Robuste Effizienz des Ressourceneinsatzes im öffentlichen Personennahverkehr

Inaugural-Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft (Dr. rer. pol.) des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin

vorgelegt von
Dipl. Wirt.-Inform. Bastian Amberg

Berlin, 2017

Erstgutachterin:

Univ.-Prof. Dr. Natalia Kliewer
Freie Universität Berlin

Zweitgutachter:

Univ.-Prof. Dr. Jan Fabian Ehmke
Europa-Universität Viadrina Frankfurt (Oder)

Tag der Disputation:
2. Mai 2017

Abstract

Diese Dissertation behandelt mathematische Optimierungsmodelle und -methoden zur Ressourceneinsatzplanung mit gleichzeitiger Berücksichtigung von Kosteneffizienz und Robustheit beim Einsatz der Ressourcen. Im Anwendungsfeld des ÖPNV werden Lösungsansätze zur Umlaufplanung für Fahrzeuge und zur Dienstplanung für Fahrer entwickelt, die eine Beschränkung aktueller Ansätze aufheben: Bisher wird in Optimierungsansätzen zur Ressourceneinsatzplanung hauptsächlich eine reine Minimierung der geplanten Kosten des Ressourceneinsatzes verfolgt. Dies führt üblicherweise zu straffen Plänen, die bei ihrer Ausführung eher anfällig für Verspätungen sind. Aufgrund fehlender Pufferzeiten und der Abhängigkeiten zwischen Umläufen und Diensten können dann – ohne Eingriff – selbst kleine Verspätungen zu kaskadierenden Verspätungseffekten führen. Um die Funktionsfähigkeit der Pläne trotz äußerer Einflüsse zu gewährleisten, können in der Planung proaktiv mögliche Verspätungen antizipiert und ihre Auswirkungen berücksichtigt werden. Eine solche robuste Planung gegen Störungen und Verspätungen ist jedoch für den gemeinsamen Einsatz von Fahrzeugen und Fahrern kaum untersucht.

In dieser Arbeit werden die gegenseitigen Abhängigkeiten zwischen Umlauf- und Dienstplänen berücksichtigt und ein Baukasten mit neuen Lösungsansätzen vorgeschlagen. Diese beziehen die potenzielle Propagation möglicher Verspätungen in die Planung ein, bei gleichzeitiger Verfolgung von Kosteneffizienz. Dabei berücksichtigen die Ansätze in verschiedenem Maße die Verknüpfung der Planungsebenen Fahrplan-, Umlauf- und Dienstplanung. Es wird untersucht, in welchem Ausmaß die verschiedenen Planungsmethoden zu Plänen mit unterschiedlicher Qualität bezüglich der Zielsetzungen Robustheit und Kosteneffizienz führen.

In den Lösungsansätzen werden netzwerkfluss- und pfadbasierte Modellformulierungen basierend auf Time-Space-Netzwerken kombiniert. Zum Lösen wird Column Generation in Kombination mit Lagrange-Relaxation genutzt. Dabei werden die Pricing-Probleme zur Bestimmung robuster und effizienter Umläufe und Dienste als ressourcenbeschränkte Kürzeste-Wege-Probleme formuliert. Die berechneten Pläne werden zusätzlich durch stochastische Simulation evaluiert. Diese bildet eine reine Bewertung der erwarteten Verspätungspropagation als auch eine Bewertung bei Einbeziehen dispositiver Maßnahmen während der Planausführung ab.

Die Ergebnisse unterstreichen die Relevanz der kombinierten Betrachtung von Umlauf- und Dienstplanung und des Einbeziehens potenzieller Verspätungspropagation in die Planung. Mit den vorgeschlagenen Ansätzen kann die erwartete Verspätungspropagation im Vergleich zu rein kostenminimierender Planung drastisch reduziert werden ohne zusätzliche Kosten. Dabei führt das Abbilden von Propagationseffekten innerhalb der Lösungsverfahren zu einer adäquaten Verteilung von Pufferzeiten in den Plänen inklusive passender, robuster Verknüpfung von Fahrten und Aufgaben unterschiedlicher Ressourcen. Zusätzlich werden durch die Analyse der Planstrukturen der verschiedenartig erstellten Ressourceneinsatzpläne Prinzipien erfolgreich robust effizienter Pläne identifiziert.

Abstract

This thesis deals with mathematical optimization models and methods for resource scheduling with simultaneous consideration of cost-efficiency and robustness during operations. In the field of public transit, solution approaches are developed for vehicle scheduling and crew scheduling that overcome a particular limit of current approaches: Current optimization approaches mainly aim at pure minimization of planned costs for resource usage. This usually results in tight schedules which are more susceptible to delays during operations. Due to the lack of buffer times and due to dependencies between scheduled vehicles and drivers – without intervention – even small delays may cause cascading additional delays. In order to ensure the functionality of the computed schedules despite external factors, possible delays and their effects should already be considered in the planning phase. However, to date there is only little research assessing robust scheduling against disruptions and delays affecting usage of both vehicles and crews.

In this work, the mutual dependencies between vehicle and crew schedules are considered and a modular system with new solution approaches is proposed. The approaches consider the propagation of possible delays during scheduling while simultaneously aiming at cost-efficient resource schedules. In addition, the planning stages vehicle scheduling and crew scheduling and aspects of timetabling are integrated to different extends. It is examined to what extend the various scheduling approaches result in schedules with different qualities in terms of the objectives robustness and cost-efficiency.

In the proposed solution approaches network flow based model formulations are combined with path based formulations based on Time-Space-Networks. As solution method column generation is used in combination with Lagrangean relaxation. The pricing problems for the generation of robust and efficient vehicle blocks and duties are modelled as resource constrained shortest path problems. The computed schedules are evaluated with the help of a stochastic simulation. This simulation incorporates a pure evaluation of the expected propagated delay as well as an evaluation when dispatching strategies are considered during operation of the schedules.

The results stress the relevance of the combined consideration of vehicle and crew scheduling and the consideration of possible delay propagation when scheduling re-

sources. Compared to pure cost-efficient scheduling, the proposed approaches are able to decrease the expected delay propagation drastically without additional costs. Modelling delay propagation effects within the solution methods results in an adequate distribution of buffer times in the schedules including the construction of suitable, robust links between trips and tasks of different resources. Further, based on the analysis of the variously computed resource schedules and their structures, principles of successful robust efficient schedules are identified.

Inhaltsverzeichnis

I Stand der Technik und Forschung **1**

1 Einleitung **3**

2 Planung und Steuerung des ÖPNV-Betriebs **9**

2.1 Fahrplanerstellung 10
2.2 Umlaufplanung 12
2.3 Dienstplanung 13
2.4 Steuerung der Planausführung 15
2.5 Robuste Planung 19
2.6 IT-Systeme zur Planung und Steuerung 20
2.7 Zusammenfassung und Implikationen 21

3 (Robuste) Umlauf- und Dienstplanung: Stand der Forschung **23**

3.1 Ausgewählte Konzepte des Operations Research 24
3.1.1 Optimierung unter Unsicherheit 25
3.1.2 Optimierung mit mehreren Zielkriterien 27
3.2 Kosteneffiziente Umlauf- und Dienstplanung 29
3.2.1 Lösungsansätze zur Umlaufplanung 31
3.2.2 Lösungsansätze zu kombinierter Umlauf- und Dienstplanung . 33
3.3 Robuste Umlauf- und Dienstplanung 37
3.3.1 Lösungsansätze zur Umlaufplanung 38
3.3.2 Lösungsansätze zu kombinierter Umlauf- und Dienstplanung . 41
3.4 Robuste Planung in weiteren Verkehrsdomänen 42
3.4.1 Lösungsansätze im Flugverkehr 42
3.4.2 Lösungsansätze im Bahnverkehr 52
3.5 Zusammenfassung und Implikationen 57

4 Forschungsbedarf und Zielsetzung **61**

II Neue Lösungsansätze für Robuste Effizienz 65

5 Bewertung von Effizienz und Robustheit 67
5.1 Kennzahlen für Kosteneffizienz 68
5.2 Kennzahlen für Robustheit 70
5.2.1 Verspätungspropagation 70
5.2.2 Pünktlichkeit 72
5.2.3 Tauschmöglichkeiten von Ressourcen 73
5.3 Simulation zur Evaluation der Umlauf- und Dienstpläne 75
5.3.1 Monte-Carlo-basierte Simulation 76
5.3.2 Agentenbasierte Simulation 82
5.4 Zusammenfassung und Implikationen 83

6 Übersicht und Zusammenhang der Lösungskomponenten 85
6.1 Zugrunde liegende Time-Space-Netzwerkstruktur 86
6.2 Basisoptimierungsmodelle zum Einbeziehen von Verspätungspropagation 86
6.3 Basislösungschemata zur Erstellung robuster und kosteneffizienter Pläne 88
6.3.1 Schema zur Umlaufplanung 88
6.3.2 Schema zur sequenziellen Planung 89
6.3.3 Schema zur teilintegrierten Planung 89
6.3.4 Schema zur integrierten Planung 90
6.4 Zusammenfassung 90

7 Robuste Effizienz durch teilintegrierte Planung 91
7.1 Problemdefinition R-MDPVCSP 93
7.2 Lösungsschema für R-MDPVCSP 93
7.3 Globale Flussdekomposition als Vergleichsgrundlage für Lösungen der Umlaufebene 95
7.3.1 Modellierung 97
7.3.2 Lösungsverfahren 98
7.4 Umlaufebene – Generierung einer robusten Netzwerkflusslösung . . . 107
7.4.1 Modellierung 109
7.4.2 Lösungsverfahren 112
7.5 Dienstebene 116
7.5.1 Modellierung 117
7.5.2 Lösungsverfahren 118
7.6 Verknüpfung der Umlauf- und Dienstebene 122
7.6.1 Modellierung 122

7.6.2 Lösungsverfahren 123
7.7 Zusammenfassung und Implikationen 127

8 Robuste Effizienz durch integrierte Planung 129
8.1 Problemdefinition R-MDIVCSP 130
8.2 Modellierung des R-MDIVCSP 131
8.3 Lösungsverfahren 134
8.3.1 Initialisierung 135
8.3.2 Master-Problem: Lagrange-Relaxation 137
8.3.3 Pricing-Probleme zur Generierung von Umläufen und Diensten 139
8.3.4 Ganzzahlige Lösungen 141
8.4 Zusammenfassung und Implikationen 143

9 Robuste Effizienz durch integrierte Planung mit Zeitfenstern 145
9.1 Problemdefinition R-MDIVCSP-TW 147
9.2 Retiming-Ansatz als Vergleichsgrundlage 147
9.3 Modellierung des R-MDIVCSP-TW 149
9.4 Lösungsverfahren 154
9.4.1 Generelles Verfahren 154
9.4.2 Problemgrößenreduktion durch Preprocessing-Techniken . . . 156
9.4.3 Problemgrößenreduktion durch Kantengenerierung zur Laufzeit 159
9.5 Zusammenfassung und Implikationen 173

10 Ein hybrider Evolutionärer Algorithmus zum Erreichen Robuster Effizienz 175
10.1 Problemdefinition 2-SDIVCSP 177
10.2 Literaturübersicht – Evolutionäre Mehrzielalgorithmen in Fahrzeugeinsatz- und Personalplanung 177
10.3 Modellierung 2-SDIVCSP 182
10.4 Lösungsverfahren 184
10.4.1 Initialisierung 185
10.4.2 Repräsentation 187
10.4.3 Variationsoperatoren 188
10.4.4 Dekoder 192
10.4.5 Fitnessberechnung 195
10.4.6 Terminierung 197
10.5 Zusammenfassung und Implikationen 197

III Evaluation der neuen Ansätze und Erkenntnisse 199

11 Vergleichende Analysen numerischer Experimente 201
11.1 Konfiguration der Testumgebung 204
11.1.1 Instanzenauswahl 205
11.1.2 Wesentliche Parameter der Simulationsumgebung 205
11.1.3 Wesentliche Parameter der Optimierungsverfahren 206
11.2 Einfluss der Planungsmethodik auf Robustheit und Effizienz 208
11.2.1 Berücksichtigen von Verspätungspropagation in der Planung . 208
11.2.2 Auswirkungen umlaufbasierter Planung des Ressourceneinsatzes 215
11.2.3 Gemeinsame Planung von Umläufen und Diensten im Vergleich 223
11.2.4 Zusammenfassung der Untersuchungen 231
11.3 Einflussfaktoren in den Planstrukturen auf Verspätungspropagation . 232
11.3.1 Verdichtungs- und Vernetzungseffekt der Ressourcenebenen . . 233
11.3.2 Planstruktur-Indikatoren zu Verdichtung und Vernetzung . . . 242
11.3.3 Zusammenfassung der Untersuchungen 248
11.4 Einfluss der Ausweitung der betrachteten Planungsaufgabe 249
11.4.1 Fahrplananpassung basierend auf geplantem Ressourceneinsatz 250
11.4.2 Integrierte Planung mit Zeitfenstern im Vergleich 257
11.4.3 Zusammenfassung der Untersuchungen 264
11.5 Einfluss der Ausweitung des betrachteten Robustheitsbegriffs 265
11.5.1 Berücksichtigen von Ressourcentausch in der Ausführung . . . 266
11.5.2 Zusammenfassung der Untersuchungen 272
11.6 Konsequenzen zum Erreichen Robuster Effizienz in der Planung . . . 273

12 Zusammenfassung und Ausblick 277

Literatur 286

A Anhang - Beispiel Konvergenz der Simulation 301

B Anhang - Beispiel Approximation pareto-optimale Lösungsmenge 303

Abbildungsverzeichnis

2.1 Kreislauf zwischen Planungs- und Steuerungsaufgaben im ÖPNV . . . 9
2.2 Grundbegriffe der Umlaufplanung 13
2.3 Grundbegriffe der Umlauf- und Dienstplanung 15
2.4 Zusammenhang von primärer (*gelb*) und sekundärer (*rot*) Verspätung 17

3.1 Grundlagen des Stands der Forschung für Lösungsansätze zur robusten und kosteneffizienten Ressourceneinsatzplanung im ÖPNV 23

4.1 Robuste Effizienz der zu erstellenden Pläne als Zielsetzung für die Weiterentwicklung aktueller Lösungsansätze 63

5.1 Kennzahlen zur Bewertung von Robustheit und Kosteneffizienz 67
5.2 Tauschpunkt zur Umverteilung von Aufgabensequenzen und Pufferzeiten zur Vermeidung von Verspätungspropagation bei unterschiedlich geplanten Aufgabenzuordnungen (I) und (II) 74
5.3 Simulationsframework zur Evaluation der Robustheit des geplanten Ressourceneinsatzes 77
5.4 Dispositionsregeln zum Fahrzeugtausch 80
5.5 Dispositionsregeln für Dienständerungen 80

6.1 Überblick über Basislösungsschemata zur Berücksichtigung von Kosteneffizienz und Verspätungspropagation 88

7.1 Schritte der teilintegrierten Umlauf- und Dienstplanung mit Einbezug von Robustheitsaspekten in die Planung 91
7.2 Teilintegrierte Umlauf- und Dienstplanung mit Berücksichtigung der Zielsetzungen Kosteneffizienz und Robustheit 94
7.3 Flussdekomposition mit Berücksichtigung von Verspätungspropagation in einem Time-Space-Netzwerk, angelehnt an [Amberg et al., 2011] 97
7.4 Flussdekomposition mit Berücksichtigung von Verspätungspropagation in zwei Time-Space-Netzwerken bei unterschiedlich robusten Flusslösungen 109

8.1 Verknüpfung der Planungsprobleme zur integrierten Umlauf- und Dienstplanung mit Einbezug von Robustheitsaspekten in die Planung 129

8.2 Integrierte Umlauf- und Dienstplanung mit Berücksichtigung der Zielsetzungen Kosteneffizienz und Robustheit orientiert an [Amberg, 2014] 134

9.1 Potenziale durch Verschiebeintervalle in Bezug auf Kosteneffizienz und Robustheit des Ressourceneinsatzes 145

9.2 Einsatz von Retiming zur Reduktion von Verspätungspropagation: (A) vorhandene Pufferzeit kostenneutral umverteilen, (B) zusätzliche Pufferzeit einfügen 148

9.3 Einbeziehen von Kantengenerierung in das integrierte Lösungsverfahren zur Umlauf- und Dienstplanung (mit Zeitfenstern) 162

10.1 Pareto-optimale Lösungsmenge bezüglich Robustheit und Kosteneffizienz175

10.2 Dekoderbasierte problemspezifische Erzeugung eines Individuums . . 185

10.3 Repräsentation eines Individuums abbildbar als unterschiedliche Permutationen der gleichen Dienststückmenge 188

10.4 Cycle-Crossover-Operator zur Rekombination zweier Individuen und gegebenenfalls entstehende Zyklen, die pro Individuum zu einem Zyklus anzupassen sind 190

10.5 Zufälliger Tausch zweier Werte in der Adjazenzrepräsentation G1 als Mutationsoperator und anschließende Anpassung der Binärcodierung G2. 191

11.1 Übersicht der aufeinander aufbauenden Studien und Aufgliederung der drei Hauptforschungsfragen - Teil 1 202

11.2 Übersicht der aufeinander aufbauenden Studien und Aufgliederung der drei Hauptforschungsfragen - Teil 2 203

11.3 Gegenüberstellung geplanter Kosten und erwarteter Verspätungspropagation bei Umlaufplanung (vs) und integrierter Planung (int) mit Pufferzeiten basierend auf fixen Vorgaben (buf) und basierend auf der Berücksichtigung von propagierten Verspätungen (deP) 211

11.4 Eingeplante Pufferzeiten und potenzielle Verspätungspropagation im Vergleich zu rein kostenminimierender Planung am Beispiel der Testinstanz A 214

11.5 Effektgröße zusätzlicher Verspätungspropagation durch Einbeziehen der Dienstebene gegenüber der Propagation auf Umlaufebene im Vergleich von sequenziell und teilintegriert erstellten Plänen am Beispiel der Testinstanz B 223

11.6 Gegenüberstellung geplanter Kosten und erwarteter Verspätungspropagation bei sequenzieller (seq), teilintegrierter (pInt) und integrierter Planung (int) mit Berücksichtigen von Verspätungspropagation (deP) 226

11.7 Eingeplante Pufferzeiten, Anzahl Servicefahrten vor Dienstpausen und potenzielle Verspätungspropagation im Vergleich der Planungsvarianten am Beispiel der Testinstanz C 230

