit vorgeschlagenen Modellierung mit Hilfe der Standard-Optimierungssoftware gelöst werden können. Es erlaubt eine Einbeziehung von verschiedenen Modellerweiterungen, die die zusätzlichen praktischen Anforderungen aus der Praxis abbilden. In Kapitel 6 wurden mehrere Erweiterungen des in Kapitel 4 konzipierten Grundmodells vorgenommen. Dazu zählen die Linienberücksichtigung, das Konzept der Depotgruppen, verschiedene Arten von Kapazitätsanforderungen und die Verschiebeintervalle für Fahrten.

An Beispielaufgaben aus der Praxis wurde in Kapitel 7 die praktische Einsatzfähigkeit der entwickelten Verfahren belegt. In diesem Kapitel wurden organisatorische Maßnahmen zur Kosteneinsparung vorgeschlagen und anhand der numerischen Ergebnisse deren Auswirkungen diskutiert. Die Erkenntnisse dieses Kapitels erlauben es, die Auswirkungen einzelner Maßnahmen in einem Verkehrsunternehmen abzuschätzen.

Neben dem neuen Modell und dem exakten Lösungsansatz wurden in dieser Arbeit Methoden entwickelt, die die Modellgröße oder Komplexität reduzieren und dabei immer noch sehr gute bis optimale Lösungen des gleichen Problems in vertretbarer Zeit ermöglichen. Wie das Kapitel 8 zeigt, wurde dabei der Komplexität extrem großer Probleminstanzen von mehreren Seiten entgegengewirkt. Durch eine Dekomposition der Original-Problemstellung in mehrere unabhängige Teilprobleme können die Problemstellungen vereinfacht werden, ohne Einbußen in der Genauigkeit der Kostenoptimalität zu haben. Weiterhin wurde eine parametergesteuerte Reduktion der Problemgröße ermöglicht, die die Genauigkeit der Berücksichtigung von operativen Kosten zugunsten der besseren Laufzeiten geringfügig opfert, die Fahrzeugflotte aber weiterhin minimal hält.

Um die Lösungszeiten weiter zu reduzieren bzw. um extrem große Probleminstanzen lösen zu können, die mit keinem existierenden exakten Verfahren lösbar sind, wurden approximative Modelle und Verfahren konzipiert und implementiert. Dabei wurden in einer Zwei-Phasen Heuristik in Abschnitt 8.4 einige Fahrplanfahrten im Vorfeld der Optimierung, basierend auf den exakt optimalen Lösungen von mehreren vereinfachten Problemen, zu Teilumläufen miteinander verknüpft. Auf diese Weise kann der Lösungsraum der mathematischen Modelle erheblich reduziert werden, ohne sich zu weit von der exakt optimalen Lösung zu entfernen.

Die in dieser Arbeit vorgeschlagenen exakten und heuristischen Verfahren basieren auf den Methoden der mathematischen Optimierung, die in den Standard-Optimierungsbibliotheken enthalten sind. Daher können sowohl das exakte Ver-

fahren, als auch die Heuristik durch die zukünftig absehbaren Fortschritte in der Rechentechnik und in den Optimierungsbibliotheken noch schneller und leistungsfähiger werden. Es ist zu erwähnen, dass dies alleine durch die Updates der entsprechenden Optimierungsbibliothek ohne zusätzlichen Implementierungsaufwand möglich ist.

Die Konzeption und Umsetzung der Modelle und Algorithmen beachten die Aspekte der Integration in ein Entscheidungsunterstützungssystem für die Planung im ÖPNV. Unterstützt wird ein komplexer Prozess mit zahlreichen Aufgaben, deren Formalisierungsgrad stark variiert. Der Entwurf eines Systems zur Entscheidungsunterstützung sollte nach [Suhl, 1995] nicht nur den möglichen Automatisierungs- und Optimierungsgrad einer Planungsaufgabe perfekt ausnutzen, sondern auch dem Expertenwissen, der Kreativität und nicht zuletzt der Intuition eines kompetenten menschlichen Entscheider Rechnung tragen. Aus diesem Grund wurde die implementierte Umlafoptimierungssoftware mit zahlreichen Einstellungs- und Steuerungsmöglichkeiten versehen. Nach einem Optimierungslauf werden dem Anwender mehrere Entscheidungsvorschläge geliefert, die alle kostenoptimal sind, aber je nach verwendeter Dekompositionsstrategie unterschiedliche Eigenschaften aufweisen. Diese Entscheidungsvorschläge können mit Hilfe der grafischen Oberfläche von *ptv interplan* ausgewählt und bearbeitet werden.