11.8 Aufschlüsselung der gemessenen Verspätungspropagation (*rot gestrichelte Linien*) sowie der nachweisbaren Propagationseffekte in Abhängigkeit der durchgeführten Simulationsvarianten 235

11.9 Gegenüberstellung und Aufschlüsselung von Verspätungspropagation (vertikale Achse) verschiedenartig erstellter Lösungen (horizontale Achse) unterschiedlicher Instanzen (Zeilen) und Planungsschemata (Spalten) in Verdichtungseffekt (*grau*) und Vernetzungseffekte (*hellblau, blau*) 237

11.10 Häufigkeiten der Pufferzeiten in allen Dienststücken für (robust) integriert erstellte Lösungen exemplarisch für Testinstanz B 239

11.11 Anzahl von Dienststückwechseln kategorisiert nach zur Verfügung stehenden Pufferzeiten für integriert, teilintegriert und sequenziell erstellte Lösungen für Testinstanzen A, B, C 241

11.12 Kennzahl zur Verdichtung mit verschiedenen Schwellenwerten (Bilder I-III) gegenübergestellt der Verspätungspropagation auf Dienststückebene für sequenziell (*Dreiecke*), teilintegriert (*Vierecke*) und integriert (*Kreise*) erstellte Lösungen 245

11.13 Kennzahl zur Vernetzung mit verschiedenen Schwellenwerten (Bilder I-III) gegenübergestellt der entstehenden zusätzlichen Verspätungspropagation des kombinierten Ressourceneinsatzes für sequenziell (*Dreiecke*), teilintegriert (*Vierecke*) und integriert (*Kreise*) erstellte Lösungen 247

11.14 Häufigkeiten der Pufferzeiten zwischen Aufgabenwechseln in allen Dienststücken des geplanten Ressourceneinsatzes am Beispiel (robust) integriert erstellter Lösungen für Testinstanz B (Bilder I-IV), sowie zugehöriger AP-Indikator 255

11.15 Anpassungen der Ganglinie (*hellblau*) und Verteilung von Pufferzeiten (*grau*) durch Retiming am Beispiel einer kostenminimalen (robust dekomponierten) Umlauf- und Dienstplanlösung für Testinstanz B, inklusive markierter wesentlicher Änderungen (*orangefarbene Pfeile*) . . 256
11.16 Gegenüberstellung geplanter Kosten und erwarteter Verspätungspropagation bei sequenzieller (seq), teilintegrierter (pInt), integrierter (int) und integrierter Planung mit Zeitfenstern (intTW) mit Berücksichtigen von Verspätungspropagation (deP) 260
11.17 Gegenüberstellung von Propagationseffekten (*schwarz*) im eingeplanten Ressourceneinsatz, sowie Fahrtenverschiebungen (*dunkelblau*), Anpassungen der Ganglinie (*hellblau*) und Verteilung von Pufferzeiten (*grau*) durch integrierte (robuste) Planung mit Zeitfenstern im Vergleich zu integrierter Planung für Testinstanz C 262
11.18 Gegenüberstellung der Kennzahlbewertung zur Verdichtung (Bild I) und Vernetzung (Bild II) des (robust) integriert (mit Zeitfenstern) eingeplanten Ressourceneinsatzes für Testinstanz C 263
11.19 Gegenüberstellung geplanter Kosten und erwarteter Verspätungspropagation bei sequenzieller (seq), teilintegrierter (pInt), integrierter (int) und integrierter Planung mit Zeitfenstern (intTW) mit Berücksichtigen von Verspätungspropagation (deP) und Variation der Robustheitsevaluation 268
11.20 Anzahl Tauschpunkte mit Bezug zur Anzahl eingeplanter Dienststücke (poW) im durch sequenzielle, teilintegrierte, integrierte und integrierte Planung mit Zeitfenstern erstellten Ressourceneinsatz, sowie tatsächlich durchgeführte Tausche (Swaps) in Bezug zur auftretenden Verspätungspropagation durch Vernetzung der Ressourcen 271
11.21 Konsequenzen zum Erreichen Robuster Effzienz basierend auf den durchgeführten Untersuchungen im Hinblick auf die Dimensionen Planungsmethode, Planstruktur und Robustheitsaspekt 275

A.1 Erwartete sekundäre Verspätung in Sekunden je Fahrt (F1, F2, F3) bei primären Verspätungen generiert mit verschobener Weibull-Verteilung (W), Gleichverteilung (U) und Exponentialverteilung (E) 301

B.1 Gegenüberstellung geplanter Kosten und erwarteter Verspätungspropagation der initialen Population (initial) und der finalen Population (final), sowie bei rein kostenminimierender Planung (int) 303

Tabellenverzeichnis

2.1 Anforderungen an die durch einen Fahrplan vorgegebene Leistungserbringung aus Unternehmens- und Kundensicht 12
2.2 Dispositive Maßnahmen und Auswirkungen auf den geplanten Ressourceneinsatz 18

3.1 Kategorisierte Übersicht existierender Ansätze aus der Literatur mit Bezug zu robuster Planung des Ressourceneinsatzes 59

4.1 Zusammenfassung der in Abschnitt 3.5 beschriebenen Potenziale und Grenzen existierender Ansätze zu robuster Ressourceneinsatzplanung in Bus-, Flug- und Bahnverkehr 61

6.1 Basisoptimierungsmodelle zur Minimierung von Kosten (C) und/oder Verspätungspropagation (P) 87

7.1 Generelle Ressourcenverbräuche auf den Kanten im Dienstelement-Abschnitt basierten Time-Space-Dienstgenerierungsnetzwerk nach [Amberg, 2014] 121

9.1 Initiale Netzwerkgrößen bei Einsatz von Kantengenerierung im Vergleich für rein kosteneffiziente Planung 165
9.2 Performanz des integrierten Lösungsverfahrens (mit Zeitfenstern) mit und ohne zusätzliche Kantengenerierung 170

11.1 Eigenschaften der verschiedenen Diensttypen (Zeiten in hh:mm) . . . 205
11.2 Anzahl Umläufe und Dienste bei Planung mit zuvor fixierten Pufferzeiten 212
11.3 Vergleich von Kosteneffizienz und Robustheit bei sequenzieller und teilintegrierter Planung auf Umlauf- (VS) und Dienstebene (CS), sowie beim gemeinsamen Ressourceneinsatz (VCS) 219
11.4 Gegenüberstellung von Umlauf- und Dienstanzahl und erwarteter Verspätungspropagation für ausgewählte Lösungen aus Abbildung 11.6 . 227

11.5 Rechenzeiten der verschiedenen Lösungsschemata im Überblick (in hh:mm:ss) 227
11.6 Einfluss von nachträglichem Retiming auf die Robustheit von mit unterschiedlichen Planungsmethoden erstellten Ressourceneinsatzplänen 252
11.7 Vergleich robust integriert erstellter Pläne (intR) mit nachträglichem Retiming (retiming) und robust integriert erstellter Pläne mit Zeitfenstern (intTW) mit ähnlicher Robustheit 259
11.8 Gemessene Verspätungspropagation und Anzahl Tauschoperationen während der Ausführung als arithmetischer Mittelwert über die Simulationsläufe am Beispiel von Testinstanz C 269

Liste der Algorithmen

7.3.1 Globale Flussdekomposition zur Minimierung der Verspätungspropagation durch Einsatz eines zweiphasigen Column Generation Schemas . 99
7.3.2 Label Setting mit Berücksichtigung der Verspätungspropagation 103
8.3.1 Lösungsverfahren für R-MDIVCSP 136
9.4.1 Pufferzeit-Zubilligen-Heuristik (Variante B) für die Berechnung von Verschiebeintervallen 158
9.4.2 Lösungsverfahren für R-MDIVCSP-TW mit Kanten- und Spaltengenerierung 161
10.4.1 Problemspezifischer SPEA2 für kosteneffiziente und robuste Umlauf- und Dienstplanung 186

Abkürzungsverzeichnis

2-SDIVCSP .... Bi-Objective Single-Depot Integrated Vehicle and Crew Scheduling Problem (Formulierung (10.1)-(10.9))
BVG ..... Berliner Verkehrsbetriebe
CG-P ..... Column Generation Pricing Problem
CSP ..... Crew Scheduling Problem
EA ..... Evolutionary Algorithm
FIFO ..... First-In-First-Out
GFD ..... Global Flow Decomposition
GFD-R-VSP I . Global Flow Decomposition for Robust Vehicle Scheduling I (Formulierung (7.4)-(7.6))
GFD-R-VSP II Global Flow Decomposition for Robust Vehicle Scheduling II (Formulierung (7.11)-(7.14))
GFD-R-VSP II Pow Global Flow Decomposition for Robust Vehicle Scheduling based on Pieces of Work (Formulierung (7.53)-(7.56))
INT ..... Integrated Scheduling
IP ..... Integer Program
ITCS ..... Intermodal Transport Control System
IVCSP ..... Integrated Vehicle and Crew Scheduling Problem
LIFO ..... Last-In-First-Out
LP ..... Linear Program
MATSim ..... Multi-Agent Transport Simulation
MCFP ..... Multi-Commodity-Minimal-Cost-Flow-Problem
MDVSP ..... Multiple-Depot Vehicle Scheduling Problem
MIP ..... Mixed Integer Program
MOEA ..... Multi-Objective Evolutionary Algorithm
ÖPNV ..... Öffentlicher Personennahverkehr
ÖPV ..... Öffentlicher Personenverkehr
P-VC ..... Propagation based on Vehicle and Crew Schedule (Formulierung (5.4)-(5.6))
PD ..... Primary Delay
PINT ..... Partial Integrated Scheduling

Pow ..... Piece of Work, engl. für Dienststück
R-CSP ..... Robust Crew Scheduling Problem (Formulierung (7.44)-(7.46))
R-CSP-E ..... Robust Crew Scheduling Problem with Elastic cost constraint (Formulierung (7.47)-(7.51))
R-MDIVCSP .. Robust Integrated Multiple-Depot Vehicle and Crew Scheduling Problem (Formulierung (8.1)-(8.8))
R-MDIVCSP-TW Robust Integrated Multiple-Depot Vehicle and Crew Scheduling Problem with Time Windows (Formulierung (9.1)-(9.8))
R-MDPVCSP .. Robust Multiple-Depot Partial Integrated Vehicle and Crew Scheduling Problem
R-MDVSP ..... Robust Multiple-Depot Vehicle Scheduling Problem (Formulierung (7.20)-(7.25))
R-MDVSP-E .. Robust Multiple-Depot Vehicle Scheduling Problem with elastic cost constraint (Formulierung (7.26)-(7.33))
RBL ..... Rechnergestütztes Betriebsleitsystem
RCSP ..... Ressource Constrained Shortes Path Problem
SD ..... Secondary Delay
SDVSP ..... Single-Depot Vehicle Scheduling Problem
SEQ ..... Sequential Scheduling
SPEA2 ..... Strength Pareto Evolutionary Algorithm 2
SPP ..... Set-Partitioning-Problem
TSN ..... Time-Space-Network
VDI ..... Verein Deutscher Ingenieure
VDV ..... Verband Deutscher Verkehrsunternehmen
VSP ..... Vehicle Scheduling Problem

Teil I

Stand der Technik und Forschung

Kapitel 1

Einleitung

Im öffentlichen Personennahverkehr (ÖPNV) wurden in Deutschland im Jahr 2015 11,0 Milliarden Personen befördert. Knapp die Hälfte davon im Busverkehr, die andere Hälfte mit Straßenbahnen und Eisenbahnen. Pro Tag legten somit im Durchschnitt 30,1 Millionen Personen eine Fahrt mit einem öffentlichen Verkehrsmittel zurück, die eine Länge von 50 Kilometer nicht überschritt und deren Fahrtdauer kleiner als 60 Minuten war (siehe [Statistisches Bundesamt, 2016]).

Bei der Bedienung dieser Fahrten durch die einzelnen Verkehrsunternehmen sind allerdings Störungen und aus ihnen resultierende Verspätungen unvermeidbar. Dies gilt insbesondere für den Busverkehr, der sich im Vergleich zum Schienenverkehr den Fahrweg weitgehend mit anderen Verkehrsteilnehmern teilt. Exemplarisch hierfür stehen die Statistiken, die jährlich über die Qualität der von den Berliner Verkehrsbetrieben (BVG) erbrachten Leistung veröffentlicht werden und zur Bewertung der im Verkehrsvertrag mit dem Land Berlin vereinbarten Leistungserbringung herangezogen werden. Die BVG bedient im Busverkehr 151 Linien mit 1.392 Fahrzeugen. In 2013 wurden im Busverkehr lediglich 85,1% aller geplanten Fahrgastfahrten mit einer geringeren Verspätung als dreieinhalb Minuten ausgeführt. Die entsprechenden Quoten bei der Straßenbahn lagen bei 91,8% und bei der U-Bahn bei 97,3% (siehe [Center Nahverkehr Berlin, 2015]).

Die ursprünglichen, primären Verspätungen basieren dabei auf externen Ursachen wie der Verkehrslage oder Witterungseinflüssen. Sie sind deshalb vom Verkehrsunternehmen nicht unmittelbar zu beeinflussen, wirken aber auf die Planausführung der eingesetzten Fahrzeuge und Fahrer. So kann es beispielsweise zu Konflikten für die weitere Planausführung kommen, wenn ein Bus für die planmäßige Ausführung einer Linienfahrt benötigt wird, dieser aber noch mit einer Verspätung eine andere Linienfahrt bedient. Erfolgt kein Eingriff eines Disponenten des Verkehrsbetriebes, entscheidet allein die Beschaffenheit der Einsatzpläne des betroffenen Fahrzeuges und des betroffenen Fahrers darüber, ob die Verspätung entweder abgebaut oder im Betriebsablauf weitergereicht wird. Auch eine Propagation der Verspätung auf weitere Fahrzeuge und Fahrer ist dabei über die Sicherung von Anschlussverbindungen oder

zuvor eingeplante Fahrer- und Fahrzeugwechsel möglich. Die *Robustheit* der Pläne gegenüber Verspätungen, also ihre Funktionsfähigkeit trotz äußerer Einflüsse, wird daher durch zwei Faktoren bestimmt:

- dem Grad der Vernetzung der Planaktivitäten,
- der Verteilung von Pufferzeiten zwischen der Ausführung der Planaktivitäten.

In der *Umlaufplanung* werden Fahrzeuge geplanten Fahrten zugeordnet und in der *Dienstplanung* Dienste so für Fahrer erstellt, dass alle Fahrzeugaktivitäten abgedeckt werden. Die beiden komplexen Planungsaufgaben sind wesentliche und für ÖPNV-Unternehmen wirtschaftlich bedeutende Planungsprobleme in einem umfangreichen, mehrstufigen Planungsprozess, dessen Bearbeitung zunehmend durch leistungsfähige Softwarelösungen unterstützt wird. Insbesondere zur Ressourceneinsatzplanung beinhalten moderne Planungssysteme (wie z.B. HASTUS, IVU.plan, MOBILE-PLAN) neben umfangreichen grafischen Möglichkeiten für eine manuelle Bearbeitung auch leistungsfähige mathematische Optimierungskomponenten zur automatisierten Lösung. Diese profitieren neben dem technologischen Fortschritt, der sich in stark gestiegener Rechenleistung widerspiegelt, vor allem auch vom Fortschritt der Forschung der letzten Jahre aus dem Bereich des Operations Research. Zahlreiche Ansätze (Modelle, Methoden, Algorithmen) zur Lösung der Planungsprobleme ermöglichen es, immer größere Problemstellungen einer Planungsebene in kürzerer Zeit bearbeiten zu können. Darüber hinaus rückt statt einer traditionellen, sequenziellen Bearbeitung eine integrierte Betrachtung der Planungsaufgaben in den Vordergrund, so dass Problemstellungen verschiedener Planungsebenen auch teilweise gemeinsam gelöst werden können, wodurch größere Freiheitsgrade für die Planung entstehen.

Der Fokus aktueller Optimierungsmethoden im ÖPNV liegt auf einer Minimierung der geplanten Kosten. Ihre Verwendung bei der Umlaufplanung bzw. der Dienstplanung hilft zusätzliche Entscheidungsmöglichkeiten für einen kostenoptimierten Ressourceneinsatz aufzuzeigen und kann daher einen enormen Vorteil bezüglich der *Kosteneffizienz* von Umlauf- und Dienstplänen bringen. Diese können so gestaltet werden, dass eine kostengünstigere Bedienung des Fahrplans möglich ist, indem durch raffinierte Planung weniger Fahrzeuge, weniger Leerfahrtkilometer oder weniger Arbeitsstunden für die gleiche Leistungserbringung eingesetzt werden. Ressourcen werden dabei hauptsächlich durch einen Abbau von vermeidbaren Pufferzeiten und einer stärkeren Vernetzung der Planaktivitäten freigesetzt.

In besonders stark vernetzten Plänen können einzelne Störungen während der Planausführung allerdings weitläufige Effekte haben. Je knapper und vernetzter der Ein-

satz von Fahrzeugen und Fahrerdiensten erfolgt, desto stärker wird auch die Ausbreitung von Verspätungen während der Ausführung möglich.

Um die Propagation von Verspätungen einzudämmen und kaskadierende Effekte im Transportnetzwerk zu vermeiden, können zum einen gehäuft Dispositionsentscheidungen getroffen werden. Die ÖPNV-Unternehmen können im laufenden Betrieb auf Verspätungen reagieren und versuchen, durch geschickte Um-Disposition der Ressourcen den Fahrplan weitestgehend einzuhalten. Im äußersten Fall auch durch den Ausfall oder der Verkürzung von Linienfahrten, oder aber durch den Einsatz von Reservefahrzeugen oder zusätzlichem Fahrpersonal. Einer extrem kosteneffizienten Planung muss also nicht zwangsläufig eine kosteneffiziente Ausführung folgen. Zum anderen können aber auch bereits im Vorfeld bei der Planung von Umläufen und Diensten mögliche spätere Verspätungen antizipiert und berücksichtigt werden, um dadurch die Auswirkungen von Störungen während der Ausführung zu verringern und die Funktionsfähigkeit der Pläne trotz äußerer Einflüsse zu gewährleisten. An dieser Stelle setzt die vorliegende Arbeit an.

Zielsetzung und Forschungsfragen

Die Zielsetzung der vorliegenden Arbeit besteht in der **Entwicklung von Optimierungsmodellen und -methoden**, die eine *Robuste Effizienz* von Umlauf- und Dienstplänen im ÖPNV erlauben. Die erzeugten Pläne sollen also sowohl einen geringen Ressourceneinsatz bei hoher Auslastung der Ressourcen erlauben, als auch bei ihrer Ausführung möglichst unempfindlich gegenüber von (kleineren) Störungen verursachten Verspätungen sein. Die beiden Faktoren Kosteneffizienz und Robustheit werden dafür zusammen berücksichtigt und ihre gemeinsame Optimierung wird während der Planerstellung verfolgt. Der Schwerpunkt liegt dabei auf einer integrierten Betrachtung der Umlauf- und Dienstplanung. Da bei einer integrierten Betrachtung Fahrzeug- und Fahrereinsätze simultan geplant werden, ergeben sich höhere Freiheitsgrade für die Dienstplanung als es im Vergleich mit einer traditionellen sequenziellen Bearbeitung der Planaufgaben der Fall ist. Insbesondere können bei der integrierten Planerstellung die gegenseitigen Abhängigkeiten zwischen Umläufen und Diensten berücksichtigt werden, um eine mögliche Propagation von Verspätungen sowohl innerhalb als auch zwischen diesen beiden Ressourcenebenen zu minimieren.

Zur Erreichung der Zielsetzung werden folgende **Forschungsfragen** untersucht:

F1. Welchen Einfluss haben unterschiedliche Planungsmethoden auf Robustheit und Effizienz der erstellten Pläne?

F2. Wieviel Robustheit kann in kostenoptimalen Plänen realisiert werden?

F3. Welche Planungsempfehlungen können für den robusten und kosteneffizienten Fahrzeug- und Personaleinsatz hergeleitet werden?

Methodische Verortung der Arbeit

Mit der gewählten Zielsetzung folgt diese Arbeit dem Forschungsparadigma *gestaltungsorientierter Wirtschaftsinformatik* (siehe [Österle et al., 2010]) bzw. dem angelsächsischen Äquivalent *Design Science Research* (siehe [Hevner et al., 2004]) an der Schnittmenge mit angewandtem *Operations Research*. Die entwickelten Optimierungsmodelle und -methoden bilden dabei die als Ergebnis zu erstrebenden, innovativen Artefakte. Für unterschiedliche Entscheidungssituationen zeigen sie konkrete Möglichkeiten für einen robusten und kosteneffizienten Ressourceneinsatz auf. Sie können somit als Bausteine für IT-Systeme zur Entscheidungsunterstützung eingesetzt werden. Darüber hinaus können mit ihrer Hilfe Untersuchungen durchgeführt werden, um Leitlinien für den Ressourceneinsatz abzuleiten.

Für die Entwicklung und den Entwurf der Artefakte werden im Rahmen dieser Arbeit Methoden des Operations Research verwendet, wobei mathematische Optimierung, Simulation und Aspekte der Datenanalyse miteinander verknüpft werden. Die einzelnen Artefakte werden dabei aufeinander aufbauend gestaltet und bilden so in ihrer Gesamtheit einen Entwicklungsprozess ab. Für eine Diskussion über philosophische Unterschiede zwischen Operations Research und Design Science Research sei an dieser Stelle auf [O'Keefe, 2014] verwiesen.

Das generelle Vorgehen der Arbeit orientiert sich am Erkenntnisprozess der gestaltungsorientierten Wirtschaftsinformatik (siehe [Österle et al., 2010]). Die einzelnen Prozessphasen - von der Analyse zur Erhebung der Problemstellung und der Forschungsfragen, über den Entwurf von Artefakten, bis hin zur Evaluation der geschaffenen Artefakte - spiegeln sich im dreigeteilten Aufbau der Arbeit wider. Dabei stellt die Arbeit selbst einen Bestandteil der vierten Phase, der Diffusion der Forschungsergebnisse, dar.

Gliederung der Arbeit

Die Arbeit gliedert sich in drei Teile mit insgesamt zwölf Kapiteln. Der erste Teil gibt eine Einleitung in das Themenfeld dieser Arbeit und diskutiert den aktuellen Stand der Forschung zur Umlauf- und Dienstplanung im ÖPNV und anderen Fachgebieten, die sich im Gegensatz zur Forschung im ÖPNV bereits intensiv mit der Berücksichtigung von Robustheit bei der Planung auseinandersetzen. Er schließt mit der Herleitung des Forschungsbedarfs und der Zielsetzung dieser Arbeit. Im zweiten Teil werden

basierend auf den zuvor identifizierten Forschungsmöglichkeiten stufenweise Lösungsansätze für eine Planung mit Robuster Effizienz entwickelt. Im dritten Teil werden basierend auf den zuvor entwickelten Lösungsansätzen Ergebnisse durchgeführter numerischer Experimente analysiert und daraus resultierende Erkenntnisse abgeleitet. Außerdem werden Erkenntnisse im Hinblick auf die formulierten Forschungsfragen dieser Arbeit zusammengefasst und ein Ausblick auf zukünftige Forschungsmöglichkeiten gegeben.

Teil 1: Kapitel 2 bildet den fachlichen Rahmen der Arbeit. Auf der einen Seite werden die Aufgaben der Fahr-, Umlauf- und Dienstplanung mit ihren Zielsetzungen innerhalb des Planungsprozesses im ÖPNV betrachtet. Auf der anderen Seite werden die Ausführung und Steuerung der Pläne und der Umgang mit Folgen von Störungen im Betriebsablauf gegenübergestellt. Als Verknüpfung zwischen Planung und Ausführung werden darüber hinaus Aspekte robuster Pläne erörtert.

Kapitel 3 ergänzt den fachlichen Rahmen um ein methodisches Fundament. Im Allgemeinen werden zunächst wesentliche Konzepte zur Optimierung unter Unsicherheit, sowie zur mehrkriteriellen Optimierung aus der Literatur identifiziert. Im Speziellen werden dann Lösungsansätze zu kosteneffizienter Planung sowie robuster Planung in unterschiedlichen Problemdomänen vor dem Hintergrund von mathematischer Optimierung, Simulation und Datenanalyse diskutiert.

Kapitel 4 gibt die Zielrichtung der weiteren Arbeit vor, indem Forschungsbedarf sowohl aus fachlicher Sicht als auch aus methodischer Sicht hergeleitet wird. Dabei wird das Leitbild der Robusten Effizienz für diese Arbeit präzisiert.

Teil 2: Kapitel 5 stellt den Ausgangspunkt für die Entwicklung von Lösungsansätzen im weiteren Verlauf dieses Teils der Arbeit dar. Zu Beginn wird untersucht, an welchen Merkmalen von Ressourceneinsatzplänen Kosteneffizienz und insbesondere Robustheit ausgemacht werden können. Daran anknüpfend wird ein Simulationsmodell zur Evaluation der Robustheit von Umlauf- und Dienstplänen und der Aussagekraft dieser Indikatoren entwickelt.

Kapitel 6 gibt im Anschluss einen Überblick über die in den folgenden Kapiteln ausgearbeiteten Lösungsansätze.

Kapitel 7 schlägt als grundlegenden Ansatz für Robuste Effizienz in der Umlauf- und Dienstplanung ein teilintegriertes Modell bzw. Lösungsverfahren vor. Dabei werden die zuvor beschriebenen Kennzahlen als Zielgrößen in die Planung mit einbezogen und Einflussmöglichkeiten auf die Planerstellung auf Umlaufebene, Dienstebene und in der Verknüpfung der beiden Ebenen herausgearbeitet. Gleichzeitig wird ein Bau-

kasten von Methoden-Komponenten entwickelt, der in den folgenden Kapiteln im Rahmen weiterer Ansätze ebenfalls verwendet wird.

Kapitel 8 greift den zuvor entwickelten Ansatz innerhalb eines vollständig integrierten Modells bzw. Lösungsverfahrens auf. Dabei wird bei gleichzeitiger Zielverfolgung von Kostenminimierung speziell die Berücksichtigung potenzieller Verspätungspropagation in der Planung simultan auf Umlauf- und Dienstebene abgebildet und im Verfahren umgesetzt.

Kapitel 9 erweitert die zuvor ausgearbeiteten Ansätze um eine zusätzliche Einflussgröße, indem nachträglich zeitliche Fahrtenverschiebungen auf Fahrplanebene durch Entscheidungen auf der Umlauf- und Dienstebene erlaubt werden. Dafür werden zum einen Fahrplanänderungen bei gegebenen Ressourceneinsatzplänen diskutiert. Zum anderen wird ein integriertes Modell bzw. Lösungsverfahren vorgeschlagen, das Fahrplan-, Umlauf- und Dienstebene gemeinsam betrachtet. Darüber hinaus werden Techniken zur Reduktion der entstehenden Problemkomplexität untersucht.

Kapitel 10 bildet im Vergleich zu den vorangeganenen Lösungsansätzen eine andere Herangehensweise zur Erreichung Robuster Effizienz ab. Basierend auf einer mehrkriteriellen Metaheuristik und zuvor entwickelten Lösungskomponenten wird ein hybrides Verfahren ausgearbeitet, das im Lösungsverlauf eine breite Menge von Lösungen generiert, die eine Pareto-Front zwischen den Zielen Robustheit und Kosteneffizienz formen.

Teil 3: Kapitel 11 evaluiert die vorgestellten Lösungsansätze und diskutiert robusten und kosteneffizienten Ressourceneinsatz anhand aufeinander aufbauender Studien. Diese basieren auf Experimenten mit den in den vorangegegangenen Kapiteln entwickelten Optimierungsmodellen und Lösungsverfahren sowie der vorgestellten Simulationsmodelle. Die einzelnen Studien unterscheiden sich jeweils in denen bei der Planung und der simulierten Ausführung des Ressourceneinsatzes einbezogenen Freiheitsgraden.

Kapitel 12 fasst die Ergebnisse der Arbeit zusammen und gibt neben Implikationen für die Praxis einen Ausblick auf weiterführende Forschungsaktivitäten.

Kapitel 2

Planung und Steuerung des ÖPNV-Betriebs

Planungsaufgaben im ÖPNV sind vielfältig und umfassen strategische, taktische und operative Überlegungen, jeweils bezogen auf unterschiedliche Zeiträume und Zielstellungen (siehe [Desaulniers und Hickman, 2007]). Dieses Kapitel gibt eine Übersicht über die Aufgaben und stellt dabei den Gesamtrahmen dar, in welchem die vorliegende Arbeit inhaltlich eingebettet ist. Dafür werden insbesondere die taktischen und operativen Aufgaben von der Fahrplanerstellung bis hin zur Ressourceneinsatzplanung näher erläutert und um eine Darstellung der Aufgaben in der Planausführungsphase ergänzt. Abbildung 2.1 zeigt den Kreislauf zwischen Planungs- und Steuerungsaufgaben und fasst (grün hinterlegt) den Gestaltungsraum dieser Arbeit im Kontext des Planungsprozesses im ÖPNV zusammen.

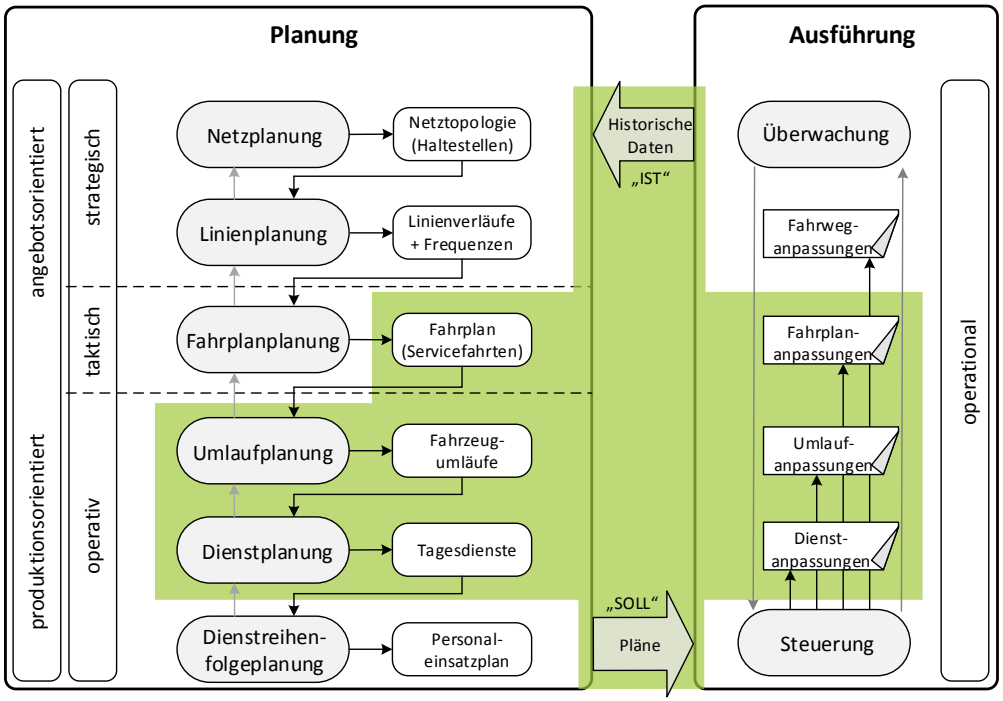

Abbildung 2.1: Kreislauf zwischen Planungs- und Steuerungsaufgaben im ÖPNV

Die *strategische* Planung umfasst neben der Ermittlung des Beförderungsbedarfes auch die Planung des Streckennetzes mit Haltestellenpositionen und zugehöriger Linienstruktur, die angibt in welcher Reihenfolge einzelne Haltestellen anzufahren sind. Ausgehend von diesen Überlegungen wird in der *taktischen* Planungsebene ein Fahrplan auf saisonaler Basis erstellt. Wurden in der strategischen Ebene hauptsächlich räumliche Aspekte bei der Planung betrachtet, kommen nun noch zeitliche Aspekte hinzu. Für die einzelnen Linien müssen Bedienhäufigkeiten, genaue Fahrtenabstände und konkrete Abfahrtszeitpunkte an Haltestellen festgelegt werden. Die dadurch entstehenden Fahrplanfahrten bilden die Grundlage für die *operative* Planung, in welcher über den Ressourceneinsatz zur Bedienung des Fahrplanes entschieden wird. Der Einsatz von Fahrzeugen und Fahren, sowie die Einhaltung des Fahrplanes muss dann anschließend in der *Ausführung* überwacht und gesteuert werden.

Für weitere, vertiefende Informationen über den gesamten Planungsprozess sei auf [Desaulniers und Hickman, 2007] und [Ceder, 2007] verwiesen. Die folgenden Abschnitte konzentrieren sich auf diese Arbeit berührende Aufgaben. So wird in Abschnitt 2.1 auf die Fahrplanerstellung eingegangen, in Abschnitt 2.2 auf die Umlaufplanung und in Abschnitt 2.3 auf die Dienstplanung. Die Aufgabe der Ausführung und Steuerung von Plänen wird in Abschnitt 2.4 thematisiert, sowie eine robuste Planung von Umläufen- und Diensten in Abschnitt 2.5. Abschnitt 2.6 gibt einen kurzen Überblick über kommerzielle Softwaresysteme zur Unterstützung der Planung und Leistungserbringung. Abschnitt 2.7 fasst dieses Kapitel im Hinblick auf die weitere Arbeit zusammen.

2.1 Fahrplanerstellung

Der Fahrplan bildet die Schnittstelle zur Kommunikation der konkreten Transportdienstleistung eines Verkehrsunternehmens an seine Kunden, indem Ankunfts- und Abfahrtszeiten an einzelnen Haltestellen und somit konkrete Zeiten für Wege im ÖPNV-Netz festgelegt werden. Aus Planungssicht markiert er gleichzeitig den Übergang von angebotsorientierter zu produktionsorientierter Planung. So sind bei der Erstellung des Serviceangebotes Fahrplanfahrt zum einen die Passagierbedürfnisse wie beispielsweise eine geringe Reise- oder Wartezeit zu berücksichtigen. Zum anderen werden aus Unternehmenssicht mit der Festlegung von Fahrplanfahrten Kostenvorgaben für den Einsatz von Fahrzeugen und Fahrern getroffen. Ein aus beider Sicht idealtypischer Fahrplan vereint eine Maximierung des möglichen Servicegrades für den Kunden und eine Minimierung potenzieller Kosten für die Bedienung des Fahrplanes durch das Unternehmen.

Vor diesem Hintergrund ist für die Fahrplanung zunächst auf Basis der betrachteten Linien und deren in Abhängigkeit des potenziellen Fahrgastaufkommens bestimmten Bedienhäufigkeiten der genaue Fahrtenabstand zwischen Fahrten zu klären, die eine Haltestelle anfahren. Dieser kann durch regelmäßige oder unregelmäßige Zeitabstände angegeben werden. Üblicherweise werden feste Taktfrequenzen bzw. Taktabfahrtszeiten vergeben. Diese dienen zum einen der besseren Merkbarkeit (für Fahrgast und Busfahrer), zum anderen wird durch feste Taktabfahrtszeiten eine einfachere (manuelle) Planbarkeit durch wiederkehrende Muster von Fahrtenfolgen geschaffen. Häufig wird durch zusätzliche Fahrten zu Hauptverkehrszeiten (z.B. 6:30 Uhr bis 8:30 Uhr, 14:30 Uhr bis 18:30 Uhr) eine bedarfsgerechte Verdichtung des Fahrangebotes erreicht, die in kürzeren Fahrtenabständen und höherer Frequenz resultiert. Ebenso können Fahrtenabstände zu Schwachverkehrszeiten (z.B. Betriebsbeginn bis 6:30 Uhr, 20:00 Uhr bis Betriebsende) wieder vergrößert werden.

Als weiterer Aspekt müssen die Abfahrtszeiten an möglichen Knotenpunkten im Nahverkehrsnetz aufeinander abgestimmt werden. Dadurch können Anschlüsse zwischen verschiedenen Linien bzw. verschiedenen Transportmodi gewährleistet werden, so dass Umsteigevorgänge für Fahrgäste mit wenig Wartezeiten verbunden sind. Attraktiv wird ein ÖPNV-Angebot durch kurze Reisezeiten innerhalb des Netzes und geringe Wartezeiten. Je ausgedünnter das Verkehrsnetz ist und desto größer einzelne Fahrtenabstände sind, desto wichtiger sind synchronisierte Abfahrtszeiten, um lange Wartezeiten zwischen zwei Fahrten zu verhindern. Die gebildeten Anschlussbeziehungen zwischen Fahrten sind insbesondere für die Ausführung und Steuerung der Pläne relevant (siehe Abschnitt 2.4).