Zusammenfassend lässt sich sagen, dass mit den vorgestellten Ergebnissen die Ziele dieser Arbeit erreicht wurden: Der vorgeschlagene Modellierungsansatz basiert auf einem Time-Space-Netzwerk und vermeidet somit die Nachteile der traditionellen connection-basierten Modelle. Der Lösungsansatz verwendet die allgemeine Branch-and-Cut-Methode, die in den Standard-Optimierungsbibliotheken enthalten ist. Im Gegensatz zu spezialisierten Algorithmen erlaubt dieser Ansatz zahlreiche Modellerweiterungen für die Berücksichtigung praktischer Anforderungen aus der Umlaufplanung. Die vorgeschlagene Zwei-Phasen-Heuristik reduziert die Lösungszeiten bzw. erlaubt eine Lösung von extrem großen Probleminstanzen. Diese Heuristik verwendet ebenfalls die Branch-and-Cut-Methode, so dass alle Modellerweiterungen weiterhin berücksichtigt werden können. Die Lücke zwischen den Anforderungen aus der Praxis der Umlaufplanung und den Möglichkeiten der existierenden theoretischen Ansätze zur Lösung des Mehrdepot-Umlaufplanungsproblems wurde damit geschlossen.

Diese Arbeit ist am Decision Support & Operations Research Lab der Universität Paderborn im Rahmen einer Kooperation mit einem Softwareanbieter für die Bereiche Transport und Verkehr - der PTV AG - entstanden. Mit der in C# entwickelten Softwarekomponente lassen sich komplexe Probleme aus der Praxis mit Hilfe der Standard-Optimierungssoftware, wie MOPS oder ILOG CPLEX, in annehmbarer Zeit optimal lösen. Die Umlafoptimierungskomponente wurde in das

Planungswerkzeug *ptv interplan* integriert und wird in realen ÖPNV-Unternehmen eingesetzt.

Die vorliegende Arbeit eröffnet ein Feld für weitere Forschungs- und Entwicklungsaktivitäten im Bereich der optimalen Planung im ÖPNV. So können zukünftig weitere Lösungsmethoden für das entwickelte Time-Space-Netzwerk-basierte Modell untersucht werden. Zu untersuchen wäre auch wie sich eine mögliche Reduktion der Flusschranken für die Kanten des Netzwerkmodells auf die Lösungszeiten auswirken könnte.

Forschungsbedarf besteht weiterhin in der Integration der Umlaufplanung mit den vor- und nachgelagerten Planungsphasen. Die existierenden theoretischen Ansätze, die eine integrierte Dienst- und Umlaufplanung umfassen, könnten mit den Vorteilen der in dieser Arbeit vorgestellten Modellierung verbunden werden, um damit zu einer praktisch einsetzbaren Lösung zu gelangen. Eine teilweise Integration der Anforderungen aus den nachgelagerten und vorgelagerten Planungsphasen wurde bereits in dieser Arbeit vorgeschlagen. Durch verschiedene Dekompositionsstrategien lassen sich teilweise die Anforderungen aus der Dienstplanung berücksichtigen. Teilentscheidungen der Fahrplankonstruktion wurden durch Verschiebeintervalle für Fahrten in die Umlaufplanung miteinbezogen. Zukünftig sollten diese Ansätze weiter verfolgt werden. Beispielsweise wird bereits die auf der Lösung des Set Partitioning Problems basierende Dekomposition um weitere Zielsetzungen aus der Dienstplanung erweitert und effizienter gemacht. Die Tendenz zeigt, dass die Grenzen zwischen den einzelnen Planungsaufgaben zukünftig immer mehr verwischen werden, was zu einer phasenübergreifenden optimalen Planung im ÖPNV führen wird.