Die unter Berücksichtigung der vorangegangenen Überlegungen erstellten Fahrten zur Fahrgastbeförderung bilden als *Servicefahrten* die grundlegende Eingabe für die Ressourceneinsatzplanung. Jede Servicefahrt kennzeichnet sich durch eine Start- und Endhaltestelle sowie eine Anfangs- und Endzeit aus. Dazwischen umfasst sie eine Folge planmäßig zusammenhängender Fahrten zwischen zwei Haltestellen einer Linie, an denen Fahrgäste ein- und aussteigen. An den Stationen zwischen Start- und Endhaltestelle wird in der Regel kein Fahrzeug bzw. Personal zur Bedienung der Servicefahrt gewechselt. Dennoch können hier kurze Warte- bzw. Standzeiten eingeplant sein, um kleine Unregelmäßigkeiten im Verkehrsfluss ausgleichen zu könnnen.

Tabelle 2.1 fasst die Anforderungen an die in einem Fahrplan konzentrierte und als Ausgangsgröße für den Ressourceneinsatz vorgegebene Leistungserbringung eines Verkehrsunternehmens zusammen. Die Tabelle verdeutlicht in welchem Rahmen durch einen Fahrplan Vorgaben für die Umlauf- und Dienstplanung festgelegt werden.

Tabelle 2.1: Anforderungen an die durch einen Fahrplan vorgegebene Leistungserbringung aus Unternehmens- und Kundensicht

Perspektive	Verkehrsunternehmen	Fahrgast
Ziel	Minimierung potenzieller Kosten der Bedienung bei gleichzeitiger Qualitätsmaximierung	Maximierung des gewünschten Service (Reisegeschwindigkeit, Komfort)
Maßnahmen	- geringer Ressourceneinsatz - koordiniertes Netzwerk - zuverlässiges Netzwerk	- häufige Bedienung der Haltestellen - geringe Wartezeiten - wenig Umsteigevorgänge - Taktfahrplan - Pünktlichkeit

2.2 Umlaufplanung

Die in der Fahrplanung erstellten Servicefahrten bilden die Eingabe für die Umlaufplanung. In diesem Planungsschritt wird der Fahrzeugeinsatz zur Bedienung des Fahrplanangebotes bestimmt und im *Umlaufplan* festgelegt. Dieser bezieht sich generell auf den Fahrplan eines Verkehrstages. Zielsetzung ist eine kostengünstige Bedienung des Fahrplanes, indem die dafür benötigte Fahrzeugflotte und die durch sie entstehenden Kosten so gering wie möglich gehalten werden.

Bei den Überlegungen zur Planerstellung sind zum einen unterschiedliche *Fahrzeugtypen* zu berücksichtigen, die sich in Kosten, Ausstattung, Kapazität und/oder Geschwindigkeit unterscheiden können. Bestimmte Servicefahrten erfordern beispielsweise aufgrund des zu erwartenden Fahrgastaufkommens Fahrzeuge mit hoher Kapazität. So werden im Berliner Nahverkehr z.B. auf Linien mit hohem Touristenaufkommen gerne Doppeldeckerfahrzeuge eingesetzt. Solche Fahrzeuge sowie Gelenkbusse sind wiederum nicht dafür geeignet, Servicefahrten entlang enger Straßenführung zu bedienen. Zum anderen können unterschiedliche *Depotstandorte* in die Planerstellung miteinfließen. Gerade in größeren ÖPNV-Räumen besitzen Verkehrsunternehmen häufig mehrere Standorte, von denen die Fahrzeuge ihre Aktivitäten starten und in welche sie (spätestens) zum Betriebsende zurückkehren. In der Regel sind Fahrzeuge dann einem konkreten Depot zugeordnet, um ihren Einsatzbereich eingrenzen und die Planung zu vereinfachen. Dabei können die einzelnen Depots unterschiedliche Kapazitäten an Stellplätzen besitzen.

Alle während des Betriebstages auszuführenden Aktivitäten eines Fahrzeuges bilden einen *Fahrzeugumlauf* bzw. *(Tages-)Umlauf*. Fahrtenfolgen zwischen zwei Depotauftenthalten werden als *Teilumlauf* bzw. *Umlaufblock* bezeichnet (siehe auch Abbildung 2.2 zur Erläuterung). Neben den Servicefahrten zur Fahrgastbeförderung und eingeplanten Standzeiten umfassen die Fahrtenfolgen eines (Teil-)Umlaufs auch

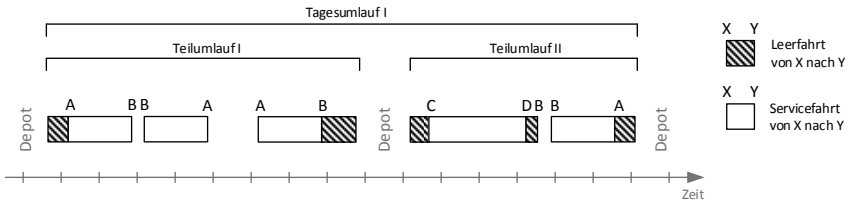

Abbildung 2.2: Grundbegriffe der Umlaufplanung

Fahrten ohne Passagiere. Solche *Leerfahrten* dienen dem Standortwechsel eines Fahrzeuges und lassen sich in *Ein- bzw. Ausrückfahrten* und *Verbindungsfahrten* unterteilen. Eine *Ausrückfahrt* bezeichnet dabei eine Fahrt von einem Depotstandort zur Starthaltestelle einer Servicefahrt. Eine *Einrückfahrt* führt von der Endhaltestelle einer Servicefahrt in ein Depot. Mit *Verbindungsfahrten* werden zwei Servicefahrten miteinander verknüpft, wenn sich die Endhaltestelle der ersten Servicefahrt, mit der Starthaltestelle der anschließenden Servicefahrt voneinander unterscheiden. Sie können als zusätzliche Fahrten eingesetzt werden, um durch größere Freiheitsgrade in der Planung die Fahrzeuganzahl für die Bedienung der Servicefahrten zu reduzieren (vgl. z.B. [Kliewer, 2005]). Die Minimierung der benötigten Fahrzeuganzahl entspricht häufig der primären Zielsetzung. Als weitere Zielsetzungen können eine Minimierung der insgesamt zurückzulegenden Fahrstrecke oder eine Minimierung der Standzeiten in der Umlaufplanung verfolgt werden.

Als zusätzlich Freiheitsgrade oder Einschränkungen für die Planung können weitere Nebenbedingungen in Betracht kommen. So kann es beispielsweise erlaubt sein, dass Fahrzeuge am Ende des Arbeitstages in ein anderes Depot einrücken können, als in das, aus dem sie gestartet sind. Dabei ist auf die korrekte Anzahl an Fahrzeugtypen für den nächsten Einsatztag zu achten. Oder als eine Vorgabe darf jeder Umlauf nur Servicefahrten einer Linie umfassen. Eine solche linienreine Planung soll zum einen den Ressourceneinsatz leichter planbar machen und zum anderen sollen so mögliche Verspätungseffekte bei der Ausführung isoliert auf einer Linie verbleiben.

2.3 Dienstplanung

Im Vergleich zur Umlaufplanung gestaltet sich die (Tages-)Dienstplanung aufgrund gesetzlicher, tariflicher und betrieblicher Vorgaben deutlich komplizierter. In diesem Planungsschritt werden auf Basis des Fahrplanes bzw. auf Basis der in der Umlaufplanung festgelegten Fahrtenfolgen anonyme *(Tages)-Dienste* erstellt und im *Dienstplan* zusammengefasst. Ziel ist eine möglichst kostengünstige Bedienung des Fahrplanes

zu erreichen und zu gewährleisten, dass alle auszuführenden Fahrzeugaktivitäten von einem Fahrer abgedeckt werden und somit durchführbar sind.

Für die Dienstbildung sind verschiedene *Dienstarten* zu berücksichtigen, die sich z.B. bezüglich Dienstdauer und zu Grunde liegender Pausenregel unterscheiden. So kann einem Dienst beispielsweise mindestens eine große Pause von mindestens 30 oder 45 Minuten oder aber mehrere kleine Pausen mit z.B. 15 Minuten Länge zugesprochen werden. In der Regel sind für alle Dienstarten die Lenkzeiten, also die Zeiten zu denen ein Fahrer ein Fahrzeug bewegt, entscheidend zur Pausenbildung. So wird in der Fahrpersonalverordnung[1] gesetzlich festgelegt, wie lange ein Fahrer ein Fahrzeug steuern darf und wann bzw. wie lange durch Pausen unterbrochen werden muss. Die erstellen Dienste sind nur dann gültig, wenn sie einer Dienstart zuzuordnen sind.

Die für Fahrzeuge erstellten Busumläufe und die in der Dienstplanung erstellten Fahrerdienste sind über *Ablösepunkte* miteinander verknüpft. Ablösepunkte repräsentieren eine durch Zeit und Ort vorgegebene Möglichkeit, Fahrerwechsel auf einem Umlauf durchzuführen. Generell besteht für Busfahrer die Möglichkeit das Fahrzeug im Depot oder an bestimmten Haltestellen außerhalb des Depots zu wechseln. Für alle zwischen zwei Ablösepunkten auszuführenden Aktivitäten bzw. Aufgaben sind die Fahrer aus Planungssicht an das Fahrzeug gebunden (und andersrum). Diese kleinste, aus Planungssicht nicht mehr zu unterteilende Folge von Servicefahrten, Wartezeiten und/oder Leerfahrten zwischen zwei Ablösepunkten wird als *Dienstelement* bezeichnet. Eine Folge von solchen Dienstelementen, die ununterbrochen auf einem Fahrzeug ausgeführt wird, wird zu einem *Dienststück* zusammengefasst. Eine Folge von Dienststücken bildet getrennt durch Pausen einen anonymen *(Tages-)Dienst* (siehe auch Abbildung 2.3 zur Erläuterung). In der Regel soll die Anzahl der benötigen Dienste minimiert und somit das benötigte Personal gering gehalten werden. Dabei ist abhängig vom jeweiligen Verkehrsbetrieb häufig zu berücksichtigen, dass Dienste bestimmter Dienstarten bevorzugt erstellt oder vermieden werden sollen.

Die in der Dienstplanung erstellten anonymen Dienste müssen in einer anschließenden *Dienstreihenfolgeplanung* noch zu wöchentlichen oder monatlichen Sequenzen verknüpft und mit dem *Personaleinsatzplan* konkreten Fahrern zugewiesen werden. Bei der Verteilung der Dienste steht eine reine Kostenminimierung nicht allein im Vordergrund, da dafür auch durch die vorgegebenen Ressourceneinsatzpläne recht wenig Freiheitsgrade vorhanden sind. Vielmehr besteht die Zielsetzung darin, die Vergabe von Überstunden zu vermeiden und eine gerechte Verteilung der Arbeitsleistung, sowie die Berücksichtigung individueller Präferenzen für Schichtlagen zu gewährleisten. So müssen beispielsweise die maximale Anzahl an Arbeitstagen hintereinander

[1] siehe [Bundesministerium für Verkehr, Bau- und Wohnungswesen, 2005]

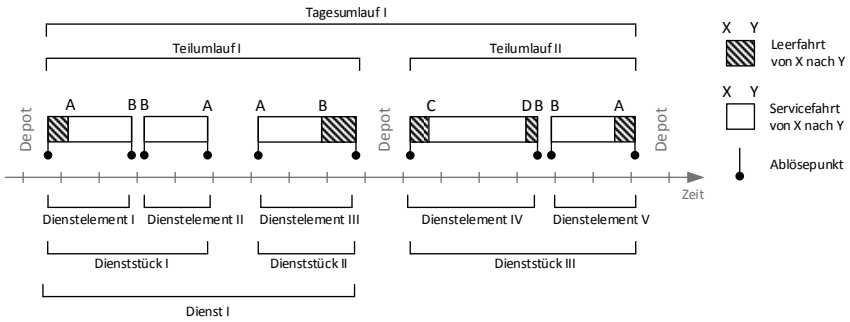

Abbildung 2.3: Grundbegriffe der Umlauf- und Dienstplanung

oder Kompatibiliäten von Schichtlagen in den Personaleinsatzplänen berücksichtigt werden. Für weiterführende Informationen über die Dienstreihenfolgeplanung sei an dieser Stelle auf [Xie und Suhl, 2015] verwiesen. Der Einfluss der Ressourceneinsatzplanung auf die Dienstreihenfolgeplanung wird beispielsweise in [Xie et al., 2012] behandelt.

2.4 Steuerung der Planausführung

Die Steuerung und Überwachung der Planausführung ist in Abhängigkeit der Betriebsgröße und der überwachten und zu synchronisierenden Transportmodi eine höchst komplexe Aufgabe.[2] Sie dient zum einen der Sicherung eines reibungslosen Betriebsablaufes. Zum anderen ermöglicht sie eine Verbesserung der Betriebsplanung. Ziel ist es, eine hohe Verlässlichkeit und somit Attraktivität der durchgeführten Transportdienstleistungen zu erreichen. Darüber hinaus sollen durch ein immer besser auf den bedienten Verkehrsraum abgestimmtes Zusammenspiel zwischen eingesetzen Fahrzeugen und Personal überflüssige Kosten vermieden und einspart werden.

Rechnergestütztes Betriebsleitsystem

Diese Aufgaben werden mit Hilfe eines *rechnergestützten Betriebsleitsytems RBL* umgesetzt, das mittlerweile auch als *Intermodal Transport Control System ITCS* bezeichnet wird. Ein solches System ermöglicht eine (ständige) Erfassung der aktuellen

[2] Beispielhaft sei hier auf den Umfang des täglichen Ressourceneinsatzes einzelner Verkehrsunternehmen in Ballungsräumen wie Berlin oder Hamburg hingewiesen: (Berlin) Die BVG bedient neben 151 Buslinien, 22 Straßenbahnlinien sowie zehn U-Bahnlinien, siehe [Berliner Verkehrsbetriebe AöR, 2017]. (Hamburg) Die Hamburger Hochbahn AG bedient 111 Buslinien mit ca. 800 Fahrzeugen sowie vier U-Bahnlinien, siehe [Hamburger Hochbahn AG, 2017].

Fahrzeugpositionen über beispielsweise Satellitenordnung (wie GPS) und Kommunikation der Fahrzeuge bzw. Fahrer mit der Leitzentrale bzw. einzelnen Disponenten. Durch computergestützten Abgleich der Soll-Daten aus der Planung mit aktuellen Ist-Daten über die Fahrtenlage, sollen so früh wie möglich Abweichungen vom Plan erkannt werden. Beim Erkennen von Störungen können so Maßnahmen ergriffen werden, um potenziell resultierende Verspätungen zu vermeiden.

Umgang mit Störungen des Betriebsablaufes

Störungen des Betriebsablaufes können vielfältiger Natur sein und unterschiedliche Folgen bewirken. Als Beispiele für Störungen seien hier Staus, Straßenarbeiten, Unfälle, schlechte Witterungsbedingungen, ein übermäßiges Passagieraufkommen, sowie Fahrzeugdefekte und Fahrerausfälle genannt. Kleinere Störungen können geringe Abweichungen vom Plan erzeugen und zu (geringen) Verspätungen führen (ca. 5 bis 20 Minuten). Aufgabe im Rahmen eines *Verspätungs- bzw. Störungsmanagements* ist es dann, über Anschlussverbindungen zwischen unterschiedlichen Linien oder zwischen Bus- und Schienenverkehr zu entscheiden. Dabei sollen durch geeignete Maßnahmen eine Propagation von Verspätungen sowie zusätzlich entstehende Kosten für die Bedienung soweit möglich vermieden werden. Große, schwerwiegende Störungen können dahingegen lange Ausfälle in der geplanten Bedienung verursachen. Ziel des Störungsmanagements ist es in solchen Fällen, durch geeignete Ersatzpläne so schnell wie möglich wieder zum geplanten Betriebsablauf zurückkehren zu können und die erforderlichen Zusatzkosten dabei so gering wie möglich zu halten.[3]

Die Folgen von großen Störungen sind für die Passgiere deutlich sicht- und spürbar. Für die Verkehrsunternehmen sind sie bei Fahrtenausfällen unmittelbar mit monetären Verlusten bzw. beim Einsatz von Ersatzfahrzeugen und -fahrern mit zusätzlichen Kosten verbunden. Für ÖPNV-Unternehmen ist allerdings auch schon das Vermeiden von kleinen, alltäglichen Verspätungen bedeutend, um einen reibungslosen Betrieb wie geplant durchführen zu können und die im Verkehrsvertrag mit den Aufgabenträgern vereinbarten Qualitätsmerkmale der Leistungserbringung einhalten zu

[3] Exemplarisch für schwerwiegende Störungen seien hier zwei konkrete, nicht alltägliche Fälle des Berliner Nahverkehrs aufgeführt: (Fall 1) Aufgrund einer unfallbedingten Sperrung und ein durch regnerisches Wetter bedingtes verstärktes Verkehrsaufkommen waren Mitte September 2013 weite Teile des Straßennetzes im Nordosten der Stadt durch einen Rückstau lahmgelegt, so dass einzelne Busse nach Angaben der BVG Spitzenwerte von bis zu 70 Minuten Verspätung an diesem Tag erreicht haben. (Fall 2) Ende des Jahres 2010 gerieten aufgrund technischer Defekte mehrere Busse während ihres Einsatzes in Brand. Neben diesen kurzfristig zu behebenden Ausfällen durch Ersatzfahrzeuge, wurden als Folge mittelfristig weniger Servicefahrten angeboten und Takte ausgedünnt, da Teile der Fahrzeugflotte vorerst aus dem Verkehr gezogen werden mussten.

können. So bieten mittlerweile viele Unternehmen Garantieversprechen bzw. Pünktlichkeitsgarantien an. Liegt eine Verspätung über einer gewissen Toleranzgrenze und (nachweislich) in der Schuld des Unternehmens, werden Entschädigungen an die betroffenen Fahrgäste gezahlt. Beispielsweise gibt es in Berlin aktuell ab 20 Minuten Verspätung den Fahrpreis von den Berliner Verkehrsbetrieben zurück erstattet. Die Nürnberger Verkehrsbetriebe zahlen bereits ab 15 Minuten Verspätung den Fahrpreis zurück, das Nahverkehrsunternehmen der Stadt Essen ab zehn Minuten. Der Nordhessische Verkehrsverbund, der den ÖPNV in und um Kassel bedient, geht noch einen Schritt weiter und zahlt bereits ab fünf Minuten Verspätung eine Entschädigung an seine Fahrgäste.[4] Diese Garantien dienen zum einen einer Service- und somit Attraktivitätssteigerung des ÖPNV. Zum anderen können die Unternehmen durch ihre Kunden so frühzeitig auf mögliche „Konstruktionsfehler“ aufmerksam gemacht werden und den Fahrplan bzw. Ressourceneinsatz anpassen.

Primäre und sekundäre Verspätungen

Die durch Störungen möglicherweise verursachten *Verspätungen* werden in dieser Arbeit in die zwei Kategorien *primär* und *sekundär* eingeteilt. *Primäre* bzw. *ursprüngliche* Verspätungen sind durch das Verkehrsunternehmen in der Regel nicht beeinflussbar und unmittelbare Folge einer Störung. Der Umgang mit den primären Verspätungen wird allerdings durch die erstellten Planstrukturen eingegrenzt und vorgegeben. So können je nach Planstruktur und Verknüpfung der Umlauf- und Dienstpläne primäre Verspätungen auf folgende Fahrten im Transportnetzwerk weitergegeben werden. Es entstehen dann *sekundäre* bzw. *propagierte* Verspätungen.

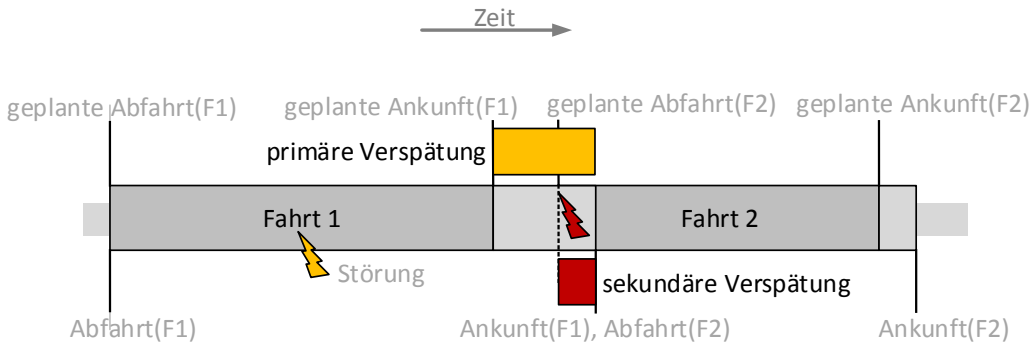

Abbildung 2.4: Zusammenhang von primärer (*gelb*) und sekundärer (*rot*) Verspätung

[4] Diese Angaben finden sich auf den Webseiten der Verkehrsbetriebe. www.bvg.de/de/ Serviceseiten/Kundengarantie, www.vag.de/service-mehr/serviceversprechen.html, www.evag.de/service/puenktlichkeitsversprechen.html, www.nvv.de/service/nvv-5-minuten-garantie (30.3.2017)

Abbildung 2.4 verdeutlicht den Zusammenhang zwischen primären und sekundären Verspätungen: Eine Störung während Fahrt 1 führt zu einer primären Verspätung, die bis zum Ende der Fahrt nicht abgebaut werden kann und somit die tatsächliche Ankunftszeit der Fahrt nach hinten verschiebt. Wird Fahrt 2 durch dieselben Ressourcen wie Fahrt 1 bedient, kann sie nicht mehr zur geplanten Abfahrtszeit starten und ihre Ausführung verschiebt sich ebenfalls nach hinten. Der sich durch diese planimmanenten Strukturen ergebende Unterschied zwischen ursprünglich geplanter und tatsächlicher Abfahrtszeit von Fahrt 2 wird dann als sekundäre Verspätung betrachtet.

Konkrete dispositive Maßnahmen

Bei der Steuerung des Betriebsgeschehens können im ÖPNV eine Reihe *dispositiver Maßnahmen* ergriffen werden, die in verschiedenster Weise Einfluss auf den Ressourceneinsatz und die Propagation von Verspätungen haben. Im Vergleich zum Schienenverkehr können diese Entscheidungen für den Einsatz von Bussen aufgrund fehlender Gleisbindung freier getroffen werden, so dass sich ein größerer Handlungsspielraum für die Disponenten bietet. Die VDV-Schrift 730[5] listet eine Vielzahl spezifischer Dispositionsmaßnahmen im Rahmen eines ITCS auf. Grundsätzlich können die getroffenen Maßnahmen Fahrweg, Fahrplan, Fahrzeug und Fahrer betreffen. Tabelle 2.2 fasst exemplarische Maßnahmen zusammen und stellt ihnen die Auswirkungen auf den Ressourceneinsatz gegenüber.

Tabelle 2.2: Dispositive Maßnahmen und Auswirkungen auf den geplanten Ressourceneinsatz

Betrifft	Maßnahme	Auswirkung
Fahrweg	Umleitung, Kurzwende, Linienteilung	Verlängerung/Verkürzung von Fahrtstrecke und -zeit
Fahrplan	Fahrtausfälle/-änderungen, Verstärker	Vergrößerung/Verringerung der auszuführenden Fahrtenanzahl bzw. -dauer
Fahrzeug	Tausch von Fahrten, Ersatzfahrzeug	Verlängerung/Verkürzung von Umläufen, Änderung der Zuordnung zu Diensten, Erhöhung des Ressourceneinsatzes
Fahrer	Tausch von Aufgaben, Ersatzfahrer	Verlängerung/Verkürzung von Diensten, Änderung der Zuordnung zu Umläufen, Erhöhung des Ressourceneinsatzes

[5] siehe [Verband Deutscher Verkehrsunternehmen VDV, 2011]

Rückfluss der Ausführungs-Daten in die Planung

Je mehr Korrekturmaßnahmen während der Ausführung zu treffen sind und je umfangreicher diese Maßnahmen ausfallen, desto verfälschter werden die ursprünglichen Fahr-, Umlauf- und Dienstpläne und desto eher können die tatsächlich entstehenden Kosten von den geplanten Kosten abweichen. Die mit Hilfe eines ICTS erfassten Ist-Daten können deshalb auch zur Analyse und somit zur Verbesserung des Betriebsablaufes eingesetzt werden, indem die Daten (statistisch) analysiert und dahingehend ausgewertet werden, ob sich die erstellten Pläne im Ablauf bewährt haben. So lassen sich im Nachhinein immer wiederkehrende Störungen, mögliche Anschlussprobleme oder eine besonders hohe Auslastung von einzelner Servicefahrten identifizieren und durch Anpassen von Fahr- und Ressourceneinsatzplänen für die Zukunft reparieren.

2.5 Robuste Planung

Um die im Vorfeld erstellten Pläne während der Ausführung weiterhin am Laufen zu halten bzw. den aktuellen Gegebenheiten anzupassen, kann mit Dispositionsmaßnahmen auf Störungen im Betriebsablauf reagiert werden. Neben diesem reaktiven Vorgehen durch das möglicherweise zusätzliche Kosten entstehen (vgl. Abschnitt 2.4), sollen durch eine *robuste Planung* bereits im Vorfeld Unsicherheiten bei der Ausführung berücksichtigt werden. Dabei können im Idealfall Informationen aus historischen Daten der Ausführung genutzt werden, um Pläne proaktiv so zu erstellen, dass sie unempfindlich gegenüber (zufälligen) Einflüssen sind und trotz (kleinerer) Störungen zu nahezu gleichen Kosten wie geplant ausgeführt werden können.

In Anlehnung an [Scholl, 2001] und der Argumentation von [Dueck, 2010] folgend werden in dieser Arbeit für eine robuste Planung grundsätzlich die zwei Aspekte *Stabilität* und *Flexibilität* in Bezug auf die Planausführung betrachtet. *Stabile* Pläne bleiben während ihrer Ausführung zulässig und erlauben eine trotz (kleinerer) Störungen sichere Durchführung der geplanten Vorgänge ohne jegliche Änderungen und Zusatzkosten. Im ÖPNV kann Stabilität beispielsweise dadurch erreicht werden, dass in den Plänen ausreichend Pufferzeiten zwischen der Bedienung einzelner Servicefahrten gewährt werden, um mögliche Verspätungen zu absorbieren. *Flexible* Pläne dahingegen erlauben eine bestmögliche Anpassung an sich ändernde Umweltbedingungen, wobei auch (geringe) Zusatzkosten für die Plananpassung in Kauf genommen werden. Im ÖPNV kann Flexibilität beispielsweise dadurch erreicht werden, dass in den Plänen Möglichkeiten zum Ressourcentausch vorgesehen sind oder Erweiterungen einzelner Umläufe bzw. Dienste um zusätzliche Aufgaben ohne Schwierigkeit durchführbar sind, um Auswirkungen von Störungen kompensieren zu können. Dabei

ist ein Plan umso flexibler, desto kostengünstiger Störungen durch einfache Maßnahmen ausgeglichen werden können. In welchem Umfang ein robuster Plan die beiden Eigenschaften Stabilität und Flexibilität für seine Ausführung beinhalten kann und sollte, ist Gegenstand aktueller Forschung im Bereich Transport und Verkehr (siehe Kapitel 3).

2.6 IT-Systeme zur Planung und Steuerung

Zur Umsetzung der in den vorangegangenen Abschnitten dieses Kapitels beschriebenen Aufgaben können Verkehrbetriebe auf vielfältige, kommerzielle Softwarlösungen zurückgreifen. Die auf dem Markt befindlichen Systeme zeichnen sich durch einen modularen Aufbau aus. So können je nach Bedürfnis gezielt einzelne Aufgaben unterstützt oder der gesamte Lenkungs- und Leistungserbringungsprozess eines Verkehrsbetriebes abgebildet werden.

Beispiele für umfangreiche Komplettlösungen stellen die Systeme MOBILE der init AG[6], IVU.suite der IVU AG[7] bzw. das System der TRAPEZE Inc.[8] dar. Diese Systeme beinhalten insbesondere Komponenten, die von einer umfassenden Planungsunterstützung bis hin zur Unterstützung der Plandurchführung und Auswertung bzw. Analyse der erbrachten Leistungen reichen. Ausschließlich auf die Planerstellung und Betriebslenkung fokussiert sind beispielsweise die Systeme HASTUS der GIRO Inc.[9] oder MTRAM der MAIOR srl.[10].

Allen genannten Systemen ist gemeinsam, dass sie Komponenten anbieten, die zunehmend Modelle und Methoden der quantitativen Entscheidungsunterstützung für die Planung des Ressourceneinsatzes einsetzen. Auf den Internetseiten und in den Kundenzeitungen der Unternehmen finden sich vereinzelt Fallstudien, die von deutlichen Kosteneinsparungen durch den Einsatz solcher Komponenten berichten. Generell können durch ihren Einsatz in kurzer Zeit verschiedene Möglichkeiten durchgespielt und kosteneffiziente Umlauf- und Dienstpläne berechnet werden (vgl. z.B. [Wren, 2004]). Für eine Übersicht über den Einsatz spezifischer mathematischer Optimierungskomponenten in Planungssystemen für den Ressourceneinsatz sei an dieser Stelle auf [Amberg, 2014] verwiesen.

[6] siehe [INIT innovation in traffic systems AG, 2017]
[7] siehe [IVU Traffic Technologies AG, 2017]
[8] siehe [Trapeze Software Inc., 2017]
[9] siehe [GIRO Inc., 2017]
[10] siehe [MAIOR srl., 2017]

2.7 Zusammenfassung und Implikationen

In diesem Kapitel wurden die Aufgaben und Ziele im Planungsprozess eines ÖPNV-Unternehmens beschrieben, die die Planung des Ressourceneinsatzes mittel- und unmittelbar betreffen (Abschnitte 2.1 bis 2.3). Die Sicht auf diese Aufgaben wurde mit einem Überblick über die Phase der Plandurchführung in Abschnitt 2.4 erweitert. Die im anschließenden Abschnitt 2.5 beschriebene Zielsetzung der Robustheit in der Ausführung eines Planes und zuvor beschriebene Zielsetzung der Kosteneffizienz bei der Umlauf- und Dienstplanung erfordern gegenseitige Abstimmung. So wird ein sehr kosteneffizient geplanter Ressourceneinsatz in der Regel so wenig Fahrzeuge und Fahrer wie möglich beinhalten und kaum Pufferzeiten aufweisen, dafür aber möglicherweise in seiner Ausführung sensibel auf Störungen reagieren und viel Änderungen bedürfen – in Abschnitt 2.6 wurde in diesem Zusammenhang auf eine zunehmende Unterstützung der Planung von kosteneffizienten Umläufen und Diensten durch Methoden des Operations Research eingegangen. Ein sehr robust geplanter Ressourceneinsatz dahingegen wird übermäßig viele Fahrzeuge und Fahrer sowie Pufferzeiten umfassen, wodurch es kaum Änderungen während der Ausführung bedarf, allerdings auch zu hohe Kosten für den Ressourceneinsatz vorgegeben werden. In der Planungsphase ist daher der richtiger Trade-Off zwischen Kosteneffizienz und Robustheit für die Ausführung zu finden. Es stellt sich dabei als wesentliche Frage: Wie kann der Ressourceneinsatz so eingeplant werden, dass der tatsächliche Betrieb kosteneffizient durchgeführt werden kann?

Kapitel 3

(Robuste) Umlauf- und Dienstplanung: Stand der Forschung

In diesem Kapitel wird der aktuelle Stand der Forschung zur Umlauf- und Dienstplanung im ÖPNV sowie zu einer robusten Planung im ÖPNV und verwandten Gebieten diskutiert. Dabei werden unterschiedliche Planungsansätze betrachtet, die von Methoden des Operations Research gestützt werden. Ausgehend von den Planungsansätzen werden Anknüpfungspunkte an existierende Modelle und Lösungsverfahren der mathematischen Optimierung sowie an den Umgang mit Simulation bzw. den Gebrauch von historischen Daten bei der Planung herausgearbeitet. Abbildung 3.1 stellt die Ansätze aus unterschiedlichen Problemdomänen und betrachteten Schwerpunktsetzungen als Fundament und tragende Säulen dieser Arbeit gemeinsam dar.

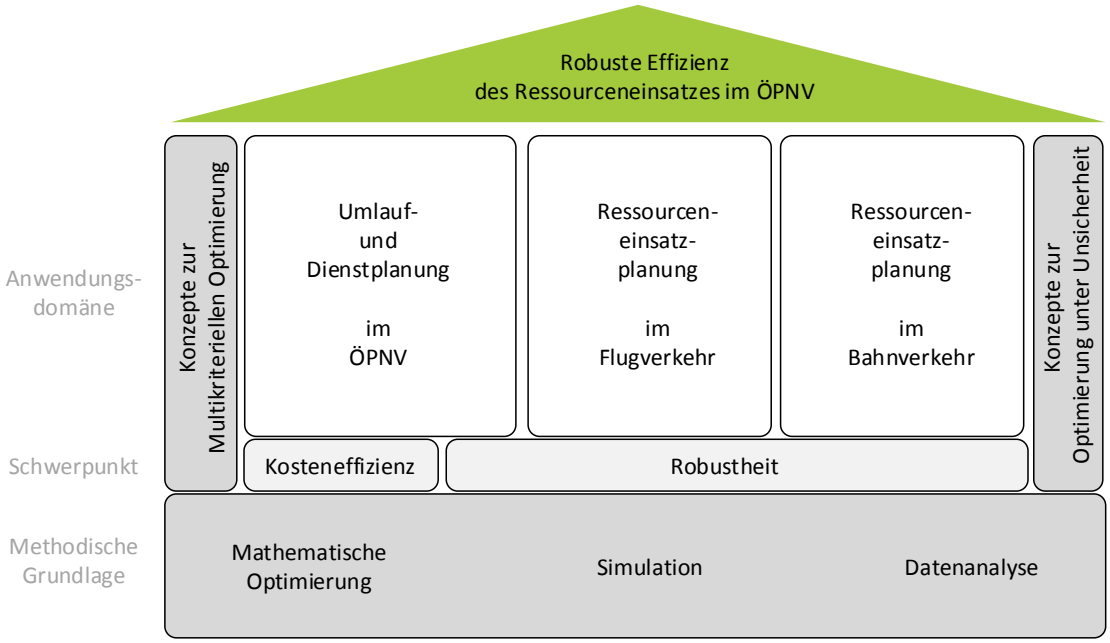

Abbildung 3.1: Grundlagen des Stands der Forschung für Lösungsansätze zur robusten und kosteneffizienten Ressourceneinsatzplanung im ÖPNV

Dieses Kapitel ist wie folgt aufgebaut: In Abschnitt 3.1 werden zunächst methodische Grundlagen der quantitativen Entscheidungsunterstützung vor dem Hintergrund robuster und kosteneffizienter Ressourceneinsatzplanung skizziert. Abschnitt 3.2 gibt einen Überblick über den derzeitigen Stand zur Optimierung in der Umlauf- und Dienstplanung im ÖPNV. Erweitert wird dieser Überblick um den gegenwärtigen Forschungstand zu robuster Planung im ÖPNV in Abschnitt 3.3 und um aktuelle Forschung in verwandten Gebieten wie Luftverkehr oder Bahn(fern)verkehr in Abschnitt 3.4. Abschnitt 3.5 fasst dieses Kapitel zusammen und stellt Potenziale und Grenzen aktueller Forschung zu robuster Planung im Personenverkehr heraus.

3.1 Ausgewählte Konzepte des Operations Research

In diesem Abschnitt werden Grundlagen aus dem Bereich des Operations Research vor dem Hintergrund dieser Arbeit erläutert. Dafür wird zunächst kurz auf mathematische Optimierung und Simulation eingegangen. Anschließend werden grundlegende Konzepte zur Optimierung unter Unsicherheit in Abschnitt 3.1.1 und wesentliche Konzepte zur mehrkriteriellen Optimierung in Abschnitt 3.1.2 erläutert.

Exkurs: Mathematische Optimierung

Mathematische Optimierung bzw. *mathematische Programmierung* bezeichnet die Maximierung oder Minimierung einer Zielfunktion von Entscheidungsvariablen unter Berücksichtigung verschiedener Gleichungs- oder Ungleichungsrestriktionen. *Zielfunktion*, *Variablen* und *Restriktionen* bilden dabei eine (reale) Entscheidungssituation als ein Modell ab. Je nachdem welche Entscheidungssituation abgebildet wird, können die Variablen unterschiedliche Ausprägungen annehmen. Die mögliche Ausprägung wird dabei durch den von den Restriktionen vorgegebenen Lösungsraum begrenzt. Für ein Optimierungsproblem soll in diesem vorgebenen Rahmen eine bestmögliche zulässige Lösung gefunden werden. Sind Zielfunktion und Restriktionen lineare Funktionen der Variablen, spricht man von *Linearer Programmierung*. Dürfen (einzelne) Variablen nur ganzzahlige Werte annehmen, spricht man von *(Gemischt)-Ganzzahliger Programmierung*. Für umfassende Hintergrundinformationen über Lineare bzw. (Gemischt-)Ganzzahlige Programmierung sei an dieser Stelle verwiesen auf [Chvatal, 1983], [Wolsey, 1998] und [Nemhauser und Wolsey, 1988]. Umlauf- und Dienstplanungsprobleme werden in der Regel als gemischt-ganzzahlige Optimierungsprobleme formuliert.