Literaturverzeichnis

- [Ahuja et al., 1983] Ahuja, R., Magnanti, T., and Orlin, J. (1983). *Network Flows: theory, Algorithms and Applications*. A Series of books in the mathematical sciences. Prentice Hall, New Jersey.
- [Andersen, 1998] Andersen, E. (1998). *Linear optimization: Theory, methods, and extensions*. Lecture notes. Odense University, DK.
- [Banihashemi and Haghani, 2000] Banihashemi, M. and Haghani, A. (2000). A model for the multiple depot transit vehicle scheduling problem with route time constraints. In anonymous, editor, *Proceedings of the 8th International Conference on Computer-Aided Scheduling of Public Transport, Berlin*.
- [Barnhart et al., 1996] Barnhart, C., Johnson, E., Nemhauser, G., Savelsbergh, M., and Vance, P. (1996). Branch-and-price: Column generation for solving huge integer programs. *Operations Research*, 4.
- [Bertossi et al., 1987] Bertossi, A., Carraresi, P., and Gallo, G. (1987). On some matching problems arising in vehicle scheduling models. *Networks*, 17:271–281.
- [Bertsekas and Castañon, 1992] Bertsekas, D. and Castañon, D. (1992). A forward/reverse auction algorithm for assymetric assignment problems. *Computational Optimization and Applications*, 1:277–297.
- [Borndörfer et al., 2004] Borndörfer, R., Grötschel, M., and Pfetsch, M. (2004). Models for line planing in public transport. Technical report, Konrad-Zuse-Zentrum für Informationstechnik, Berlin.
- [Borndörfer et al., 2002] Borndörfer, R., Löbel, A., and Weider, S. (2002). Integrierte Umlauf- und Dienstplanung im Nahverkehr. Technical report, Konrad-Zuse-Zentrum für Informationstechnik, Berlin.
- [Bussieck and Zimmermann, 1997] Bussieck, M. and Zimmermann, U. (1997). Schlußbericht: Optimale Linienführung und Routenplanung in Verkehrssystemen. Technical report, Technische Universität Braunschweig.

- [Carpaneto et al., 1989] Carpaneto, G., Dell’Amico, M., Fischetti, M., and Toth, P. (1989). A branch and bound algorithm for the multiple depot vehicle scheduling problem. *Networks*, 19:531–548.
- [Carraraesi et al., 1995] Carraraesi, P., Girardi, L., and Nonato, M. (1995). Network models, lagrangean relaxation and subgradient bundle approach in crew scheduling problems. In Daduna, J., Branco, I., ao, P., and Pinto, J., editors, *Computer-aided Transit Scheduling*, pages 188–212. Springer, Berlin.
- [Chvatal, 1983] Chvatal, V. (1983). *Linear Programming*. A Series of books in the mathematical sciences. W.H. Freeman and Company.
- [Daduna and Mojsilovic, 1988] Daduna, J. and Mojsilovic, M. (1988). Computer-aided vehicle and duty scheduling using hot programme system. In Daduna, J. and Wren, A., editors, *Computer-aided Transit Scheduling*, pages 133–146. Springer, Berlin.
- [Daduna and Paixão, 1995] Daduna, J. and Paixão, J. (1995). Vehicle scheduling for public mass transit - an overview. In Daduna, J., Branco, I., ao, P., and Pinto, J., editors, *Computer-aided Transit Scheduling*, pages 76–90. Springer, Berlin.
- [Daduna and Völker, 1997] Daduna, J. and Völker, M. (1997). Fahrzeugumlaufbildung im ÖPNV mit unscharfen Abfahrtszeiten. *Der Nahverkehr*, 11:39–43.
- [Dell’Amico et al., 1993] Dell’Amico, M., Fischetti, M., and Toth, P. (1993). Heuristic algorithms for the multiple depot vehicle scheduling problem. *Management Science*, 39:115–125.
- [Desaulniers et al., 1998] Desaulniers, G., Lavigne, J., and Soumis, F. (1998). Multi-depot vehicle scheduling problem with time windows and waiting costs. *European Journal of Operational Research*, 111:479–494.
- [Dürr and Kleibohm, 1992] Dürr, W. and Kleibohm, K. (1992). *Operations-Research: lineare Modelle und ihre Anwendungen*. Studienbücher der Wirtschaft. Carl Hanser Verlag München Wien.
- [Fischer et al., 2002] Fischer, J., Herold, W., Nastansky, L., Dangelmaier, W., and Suhl, L. (2002). *Bausteine der Wirtschaftsinformatik - Grundlagen, Anwendungen, PC-Praxis*. Erich Schmidt Verlag, Berlin.
- [Fischetti et al., 1999] Fischetti, M., Lodi, A., and Toth, P. (1999). A branch-and-cut algorithm for the multiple depot vehicle scheduling problem. Technical report, DEIS - Università di Bologna.