Exkurs: Simulation

Simulation dient als Methode zur experimentellen Analyse von (komplexen) Systemen, indem durch Experimente deren grundlegendes Verhalten untersucht und für die Wirklichkeit berechnet bzw. abgeschätzt wird. So bezeichnet Simulation „das Nachbilden eines Systems mit seinen dynamischen Prozessen in einem experimentierbaren Modell, um zu Erkenntnissen zu gelangen, die auf die Wirklichkeit übertragbar sind. Im weiteren Sinne wird unter Simulation das Vorbereiten, Durchführen und Auswerten gezielter Experimente mit einem *Simulationsmodell* verstanden.“ (siehe [Verein Deutscher Ingenieure VDI, 2013]). Sie wird insbesondere verwendet, wenn rein analytische Methoden aufgrund der Komplexität des im Modell abgebildeten Systems nicht einsetzbar sind, d.h. keine genaue Lösung über das Verhalten theoretisch hergeleitet bzw. rechnerisch ermittelt werden kann.

Sobald Teile des Simulationsmodells dem Zufall bzw. definierten Wahrscheinlichkeiten unterliegen, spricht man von *stochastischer* Simulation. Basierend auf dem Input ist der Output eines Simulationsexperiments dann nicht deterministisch bestimmbar und liefert lediglich eine Abschätzung des möglichen Verhaltens. Bei der *Monte-Carlo Simulation* werden Experimente basierend auf Zufallszahlen durchgeführt. Mit zunehmender Anzahl der voneinander unabhängigen Experimente nähert sich der arithmetische Mittelwert aus den erzeugten Zufallszahlen nach dem „Gesetz der großen Zahlen“ dem Erwartungswert an. So kann eine Annäherung an das wahrscheinliche Verhalten des abgebildeten Systems in der Realität hergestellt werden. Für weiterführende Informationen über Simulation sei an dieser Stelle auf [Law, 2014] verwiesen. Eine grundlegende Einführung in den Umgang mit Monte-Carlo Methoden wird von [Hammersley und Handscomb, 1964] gegeben. [Gentle, 2003] setzt sich vertiefend mit der Generierung von Zufallszahlen bei Anwendung von Monte-Carlo Methoden auseinander.

Für die Umlauf- und Dienstplanung ermöglicht Simulation das Durchspielen verschiedener Realitäten, die während der Ausführung der erstellten Pläne eintreffen können. So können die Umlauf- und Dienstpläne beispielsweise bezüglich ihrer Robustheit im Fall von unterschiedlich auftretenden Störungen bzw. primären Verspätungen evaluiert und miteinander verglichen werden.

3.1.1 Optimierung unter Unsicherheit

Bei vielen Entscheidungsproblemen kann zum Zeitpunkt der Entscheidungsfindung eine gewisse *Unsicherheit* bezüglich einiger Eingangsdaten vorliegen. Um mit dieser Unsicherheit im Rahmen der Entscheidungsfindung umgehen zu können, existieren

verschiedene Konzepte, von denen wesentliche im Nachfolgenden mit Blick auf die Ausrichtung dieser Arbeit kurz skizziert werden. In der Regel basieren diese Konzepte darauf, die vorhandene Unsicherheit mit Hilfe von *Szenarien* abzubilden. Solche Szenarien repräsentieren jeweils unterschiedliche Ausprägungen der betrachteten Entscheidungssituation in einer möglichen Zukunft. Je nach Ansatz sollte eine Lösung alle oder einen Teil der potenziellen Szenarien erfüllen bzw. an nicht erfüllte Szenarien leicht anpassbar sein. Darüber hinaus können unterschiedliche Wahrscheinlichkeiten für das Eintreffen einzelner Szenarien berücksichtigt werden.

Für den Fall der Umlauf- und Dienstplanung im ÖPNV können Unsicherheiten beispielsweise darin bestehen, zu welchem Zeitpunkt und in welchem Umfang während der Planausführung Störungen auftreten. Diese Störungen könnten eine Verlängerung einzelner Fahrten bedingen (vgl. Abschnitt 2.4). Mögliche Szenarien könnten dementsprechend die Dauer von Störungen bzw. primäre Verspätungen für einzelne Fahrten beinhalten.

Robust Optimization

Bei diesem Vorgehen wird eine Lösung gesucht, die ohne Änderungen für alle möglichen Szenarien ausführbar ist, d.h. auch dann wenn die ungünstigste Ausprägung der zu berücksichtigenden Unsicherheit eintritt. Das Grundkonzept basiert auf der Idee einer strikten Robustheit, die von [Soyster, 1973] eingeführt und von [Ben-Tal et al., 2009] ausführlich untersucht wird. Dabei wird eine Lösung gesucht, die den Worst-Case Zielfunktionswert über alle Szenarien minimieren kann. Ein aktueller Überblick über die Entwicklung verschiedener Konzepte von *Robust Optimization* in Theorie und praktischen Anwendungsgebieten wird in [Gabrel et al., 2014] gegeben. Darüber hinaus stellen [Ide und Schöbel, 2015] Robustheitskonzepte für die Berücksichtigung mehrfacher Unsicherheit bei Optimierungsproblemen mit mehrfachen Zielkriterien zusammmen.

Übertragen auf die Umlauf- und Dienstplanung: Die erstellten Pläne sollen auch bei Auftreten des Worst-Cases, also der größten bekannten Verspätung je Servicefahrt während der Ausführung zulässig bleiben und idealerweise keine Verspätungen propagieren.

Stochastic Programming

Es wird eine Lösung gesucht, die mit einer bestimmten Wahrscheinlichkeit für die bei der Lösungsfindung einbezogenen Szenarien zulässig ist. Dabei bildet die Zielfunktion in der Regel einen Erwartungswert über die Szenarien ab. Eine wichtige Annahme

ist, dass die Wahrscheinlichkeit für das Eintreffen jedes Szenarios als bekannt vorausgesetzt wird. Das Vorgehen ist rechentechnisch anspruchsvoll, da jedes Szenario im Lösungsprozess einzeln abzubilden bzw. zu betrachten ist. Für weiterführende Informationen über verschiedene Ausprägungen von *Stochastic Programming* sei an dieser Stelle auf [Birge und Louveaux, 1997] verwiesen.

Übertragen auf die Umlauf- und Dienstplanung: Die Pläne werden so erstellt, dass sie die erwartete Verspätungspropagation über alle Szenarien minimieren. Die potenziellen Verspätungen einzelner Fahrten werden dabei über Wahrscheinlichkeitsverteilungen abgebildet.

Recoverable Robustness

Bei diesem Vorgehen wird der Fokus auf das durch einfache Mittel erreichbare Wiederherstellen der Zulässigkeit von Lösungen gelegt. Nach [Liebchen et al., 2009] wird eine Lösung des ursprünglichen Optimierungsproblem dann als *recovery-robust* bezeichnet, wenn die Rückkehr zu dieser Lösung mit einem der zulässigen Algorithmen in allen spezifizierten Störungsszenarien erreicht werden kann. Im Lösungsprozess sind dazu sowohl Planungs- als auch zugehörige Recovery-Probleme gemeinsam zu betrachten.

Übertragen auf die Umlauf- und Dienstplanung: Davon ausgehend, dass der geplante Ressourceneinsatz aufgrund von schwerwiegenden Störungen häufig nicht eingehalten werden kann, werden Reaktionsmöglichkeiten auf Fahrtenausfälle bzw. Verspätungen in die Planung einbezogen. So kann z.B. das Bereitstellen und der Einsatz von Ersatzressourcen dazu beitragen, im Fall einer Störung mögliche Verspätungspropagation aufzufangen und wieder zum ursprünglichen Plan zurückzukehren.

3.1.2 Optimierung mit mehreren Zielkriterien

Bei der *Mehrkriteriellen Optimierung* bzw. *Mehrzieloptimierung* werden zwei oder mehrere, konkurrierende Ziele gleichzeitig verfolgt. Im Gegensatz zur Einzieloptimierung ist es hier nicht allein von Interesse, Lösungen mit optimalen Werten für eine einzelne Zielfunktion zu finden. Vielmehr sollen Lösungen gefunden werden, die *Pareto-optimal* sind. Eine Lösung wird dann als *Pareto-optimal* oder *effizient* bezeichnet, wenn es keine andere Lösung gibt, die in allen betrachteten Zielfunktionen nicht mindestens genauso gut und in mindestens einer Zielfunktion besser ist (siehe [Ehrgott, 2005]). Mehrzieloptimierungsprobleme haben also in der Regel eine Menge effizienter Lösungen, die den Trade-Off zwischen den unterschiedlichen Zielkriterien abbilden. Die Lösungen können für die Mehrzieloptimierung darüber miteiander verglichen,

ob sie einander *dominieren*. Eine Lösung dominiert eine andere, wenn sie in allen Zielfunktionswerten mindestens gleich gut und in mindestens einem besser ist.

Bezogen auf die Umlauf- und Dienstplanung im ÖPNV kann durch die nachfolgend beschriebenen Konzepte beispielsweise der Trade-Off zwischen Robustheit auf der einen und Kosteneffizienz auf der anderen Seite abgebildet werden.

Gewichtete Summe der Zielfunktionen

Bei diesem Vorgehen werden die zu betrachtenden Zielfunktionen f_i (mit $i \in \mathbb{N}$) zu einer einzelnen Zielfunktion aggregiert, wobei die Restriktionen der Problemstellung unangetastet bleiben. Die Zielfunktionen werden dafür mit Gewichten g_i normiert und zusammen addiert (vgl. Formulierung (3.1)). Dafür müssen im Vorfeld allerdings Präferenzen bezüglich der Gewichtung festgelegt werden. Auch sollten die Maßeinheiten der Zielsetzungen miteinander vergleichbar sein.

$$\sum_i g_i f_i(x) \rightarrow \text{Min} \tag{3.1}$$

Übertragen auf die Umlauf- und Dienstplanung: In der Zielfunktion wird beispielsweise zum einen die in Minuten gemessene, geplante Arbeitszeit des Ressourceneinsatzes minimiert. Zum anderen werden gleichzeitig möglicherweise auftretende Verspätungen minimiert, indem die potenziellen Verspätungsminuten stärker, schwächer oder gleich gewichtet als zusätzliche Arbeitszeitkosten in der Zielfunktion addiert werden.

ϵ-Constraint Methode bzw. Anteile von Zielen

Von den zu betrachtenden Zielfunktionen f_i (mit $i \in \mathbb{N}$) wird eine Zielfunktion f_j ausgewählt, deren Zielfunktionswert optimiert werden soll. Die anderen Zielfunktionen werden zu Nebenbedingungen umformuliert, indem obere oder untere Schranken ϵ_i für ihren Zielfunktionswert vorgegeben werden (vgl. Formulierung (3.2)-(3.3)). Der Wert ϵ_i einer Nebenbedingung kann dabei z.B. so vorgegeben werden, dass die repräsentierte Zielsetzung zu bestimmten Anteilen mindestens oder maximal erreicht wird. Um eine Rangfolge von Zielen zu berücksichtigen, kann zunächst ein Optimierungsmodell mit dem wichtigsten Ziel gelöst werden. Der erlangte Zielfunktionswert wird gemäß des zu erreichenden Anteils angepasst und als Restriktion für das nächstwichtigste Ziel formuliert (u.s.w.). Im Vergleich zur Formulierung einer gewichteten Summe von Zielfunktionen ist das Vorgehen rechentechnisch aufwändiger, da je nach Anzahl an

Zielsetzungen mehrere Optimierungsprobleme zu lösen sind. Die ϵ-Constraint Methode kann mit der Methode der gewichteten Summe kombiniert werden.

$$f_j(x) \rightarrow \text{Min} \tag{3.2}$$

$$\text{s.t.} f_i(x) \leq \epsilon_i \; \forall i, i \neq j \tag{3.3}$$

Übertragen auf die Umlauf- und Dienstplanung: Minimiere beispielsweise die potenziellen Verspätungszeiten unter der Nebenbedingung, dass maximal 105% der geplanten Kosten einer kostenoptimalen Lösung eingesetzt werden.

Elastic Constraint Methode

Diese Methode ist eine Erweiterung der ϵ-Constraint Methode, damit für das betrachtete Optimierungsproblem ganzzahlige Lösungen rechnerisch einfacher bestimmbar sind (siehe [Ehrgott, 2005]). Abweichungen s_i von als Nebenbedingungen formulierten Zielstellungen f_i (mit $i \in \mathbb{N}$) werden mit einem Bestrafungsfaktor p_i versehen und in die Zielfunktion eingebracht (vgl. Formulierung (3.4) - (3.6)). So werden einzelne Zielvorgaben nicht als harte Vorgaben erzwungen, sondern können zum Finden einer guten Lösung aufgeweicht werden.

$$f_j(x) + \sum p_i s_i \rightarrow \text{Min} \tag{3.4}$$

$$\text{s.t.} f_i(x) - s_i \leq \epsilon_i \; \forall i, i \neq j \tag{3.5}$$

$$s_i \geq 0 \tag{3.6}$$

Übertragen auf die Umlauf- und Dienstplanung: Eine strikte Vorgabe der maximal für den Ressourceneinsatz einzusetzenden Kosten muss nicht zwangsläufig eingehalten werden und wird zugunsten einer Zielsetzung nach Robustheit als weiche Vorgabe innerhalb eines tolerierbarer Rahmens betrachtet.

3.2 Kosteneffiziente Umlauf- und Dienstplanung

Seit mehreren Jahrzehnten wird bereits an mathematischen Modellen und Methoden zur Unterstützung der Ressourceneinsatzplanung im ÖPNV geforscht. Dazu werden die verschiedenen Planungsaufgaben als mathematische Optimierungsprobleme abstrahiert. Die entwickelten Ansätze lassen sich grob in Arbeiten zur Lösung von *Umlaufplanungsproblemen*, *Dienstplanungsproblemen* und *kombinierten Umlauf- und Dienstplanungsproblemen* einteilen. Je nach betrachteter Problemstellung sind sie un-

terschiedlich schwer zu lösen und erfordern verschiedene Lösungsverfahren. Im Folgenden wird ein kurzer Überblick über die genannten Planungsprobleme gegeben, sowie über etablierte Lösungsverfahren und Modellierungstechniken wesentlicher Arbeiten zur Umlaufplanung in Abschnitt 3.2.1 bzw. Umlauf- und Dienstplanung in Abschnitt 3.2.2 mit der Zielsetzung kosteneffizient zu planen. Für eine ausführliche Zusammenfassung der Entwicklung von Modellen und Methoden für eine reine Dienstplanung sei an dieser Stelle auf [Steinzen, 2007] und [Gintner, 2008] verwiesen. Eine umfassende Behandlung praktischer Anforderungen in der Dienstbildung findet sich in [Amberg, 2014], eine umfassende Behandlung praktischer Randbedingungen der Umlaufplanung beispielsweise in [Kliewer, 2005]. Eine ausführliche Zusammenstellung gegenwärtiger Forschung zu Optimierungsansätzen für die verschiedenen Planungsprobleme im ÖPNV geben [Ibarra-Rojas et al., 2015].

Umlaufplanungsproblem

Die Planungsaufgabe zur Umlaufplanung ist im sogenannten *Umlaufplanungsproblem* (engl. *vehicle scheduling problem, VSP*) abgebildet. Dabei wird für gegebene Servicefahrten eine Zuordnung zu Fahrzeugumläufen gesucht, die die Gesamtkosten des Fahrzeugeinsatzes unter Berücksichtigung folgender Restriktionen minimiert: (1) Jede Servicefahrt ist genau einem Fahrzeugumlauf zugeordnet. (2) Alle Fahrten eines Fahrzeugumlaufes sind von einem Fahrzeug nacheinander ohne zeitliche Überschneidungen ausführbar. (3) Jeder Fahrzeugumlauf endet in dem Depot, aus dem er auch gestartet ist.

Falls es genau ein Depot gibt, d.h. falls nur ein Depotstandort bzw. Abstellort existiert und die Fahrzeuge alle vom identischen Fahrzeugtyp sind (vgl. Abschnitt 2.2), handelt es sich um ein *Eindepot-Umlaufplanungsproblem* (engl. *single-depot vehicle scheduling problem, SDVSP*). Existieren mehrere Depotstandorte und/oder Fahrzeugtypen, handelt es sich um ein *Mehrdepot-Umlaufplanungsproblem* (engl. *multiple-depot vehicle scheduling problem, MDVSP*). Beim MDVSP wird jedes Fahrzeug genau einem Depot zugeordnet und für die Servicefahrten kann eine Einschränkung darin bestehen, dass sie nur von Fahrzeugen aus einem bestimmten Depot bedient werden dürfen.

Das MDVSP ist im Vergleich zum SDVSP $\mathcal{NP}$-schwer (siehe [Bertossi et al., 1987]). Das SVSP ist in polynomieller Laufzeit lösbar und wird vielfach als Unterproblem in den Planungsansätzen behandelt.

Dienstplanungsproblem

Planungsaufgaben zur Dienstplanung sind im sogenannten *Dienstplanungsproblem* (engl. *crew scheduling problem, CSP*) abgebildet. Beim CSP wird für Dienstelemente eine Zuordnung zu Diensten gesucht, die die Personalkosten minimiert und folgende Restriktionen erfüllt: (1) Jedes Dienstelement ist in genau einem Dienst enthalten. (2) Jeder Dienst erfüllt die gesetzlichen, tariflichen und betrieblichen Vorgaben.

Das CSP zählt nach [Fischetti et al., 1987] bzw. [Fischetti et al., 1989] bereits dann zu den $\mathcal{NP}$-schweren Problemen, wenn zur Dienstbildung als einzige Regel eine Beschränkung bei der Dienstdauer oder Arbeitszeit zu beachten ist.

Integriertes Umlauf- und Dienstplanungsproblem

Das *integrierte Umlauf- und Dienstplanungsproblem* (engl. *integrated vehicle and crew scheduling problem, IVCSP*) kombiniert Umlauf- und Dienstplanungsproblem. Gesucht werden Fahrzeugumläufe und Dienste, die die Gesamtkosten aus Fahrzeugeinsatz- und Personalkosten minimieren und dabei folgende Restriktionen einhalten: (1) Umläufe und Dienste bilden einen zulässigen Umlauf- bzw. Dienstplan nach den oben beschriebenen Anforderungen. (2) Umlauf- und Dienstplan sind zueinander kompatibel, d.h. jede Servicefahrt und jede Leerfahrt eines Umlaufes ist durch genau einen Dienst abgedeckt und die Dienste enthalten keine Leerfahrten, die nicht in den Umläufen vorgesehen sind.

Da MDVSP und CSP für sich genommen bereits $\mathcal{NP}$-schwere Probleme sind, ist auch das (MD-)IVCSP ein $\mathcal{NP}$-schweres Problem.

3.2.1 Lösungsansätze zur Umlaufplanung

Lösungsansätze zur Umlaufplanung können grundsätzlich danach unterschieden werden, ob sie einen *Mehrdepotfall* oder einen *Eindepotfall* betrachten. Übersichten über Ansätze zum Eindepotfall werden in [Daduna und Paixão, 1995] und [Bunte und Kliewer, 2009] gegeben. [Bunte und Kliewer, 2009] stellen außerdem Modellformulierungen für Mehrdepot-Umlaufplanungsprobleme gegenüber mit den in der Literatur üblichen Kategorisierungen der Formulierungen nach *Eingüterfluss-Problem*, *Mehrgüterfluss-Problem* (vgl. [Ahuja et al., 1993]) und *Set-Partioning Problem* (vgl. [Balas und Padberg, 1976]).

[Kliewer, 2005] bzw. [Kliewer et al., 2006] schlagen für die Modellierung als Mehrgüterfluss-Problem eine Formulierung vor, die eine Schwierigkeit bei der Formulierung von Umlaufplanungsproblemen behebt und auch größere Probleminstanzen mit einem MIP-Solver lösbar macht. Statt alle möglichen Leerfahrtverbindungen zwischen den

Servicefahrten explizit in einem sogenannten *connection-basierten* Netzwerk abzubilden, werden die Leerfahrtverbindungen implizit in einem sogenannten *Time-Space-Netzwerk (TSN)* modelliert, indem Verbindungen zwischen Gruppen von nacheinander ausführbaren Fahrten aggregiert werden. Dadurch enthält die Modellierung nur noch einen Bruchteil an Entscheidungsvariablen anderer Formulierungen und muss nicht mehr heuristisch gelöst werden. Darüber hinaus beschreibt eine zulässige Flusslösung der Formulierung ein Bündel von Umlaufplänen. Die verschiedenen Pläne eines solchen Bündels ergeben sich durch unterschiedliche Dekomposition der Flusslösung. Diese Eigenschaft wird beispielsweise von [Kliewer et al., 2008] benutzt, um die Flussdekomposition von einer zweiten Zielsetzung abhängig durchzuführen. So wird in dem Fall die Linien-Homogenität der Umläufe als Zielsetzung bei der Dekomposition berücksichtigt.

Neben den exakten Ansätzen existieren eine Reihe heuristischer Ansätze zur Lösung von Mehrdepot-Umlaufplanungsproblemen. Für Übersichten über solche heuristischen Lösungsansätze sei beispielsweise auf [Dell'Amico et al., 1993] und [Pepin et al., 2009] verwiesen.

Implikationen für diese Arbeit

Neben der Reduktion der Problemgröße bietet die TSN-Formulierung als Mehrgüterfluss-Problem nach [Kliewer et al., 2006] den Vorteil, dass die resultierende Flusslösung keinen finalen Umlaufplan darstellt, sondern mehrere gleichwertige Umlaufpläne beinhaltet. So kann durch eine geeignete Dekomposition des Lösungsflusses aus einer Menge möglicher (kostenoptimaler) Umlaufpläne derjenige Plan ausgewählt werden, der neben Kosten auch noch vor dem Hintergrund einer anderen Zielsetzung gut bewertet werden kann, z.B. mit einer möglichst geringen Propagation von Verspätungen. Dies lässt die Verwendung einer derartigen Modellierung bei der Formulierung von Umlaufplanungsproblemen im Rahmen dieser Arbeit zweckmäßig erscheinen. Demgegenüber steht der Nachteil, dass aufgrund der impliziten Modellierung von Verbindungen zwischen den Servicefahrten, keine Informationen abgebildet werden können, die die Verknüpfung zweier bestimmter Fahrten betreffen. Sollen bei der Erstellung konkreter Umläufe Informationen über zuvor auszuführende Fahrten mitberücksichtigt werden (z.B. um die Propagation von Verspätungen abzubilden) bedarf es einer Formulierung, welche die Verknüpfung von Fahrten explizit betrachtet. Dies ist beispielsweise im Rahmen einer pfadbasierten Formulierung als Set-Partitioning-Problem erreichbar. Um die Vorteile beider Formulierungen nutzen zu können, sollte daher in dieser Arbeit eine kombinierte Formulierung entwickelt werden.

3.2.2 Lösungsansätze zu kombinierter Umlauf- und Dienstplanung

Neben einer stufenweisen, sequenziellen Vorgehensweise bei der Bearbeitung der Planungsaufgaben Umlauf- und Dienstplanung wurden Lösungsansätze entwickelt, bei denen beide Planungsaufgaben kombiniert betrachtet werden, um insbesondere mehr Freiheitsgrade beim Lösen der entsprechenden Dienstplanungsprobleme nutzen zu können. Im Folgenden sollen kurz wesentliche Lösungsansätze zur partiellen Integration von Umlauf- und Dienstplanung und zur vollständigen Integration vorgestellt sowie Implikationen für diese Arbeit herausgearbeitet werden. Für eine umfangreiche Beschreibung von Lösungsansätzen zur integrierten Planung sei u.a. auf [Huisman, 2004], [Steinzen, 2007], [Weider, 2007] und [Gintner, 2008] verwiesen.

Partielle Integration von Umlauf- und Dienstplanung

Partiell integrierte bzw. teilintegrierte Lösungsansätze zeichnen sich dadurch aus, dass in den Modellen und/oder Lösungsverfahren zur Umlauf- und Dienstplanung Aspekte der jeweils nachfolgenden Planungsstufe einbezogen werden. Je nach Reihenfolge der Planungsstufen kann bei den Ansätzen zwischen *Dienstplanungsbasierten-* und *Umlaufplanungsbasierten Vorgehensweisen* unterschieden werden.

[Ball et al., 1983] formulieren ein Modell für eine gleichzeitige Umlauf- und Dienstplanung als ein Partitionierungsproblem auf einem gemeinsamen Planungsgraphen. Aufgrund der damaligen Technik (Optimierungsverfahren und Hardware) kann das Modell allerdings nicht integriert gelöst werden, sondern wird heuristisch partiell integriert gelöst, indem schrittweise Dienststücke erstellt, verbessert und zu Diensten zusammengefasst werden. Dabei kann aus der aktuellen Lösung jeweils ein kompatibler Umlaufplan abgeleitet werden. Die grundlegende Formulierung als Planungsnetzwerk wird von mehreren Autoren wie z.B. [Falkner und Ryan, 1992] und [Patrikalakis und Xerocostas, 1992] aufgegriffen, um ähnliche partiell integrierte Ansätze zu modellieren.

[Borndörfer et al., 2002] beschreiben mehrere Möglichkeiten, Aspekte der Dienstplanung bereits bei der Umlaufplanung zu berücksichtigen. So kann z.B. eine Verringerung der Kosten für Aus- und Einrückfahrten dafür sorgen, dass Umlaufplanungslösungen mehr Depotaufenthalte enhalten und so durch mehr Ablösemöglichkeiten für Dienste mehr Freiheitsgrade für die Dienstbildung entstehen. Weitere Möglichkeiten sind höhere Kosten für lange Leerfahrten, um diese zu vermeiden und so einfacher zu planende Dienste erstellen zu können und eine Beschränkung der Umlauflänge auf die maximale Dienstdauer, um sicherzustellen, dass die gebildeten Umläufe von

Diensten bedient werden können. Dahingegen erweitert [Gintner, 2008] die Umlaufplanung nicht mit Aspekten der Dienstplanung, um Freiheitsgrade für die anschließende Dienstplanung zu erhöhen, sondern nutzt die Vorteile der TSN-Formulierung des Umlaufplanungsproblems nach [Kliewer et al., 2006]. Da die Flusslösung des Umlaufplanungsproblems implizit ein Bündel mehrerer gleichwertiger Umlaufpläne repräsentiert, werden Dienste basierend auf der Flusslösung erstellt. Ein kompatibler Umlaufplan wird dann basierend auf dem erstellten Dienstplan gebildet.

Vollständige Integration von Umlauf- und Dienstplanung

Im Gegensatz zu partiell integrierten Ansätzen werden bei vollständig integrierten Ansätzen zur Umlauf- und Dienstplanung sowohl im Modell als auch im Lösungsverfahren Umlauf- und Dienstplanung gleichzeitig betrachtet. Wie bei der Umlaufplanung kann bei den Ansätzen grundsätzlich unterschieden werden, ob sie den Eindepotfall oder den Mehrdepotfall behandeln.

Der erste Ansatz zur vollständig integrierten Umlauf- und Dienstplanung wird von [Freling, 1997] für den Eindepotfall vorgeschlagen und bildet die Basis für die meisten danach entwickelten Ansätze zur integrierten Planung. Im Modell ist der Umlaufplanungsteil zur Zuordnung von Servicefahrten als Quasi-Assignment-Problem formuliert und der Dienstplanungsteil zur Abdeckung aller Dienstelemente bzw. Servicefahrten als Set-Partitioning-Problem. Gekoppelt sind Umlauf- und Dienstplanung dadurch, dass Leerfahrten eines Fahrzeuges auch von einem Dienst bedient werden müssen. Als Lösungsverfahren wird eine Kombination aus Lagrange-Relaxation und Column Generation vorgeschlagen. In einer ersten Phase wird zunächst die Kopplung zwischen Umlauf- und Dienstplanung relaxiert, um eine untere Schranke für die Lösung des Problems bestimmen zu können und in einer zweiten Phase werden dann mit Hilfe von Heuristiken zulässige Lösungen, d.h. kompatible Umlauf- und Dienstpläne, bestimmt. Pricing-Probleme zur Dienstgenerierung sind dabei als ressourcenbeschränkte Kürzeste-Wege-Probleme auf Basis von Dienststücken formuliert.

[Friberg und Haase, 1999] schlagen einen Ansatz für den Eindepotfall vor, bei welchem Umlaufplanungs- und Dienstplanungsteil jeweils als Set-Partitioning-Problem formuliert sind. Eine Kopplung zwischen den beiden Teilen wird dadurch erreicht, dass die Anzahl an Fahrzeugen und Diensten zur Bedienung von Leerfahrten vor bzw. nach Servicefahrten gleich sein muss. Gelöst wird das Modell mit einer exakten Branch-and-Price-and-Cut-Methode.

[Haase et al., 2001] formulieren für den Eindepotfall ein gemeinsames Set-Partitioning-Problem für Umlauf- und Dienstplanung mit zusätzlichen Nebenbedingungen, die sicherstellen, dass aus jeder Dienstplanungslösung ein zulässiger Umlaufplan ab-

geleitet werden kann. Als Lösungsverfahren wird eine Branch-and-Price-Methode verwendet, wobei die Pricing-Probleme zur Dienstgenerierung als ressourcenbeschränkte Kürzeste-Wege-Probleme formuliert sind.

[Borndörfer et al., 2002] schlagen für den Eindepotfall ein Modell vor, in welchem das Umlaufplanungsproblem als Netzwerkflussproblem formuliert ist und der Dienstplanungsteil als Set-Partitioning-Problem. Beide Teile sind dadurch verbunden, dass jede Leerfahrt von der gleichen Anzahl Fahrzeugen und Diensten abgedeckt werden muss. Gelöst wird das Modell mit einer Kombination aus Column Generation und Lagrange-Relaxation. Zur Bestimmung zulässiger Lösungen werden Heuristiken entwickelt.

Das erste generelle Modell zur vollständig integrierten Umlauf- und Dienstplanung im Mehrdepotfall wird von [Huisman, 2004] bzw. [Huisman et al., 2005] als Erweiterung des Eindepotfalls von [Freling, 1997] vorgeschlagen. Für den Umlaufplanungsteil wird eine Mehrgüter-Fluss-Formulierung gewählt, während das Dienstplanungsproblem wie vorher als Set-Partitioning-Problem formuliert ist. Gekoppelt sind beide Teile durch Restriktionen, die sicherstellen, dass im Dienstplan die Leerfahrten aus dem Umlaufplan bedient werden. Gelöst wird das Modell mit einer Anpassung des von [Freling, 1997] vorgeschlagenen Verfahrens.

[Borndörfer et al., 2004] greifen das von [Huisman, 2004] entwickelte Modell auf und lösen es mit einem modifizierten Variante des von [Borndörfer et al., 2002] vorgeschlagenen Lösungsverfahrens für den Eindepotfall. Der Lösungsansatz wird von [Weider, 2007] bzw. [Borndörfer et al., 2008] durch eine sogenannte „Rapid Branching“-Heuristik für die Bestimmung ganzzahliger Lösungen erweitert, mit der beim Branching im Rahmen des Branch-and-Bound-Verfahrens mehrere Variablen gleichzeitig fixiert werden. Mit dem Ansatz werden künstliche und reale Probleminstanzen mit bis zu 1.414 Servicefahrten für den Regionalverkehr und bis zu 634 Servicefahrten für den Stadtverkehr gelöst.

[Gintner, 2008] schlägt ein Modell für die integrierte Planung mit mehreren Depots vor, bei welchem der Umlaufplanungsteil basierend auf der TSN-Modellierung von [Kliewer, 2005] formuliert ist und der Dienstplanungsteil als Set-Partitioning-Problem. Durch die TSN-Formulierung des Umlaufplanungsteils werden deutlich weniger Entscheidungsvariablen und Restriktionen benötigt als in den bisher in der Literatur vorgeschlagenen Ansätzen. Als Lösungsverfahren wird eine Erweiterung des von [Freling, 1997] und [Huisman, 2004] entwickelten Verfahrens vorgeschlagen.

[Steinzen, 2007] bzw. [Steinzen et al., 2010] schlagen weitere Ansätze für den von [Gintner, 2008] verwendeten Lösungsansatz vor. Für die Column Generation Pricing-Probleme werden aus Dienststücken bestehende Dienstgenerierungsnetzwer-

ke modelliert, die auf einer Time-Space-Netzwerkstruktur basieren. Die entstehenden ressourcenbeschränkten Kürzeste-Wege-Probleme werden mit einem Algorithmus auf Basis Dynamischer Programmierung gelöst. Darüber hinaus werden zur Erstellung zulässiger Umlauf- und Dienstpläne unterschiedliche Strategien entwickelt. Die vorgeschlagenen Lösungsverfahren werden auf künstlichen und realen Probleminstanzen mit über 640 Servicefahrten und vier Depots getestet.

Vollständige Integration mit Zeitfenstern als zusätzliche Freiheitsgrade

[Kliewer et al., 2012] schlagen eine Erweiterung des vollständig integrierten Ansatzes von [Steinzen, 2007] vor, indem durch Verschiebemöglichkeiten von Servicefahrten in bestimmten Zeitintervallen zusätzliche Freiheitsgrade für die Planung von Umläufen und Diensten geschaffen werden. Verglichen mit vollständig integrierten Ansätzen ohne Zeitfenster können durch die zusätzlichen Freiheitsgrade noch effizientere Umlauf- und Dienstpläne erstellt werden – auf künstlichen und realen Instanzen mit bis zu 661 Servicefahrten und vier Depots können bei Zeitfenstern von einer bis vier Minuten bis zu 5% mehr an Fahrzeugen und Diensten eingespart werden, verglichen mit sequenziellen Ansätzen sogar bis zu 30%. Im integrierten Modell sind dabei im als TSN-Modell formulierten Umlaufplanungsteil mögliche Verschiebungen von Servicefahrten explizit über zusätzliche Kanten abgebildet. Gelöst wird das Modell mit einer Kombination aus Lagrange-Relaxation und Column Generation. Die Pricing-Probleme zur Dienstgenerierung sind auf Basis von Dienstelementen als ressourcenbeschränkte Kürzeste-Wege-Probleme formuliert, wodurch auch größere Planungsprobleme in akzeptabler Laufzeit gelöst werden können. Da die Problemkomplexität mit Einbezug von Zeitfenstern stark ansteigt, werden außerdem Heuristiken vorgeschlagen, um durch Auswahl von Servicefahrten, deren Verschiebung zu einer potenziell besseren Lösung führen kann, die Komplexität etwas zu reduzieren und die Rechenzeit zu verringern.

Implikationen für diese Arbeit

Mit partiell und vollständig integrierten Lösungsansätzen zur Umlauf- und Dienstplanung können kosteneffizientere Ressourceneinsatzpläne als mit einer sequenziellen Planung bestimmt werden. Darüber hinaus ist bei der integrierten Planung im größeren Maß eine Abstimmung zwischen Umlauf- und Dienstplan möglich. Da diese Abstimmung bzw. Vernetzung und Verknüpfung zwischen den einzelnen Ressourcenebenen Einfluss auf die Propagation von Verspätungen des gesamten Ressourceneinsatzes hat, gilt es im Rahmen dieser Arbeit zu untersuchen, wie die Vermeidung

von Verspätungspropagation während der Planausführung geeignet in integrierten Lösungsansätzen abgebildet werden kann. Als Basis für einen partiell integrierten Lösungsansatz sollte dabei auf dem Ansatz von [Gintner, 2008] aufgebaut werden, da bei diesem im Vergleich zu einer sequenziellen Planung die Kostenoptimalität der Umlaufplanung weiterhin sichergestellt ist, aber zusätzliche Freiheitsgrade für die Dienstplanung existieren. Als Basis für einen vollständig integrierten Lösungsansatz sollte auf den von [Steinzen, 2007] vorgeschlagenen Ansatz zurück gegriffen werden, da dieser auf einer ähnlichen leistungsfähigen Formulierung des Umlaufplanungsproblems als TSN-Modell basiert wie der partiell integrierte Ansatz von [Gintner, 2008]. Bei [Kliewer et al., 2012] wird zwar deutlich, dass sich durch die Möglichkeit, Fahrten bei der integrierten Planung um wenige Minuten verschieben zu können, die Komplexität des zu lösenden Planungsproblems auch bei einer TSN-Formulierung stark erhöht. Allerdings bieten die zusätzlichen Freiheitsgrade großes Potenzial für noch kosteneffizientere Ressourceneinsatzpläne. Daher sollte bei der Entwicklung von Lösungsansätzen zur Berücksichtigung von Robusheitsaspekten auch die integrierte Planung mit Zeitfenstern genauer untersucht werden. Dabei sollten insbesondere Techniken entwickelt werden, mit denen die Problemkomplexität beim Lösen reduziert werden kann.