- [Forbes et al., 1994] Forbes, M., Holt, J., and Watts, A. (1994). An exact algorithm for multiple depot bus scheduling. *European Journal of Operational Research*, 72(1):115–124.
- [Freling, 1997] Freling, R. (1997). *Models and Techniques for Integrating Vehicle and Crew Scheduling*. PhD thesis, Erasmus University Rotterdam.
- [Freling et al., 2001] Freling, R., Paixão, J., and Wagelmans, A. (2001). Models and algorithms for single-depot vehicle scheduling. *Transportation Science*, 35:165–180.
- [GIRO, 2004] GIRO (2004). www.giro.com. Technical report. Webseite.
- [Grötschel et al., 1997] Grötschel, M., Löbel, A., and Völker, M. (1997). Optimierung des Fahrzeugumlaufs im Öffentlichen Nahverkehr. In u.a., K.-H. H., editor, *Mathematik: Schlüsseltechnologie für die Zukunft; Verbundprojekte zwischen Universität und Industrie*, pages 609–624. Springer, Berlin u.a.
- [Hagemeyer, 2000] Hagemeyer, C. (2000). Branch & price verfahren für das duty scheduling problem. Technical report, Universität zu Köln. Diplomarbeit.
- [Hane et al., 1995] Hane, C., Barnhart, C., Johnson, E. L., Marsten, R. E., Nemhauser, G. L., and Sigismondi, G. (1995). The fleet assignment problem: Solving a large integer program. *Mathematical Programming*, 70/2:211–232.
- [Hanisch, 2002] Hanisch, A. (2002). *GoTo C#*. Addison-Wesley, München.
- [Hoffmann and Padberg, 1985] Hoffmann, K. and Padberg, M. (1985). LP-based combinatorial problem solving. *Annals of Operations Research*, 4:145–194.
- [Huisman, 2004] Huisman, D. (2004). *Integrated and Dynamic Vehicle und Crew Scheduling*. PhD thesis, Tinbergen Institute, Erasmus University Rotterdam.
- [Huisman et al., 2001] Huisman, D., Freling, R., and Wagelmans, A. (2001). A dynamic approach to vehicle scheduling. Technical report, Econometric Institute, Erasmus University Rotterdam. also in *Transportation Science*.
- [Huschke, 2004] Huschke, R. (2004). Neue Freiheitsgrade für die Planer. *Stadtverkehr*, 6:42–45.
- [ILOG, 2004] ILOG (2004). *ILOG CPLEX v9.0 User's Manual*. ILOG, Gentilly, France.
- [IVU, 2004] IVU (2004). www.ivu.de. Technical report. Webseite.

- [Kliwer et al., 2002] Kliwer, N., Mellouli, T., and Suhl, L. (2002). A new solution model for multi-depot multi-vehicle-type vehicle scheduling in (sub)urban public transport. In *Proceedings of the 13th Mini-EURO Conference and the 9th meeting of the EURO working group on transportation*. Politechnic of Bari.
- [Kliwer et al., 2005] Kliwer, N., Mellouli, T., and Suhl, L. (2005). A time-space network based exact optimization model for multi-depot bus scheduling. *European Journal of Operational Research (EJOR)*, erscheint 2005.
- [Lamatsch, 1992] Lamatsch, A. (1992). An approach to vehicle scheduling with depot capacity constraints. In Desrochers, M. and Rousseau, J., editors, *Computer-Aided Transit Scheduling*, pages 181–195. Springer, Berlin.
- [Landkreistag-NRW, 2004] Landkreistag-NRW (2004). Welche Zukunft hat der ÖPNV? Das EuGH-Urteil vom 24. Juli 2003 und seine Folgen. *Themenheft*, 1.
- [Löbel, 1997] Löbel, A. (1997). *Optimal Vehicle Scheduling in Public Transit*. PhD thesis, TU Berlin.
- [Löbel, 1999] Löbel, A. (1999). Solving large-scale multiple-depot vehicle scheduling problems. In Wilson, N., editor, *Computer-Aided Transit Scheduling*, pages 193–220. Springer, Berlin.
- [Löbel and Strubbe, 1996] Löbel, A. and Strubbe, U. (1996). Wagenumlaufoptimierung - Methodischer Ansatz und praktische Anwendung. In *Heureka '96: Optimierung in Verkehr und Transport*, pages 341–355. Forschungsgesellschaft für Straßen- und Verkehrswesen, Köln.
- [Lohatepanont, 2002] Lohatepanont, M. (2002). *Airline fleet assignment and schedule design: integrated models and algorithms*. PhD thesis, MIT.
- [Lohatepanont and Barnhart, 2004] Lohatepanont, M. and Barnhart, C. (2004). Airline schedule planning: Integrated models and algorithms for schedule design and fleet assignment. *Transportation Science*, 38(1):19–32.
- [Mellouli and Kliwer, 2002] Mellouli, T. and Kliwer, N. (2002). Umlaufplanung im öffentlichen Verkehr mit mehreren Depots und Fahrzeugtypen: Neue Lösungsmodelle und praktische Aspekte. In *Heureka'02: Optimierung in Verkehr und Transport*. FGSV Verlag, Köln.
- [Mercier and Soumis, 2005] Mercier, A. Cordeau, J.-F. and Soumis, F. (2005). A computational study of Benders decomposition for the integrated aircraft routing and crew scheduling problem. *Computers & Operations research*, 32:1451–1476.

- [Nachhaltigkeitsrat, 2004] Nachhaltigkeitsrat (2004). Ziele zur nachhaltigen Entwicklung in Deutschland - Schwerpunktthemen. Dialogpapier des Nachhaltigkeitsrates. http://www.nachhaltigkeitsrat.de/n_strategie/dialogpapier.
- [Nemhauser and Wolsey, 1988] Nemhauser, G. and Wolsey, L. (1988). *Integer and combinatorial optimization*. Wiley, New York.
- [PTV, 2004] PTV (2004). www.ptv.de. Technical report. Webseite.
- [Reeves, 1995] Reeves, C. (1995). *Modern heuristic techniques for combinatorial problems*. McGraw-Hill Book Company, London.
- [Reichmann, 1997] Reichmann, T. (1997). *Controlling mit Kennzahlen und Managementberichten: Grundlagen einer systemgestützten Controlling-Konzeption*. Vahlen, München.
- [Ribeiro and Soumis, 1994] Ribeiro, C. and Soumis, F. (1994). A column generation approach to the multiple-depot vehicle scheduling problem. *Operations research*, 42(1):41–52.
- [Schäpers et al., 2002] Schäpers, A., Huttary, R., and Bremes, D. (2002). *C# Kompendium*. Markt+Technik, München.
- [Schwichtenberg and Eller, 2004] Schwichtenberg, H. and Eller, F. (2004). *Programmierung mit der .NET Klassenbibliothek*. Addison-Wesley Verlag, München.
- [Silva et al., 2000] Silva, G., Gualda, N., and Kwan, R. (2000). Bus scheduling based on an arc generation - network flow approach. In *Proceedings of CASPT - 2000, 8th International Conference on Computer-Aided Scheduling of Public Transport, Berlin*.
- [Soumis et al., 1980] Soumis, F., Ferland, J.-A., and Rousseau, J. (1980). A model for large-scale aircraft routing and scheduling problems. *Transportation Research B*, 14B(1/2):191–201.
- [Suhl, 1995] Suhl, L. (1995). *Computer-Aided Scheduling: an Airline Perspektive*. Gabler Edition Wissenschaft. Gabler Verlag, Deutscher Universitäts-Verlag, Wiesbaden.
- [Suhl, 2000] Suhl, U. (2000). MOPS - Mathematical OPTimization System. *OR News*, 8:11–16.
- [Wright, 1997] Wright, S. (1997). *Primal-dual interior-point methods*. SIAM.