3.3 Robuste Umlauf- und Dienstplanung

Neben der Integration verschiedener Planungsstufen ist der Aspekt der Robustheit von Plänen zu einem wichtigen Thema in der Forschung zu Verkehrssystemen geworden. Ansätze, welche Robustheit in verschiedenster Ausprägung berücksichtigen, sind allerdings hauptsächlich im Flugverkehr bzw. Bahn(fern)verkehr zu finden (siehe Abschnitt 3.4). Im Bereich des Busverkehrs im ÖPNV existieren nur sehr wenige Ansätze, die auf eine robuste Umlaufplanung und/oder robuste Dienstplanung abzielen. Generell lassen sich diese Ansätze in reaktive online Ansätze und proaktive offline Ansätze zur Ressourceneinsatzplanung unterscheiden. Darüber hinaus existieren Ansätze zur Neuplanung im Fall schwerer Störungen bzw. bei Fahrzeugausfällen während der Ausführung. Für einen Überblick über Lösungsansätze zur Neuplanung des Fahrzeugeinsatzes bei umfangreichen Betriebsstörungen sei an dieser Stelle auf [Visentini et al., 2014] verwiesen. Im Folgenden werden Ansätze zur robusten Umlaufplanung und robusten Dienstplanung mit ihren speziellen Charakteristika diskutiert, wobei Abschnitt 3.3.1 Ansätze zur Umlaufplanung und Abschnitt 3.3.2 Ansätze zur Umlauf- und Dienstplanung umfasst. Bei der Analyse der einzelnen Ansätze werden jeweils Potenziale und Grenzen dargelegt. Die sich aus der Betrach-

tung aller Ansätze für diese Arbeit ergebenen Implikationen werden zur besseren Übersicht gemeinsam in Abschnitt 3.5 zusammengefasst und bei der Formulierung des Forschungsbedarfs in Kapitel 4 aufgegriffen.

3.3.1 Lösungsansätze zur Umlaufplanung

[Huisman et al., 2004] stellen ein Online-Verfahren vor, in welchem ein Umlaufplan nicht im Vorfeld für den gesamten Tag konstruiert, sondern mehrere Umlaufpläne dynamisch während der Ausführung erstellt werden. Um mit auftretenden Störungen während eines Tages besser umgehen zu können, wird der Tag in verschiedene Perioden unterteilt. Zu einem bestimmten Zeitpunkt wird ein Umlaufplan für den nächsten Zeitabschnitt ab diesem Zeitpunkt erstellt, basierend auf den für diese nächste Periode verfügbaren Informationen. Die entscheidende ***Annahme*** der Autoren ist hierbei, dass die Zukunft jederzeit vollständig bekannt ist. Sie wird mit Hilfe von (historischen) Verspätungsszenarien abgebildet, die jeweils konkrete Tage aus der Vergangenheit repräsentieren. Die Qualität der erstellen Pläne wird anhand dreier ***Kennzahlen*** beurteilt: der Anzahl der benötigten Fahrzeuge, der Anteil an Fahrten, die verspätet starten, sowie anhand eines fiktives Maßes für sich möglicherweise ergebende Verspätungskosten. Getestet wird der Ansatz auf einer ***Testinstanz***, die den Regionalverkehr in den Niederlanden abbildet und 1.104 Servicefahrten sowie vier Depots umfasst, wobei nur eine sehr kleine Anzahl von Fahrten aus allen Depots bedient werden dürfen. Da trotzdem aufgrund der Problemgröße ein exaktes Lösen des Mehrdepot-Optimierungsmodells für mehrere Szenarien – es werden für die Instanz 10 Szenarien genutzt – hintereinander schwierig ist, wird im Lösungsverfahren eine Cluster-Reschedule Heuristik eingesetzt, mit der Fahrten auf Depots aufgeteilt werden um Eindepot Probleme lösen zu können. Als ***Vergleich*** dienen Ergebnisse diverser Umlaufplanungen ohne und mit festen Pufferzeiten im Bereich von 2 bis 10 Minuten zwischen je zwei Fahrten, sowie eine zusätzlich durchgeführte Sensitivitätsanalyse der errechneten Umlaufpläne, bei der Fahrtzeiten zu Grunde gelegt werden, die aus einer Normalverteilung gezogen werden. Die Autoren kommen zu dem ***Schluss***, dass durch den dynamischen Ansatz weniger Fahrten verspätet starten und Verspätungen reduziert werden können, indem nur ein paar Fahrzeuge mehr eingeplant werden. Ein Nachteil des vorgeschlagenen Verfahrens liegt allerdings darin, dass Pläne erst kurz vor Ausführung bekannt werden und somit eine laufend hohe Anpassungsfähigkeit seitens der Planer und auch Fahrer verlangt wird. Darüber hinaus müssen noch die Abhängigkeiten mit den anderen Planungsstufen wie z.B. der Dienstplanung im Verfahren berücksichtigt werden.

[Kramkowski et al., 2009] schlagen einen Offline-Ansatz für verspätungstolerante Umlaufplanung vor. Um den bestehenden Planungsprozess beizubehalten und die Planung zugleich um Robustheitsüberlegungen zu ergänzen, sollen potenzielle Störungen durch geschicktes Verteilen von Pufferzeiten bereits während der Planungsphase berücksichtigt werden. Dafür wird die (monokriterielle) Metaheuristik Simulated Annealing for Noise Environments von [Branke et al., 2008] mit zwei einfachen Nachbarschaftsoperatoren als Verbesserungsheuristik für bestehende Umlaufpläne eingesetzt. Mit der ***Annahme***, dass Verspätungen bei Leerfahrten sowohl aus Passagier- als auch aus Unternehmenssicht von geringerem Interesse sind, bezieht sich die Messung der Verspätungstoleranz der erstellten Umlaufpläne auf sekundäre bzw. propagierte Verspätungen von Servicefahrten. Dabei werden für eine unverzerrte Messung keine dispositiven Maßnahmen während der Ausführung abgebildet. Die Lösungen werden anhand folgender ***Kennzahlen*** beurteilt: geplante Fahrzeugkosten und erwartete, fiktive Verspätungskosten in Abhängigkeit der Länge von primären und propagierten Verspätungen plus einen Bestrafungsfaktor für propagierte Verspätungen. Die propagierten Verspätungen ergeben sich dabei als über mehrere Szenarien gebildeter Durchschnitt der erwarteten Länge einer sekundären Verspätung pro Servicefahrt ($E(SD)$). Beim Vergleich dieser Kennzahl werden Unterschiede aufgrund der Frequenz der Verspätungen (entpricht der Wahrscheinlichkeit von sekundären Verspätungen größer als Null $P(SD > 0)$) oder Länge der Verspätungen (entspricht der erwarteten Länge von sekundären Verspätungen größer als Null $E(SD|SD > 0)$) berücksichtigt. Der Ansatz wird auf ***Testinstanzen*** mit 424 und 426 Fahrten, sowie einer größeren Instanz mit mehreren Fahrzeugtypen und 1296 Fahrten getestet. Als Initiallösung dient jeweils eine mit dem Mehrgüter-Flussmodell von [Kliewer et al., 2006] berechnete, kostenoptimale Lösung, die nach dem Prinzip First-In-First-Out (FIFO) dekomponiert wurde. Die jeweilige Initiallösung dient auch als ***Vergleich*** zur Bewertung der neu errechneten Pläne. Da bereits während des Verfahrens Monte-Carlo Simulation eingesetzt wird, wird in den zur Beurteilung dienenden Simulationsläufen noch einmal eine andere Verteilung zur Generierung primärer Verspätungen benutzt. Diese Verteilung entspricht einer Approximation der in [Huisman et al., 2004] beschriebenen Verspätungsszenarien. Als ***Fazit*** ergibt sich, dass eine Erhöhung der geplanten Kosten zu leicht robusteren Umläufen führt. Dabei sind keine Zusatzfahrzeuge nötig, da die gewählten Fahrzeugfixkosten die variablen Kosten deutlich dominieren. Betrachtet man allerdings die sich aus geplanten Kosten plus fiktiven Verspätungskosten ergebenden Gesamtkosten für die einzelnen Lösungen, dann ist die ursprüngliche, etwas unrobustere Initiallösung pro Instanz immer die beste Wahl. Die Kennzahlen $E(SD)$, bzw. $E(SD|SD > 0)$ und $P(SD > 0)$ bilden die Verspätungseffekte deutli-

cher ab, als es durch die reine Betrachtung der fiktiven Kostenfunktion während des Lösungsverfahrens ermöglicht wird. Darüber hinaus sollte untersucht werden, ob ein robuster Umlaufplan zwangsläufig zu einem insgesamt robusteren Ressourceneinsatz führt.

[Naumann et al., 2011] stellen einen Ansatz für verspätungstolerante Umlaufplanung vor, der auf Stochastischer Programming basiert. Die Modellierung basiert auf der deterministischen TSN-Formulierung für Umlaufplanung nach [Kliewer et al., 2006]. Sie wird zusätzlich um alle möglichen Verbindungskanten zwischen einzelnen Knoten ergänzt, um die direkte Auswirkung einer primären Verspätung zwischen zwei aufeinanderfolgenden Fahrten abbilden zu können. Eine Propagation von Verspätungen über mehrere Fahrten wird dabei nicht betrachtet. Ziel ist es, einen Umlaufplan zu erstellen, der die geringsten erwarteten Kosten über alle bei der Lösung betrachteten Verspätungsszenarien hat. Die Umlaufpläne werden deshalb anhand der folgenden ***Kennzahlen*** beurteilt: geplante Kosten und Bestrafungskosten. Die Höhe der Bestrafungskosten ist abhängig von den sich aufgrund der Szenarien ergebenden primären Verspätungen und ihrer Propagation auf eine nachfolgende Fahrt und - ähnlich wie in den zuvor vorgestellten Ansätzen - der Wahl eines Bestrafungsfaktors für solche Verspätungen. Für die vier ***Testinstanzen*** mit 313 bis 662 Fahrten werden jeweils 100 Verspätungsszenarien in der Optimierung genutzt, die Verspätungen beinhalten, welche aus einer Exponentialverteilung gezogen werden. Als ***Vergleich*** für die erstellten Umlaufpläne werden kostenoptimale Umlaufpläne und solchen Umlaufpläne herangezogen, die feste Pufferzeiten zwischen den Fahrten beinhalten. Zusätzlich wird eine Sensitivitätsanalyse mit jeweils 300 Szenarien durchgeführt. Die Szenarien werden mit der gleichen wie während der Optimierung genutzten Verteilungsfunktion für Verspätungen erstellt. Mit der Simulation wird eine genaue Abbildung der Verspätungspropagation möglich und der Effekt von Verspätungen im Gegensatz zur Optimierung nicht mehr unterschätzt. ***Erkenntnisse*** Der vorgeschlagene, stochastische Ansatz ist in Bezug auf geplante Kosten und Verspätungsreduktion deutlich besser, als der simple Ansatz, welcher ohne historische Verspätungsinformationen arbeitet und fixe Pufferzeiten berücksichtigt. Da das Verfahren selbst keine korrekte Verspätungspropagation abbildet, ist es notwendig durch Variieren des Bestrafungsfaktors eine Reihe verschiedener Umlaufpläne zu erstellen, die dann durch Simulation evaluiert werden und so eine in Bezug auf potenziell auftretende Verspätungspropagation geeignete Lösung ausgewählt werden kann. Es werden keine Informationen über Laufzeiten des Verfahrens bzw. enstehende Netzwerkgrößen gegeben. Durch die direkte Verbindung aller Fahrten im Netzwerkmodell erscheint solch ein stochasti-

scher Ansatz allerdings für größere Netzwerke oder für das Einbeziehen zusätzlicher Ressourcen ungeeignet.

3.3.2 Lösungsansätze zu kombinierter Umlauf- und Dienstplanung

In [Huisman und Wagelmans, 2006] wird ein dynamischer Ansatz diskutiert, welcher sowohl Umlauf- als auch Dienstplanung umfasst. Der Ablauf des Verfahrens und die grundlegenden Annahmen entsprechen dem von [Huisman et al., 2004] vorgestellten Verfahren zur dynamischen Umlaufplanung. Zusätzliche ***Annahmen*** bestehen darin, dass Fahrten nur aufgrund von Umlaufverspätungen verspätet starten können. Verspätungen werden also nur von Fahrzeugen induziert und Fahrzeugwechsel von Fahrern erzeugen keine Verspätungen. Da Passagiere eine höhere Priorität als Fahrer haben, werden eher Dienstregeln verletzt als Fahrten verspätet. Beurteilt wird eine Lösung durch sechs ***Kennzahlen***: die Gesamtanzahl benötigter Fahrzeuge und Fahrer, der prozentuale Anteil der zu spät startenden Fahrten, künstliche Verspätungskosten, der prozentuale Anteil der Dienste bei denen mindestens eine Dienstregel verletzt wird, und künstliche Kosten für diese Dienstregelverletzungen. Experimente werden auf zwei sehr kleinen ***Testinstanzen*** mit 164 bzw. 304 Fahrten aus dem gleichen Verkehrsgebiet wie in [Huisman et al., 2004] durchgeführt. Im durch die größere Instanz gebildeten Mehrdepotfall wird eine Cluster-Reschedule Heuristik benutzt, so dass Mehrdepot-Probleme als mehrere Eindepot-Probleme gelöst werden können. ***Verglichen*** werden die Ergebnisse mit einer statischen offline Umlauf- und Dienstplanung ohne und mit Pufferzeiten von 2 bis 10 Minuten zwischen je zwei Fahrten umgesetzt mit dem integrierten Modell von [Huisman et al., 2005], sowie mit Hilfe einer Sensitivitätsanalyse, bei der im Gegensatz zu den benutzten zehn Szenarien, leicht abgewandelte Verspätungen eingesetzt werden, die aus einer Normalverteilung gezogen werden. Als ***Erkenntnis*** ergibt sich, dass der dynamische Ansatz für Umlauf- und Dienstplanung bei weitem nicht so gut funktioniert wie für die reine Umlaufplanung. Selbst für die kleinere Instanz sind die Laufzeiten für eine Praxisanwendung zu groß, und beim etwas größeren Mehrdepot-Fall ist der statische Ansatz mit Pufferzeiten in allen Belangen besser. Darüber hinaus sind die bei diesem Extrembeispiel für ein Verfahren mit ständigen Anpassungen in der Ausführungsphase getroffenen Annahmen in praktischer Umgebung schwierig zu halten, da z.B. auch Fahrer Verspätungen erzeugen bzw. propagieren. Auch erfordert das vorgestellte Vorgehen tiefgreifende organisatorische Veränderungen im Planungs- bzw. Ausführungprozess. Neben der Bereitschaft seitens der Busfahrer Dienstaufgaben immer nur kurzfristig zu kennen

und zu erfüllen, müssen Arbeitszeitregelungen und Regeln der tagesübergreifenden Dienstreihenfolgeplanung in dynamischem Umfeld erfasst und erfüllt werden können, wozu wiederum leistungsfähige online arbeitende Verfahren nötig sind.

3.4 Robuste Planung in weiteren Verkehrsdomänen

Die Forderung nach Robustheit bei der Ausführung von Plänen ist in vielen Transportbereichen zu finden. Im Folgenden werden vergleichend Ansätze aus den Domänen *Flugverkehr* in Abschnitt 3.4.1 und *Bahnverkehr* in Abschnitt 3.4.2 vorgestellt. In diesen Bereichen beschäftigt sich die Forschung bereits seit einigen Jahren intensiv mit robuster Planung. Bei der Zusammenstellung werden insbesondere der Einfluss der gewählten Planungsmethoden auf die Robustheit der einzelnen Pläne, der Umgang der Ansätze mit Verspätungseffekten und die gewählte Evaluation der vorgeschlagenen Methoden herausgearbeitet. Die Veröffentlichungen zum Flugverkehr konzentrieren sich im Allgemeinen eher auf die Planung des Ressourceneinsatzes, wohingegen die Veröffentlichungen zum Bahnverkehr eher Fahrplanungs- und Recovery-Ansätze zum Thema haben. Die sich durch die Diskussion der einzelnen Forschungsbeiträge für diese Arbeit ergebenen Implikationen werden in Abschnitt 3.5 mit den Betrachtungen zu robuster Umlauf- und Dienstplanung zusammengeführt und zur Ableitung des Forschungsbedarfs in Kapitel 4 verwendet.

3.4.1 Lösungsansätze im Flugverkehr

Die Planungsaufgaben für den Ressourceneinsatz im Flugverkehr werden ähnlich wie im ÖPNV in einem mehrstufigen Prozess bearbeitet. Nachdem die zu bedienenden Flüge im *Flight Scheduling* festgelegt werden, wird der Flugzeugeinsatz im Gegensatz zur Umlaufplanung im ÖPNV in zwei Stufen geplant. Im *Fleet Assignment* wird zunächst bestimmt, welche Flüge durch einen bestimmten Flugzeugtyp ausgeführt werden sollen. Im anschließenden *Aircraft Routing* werden jeweils für die Flüge eines Flugzeugtyps Routen – als Äquivalent zu Umläufen im ÖPNV – aus Sequenzen von Flügen gebildet, die von einem Flugzeug bedient werden sollen. Dabei spielen Leerflüge bzw. Leerflugverbindungen kaum eine Rolle. Für die nach Flugzeugtypen aufgeteilten Flüge werden darüber hinaus im *Crew Pairing* Sequenzen von anonymen Diensten gebildet. Wie im ÖPNV können Dienste dabei durchaus Leerfahrt- bzw. Transferverbindungen beinhalten, in diesem Fall z.B. Flüge als Passagier. Bei der Ausführung werden grundsätzlich Flüge mit einer Toleranz von 15 Minuten als

pünktlich angesehen. Beim ÖPNV bewegt sich dieser Wert häufig bei einer bis drei oder vier Minuten.

Zu den Problemstellungen im Flugverkehr existieren im Gegensatz zu ähnlichen im ÖPNV eine Vielzahl von Veröffentlichungen, die bereits Robustheitsaspekte in der Planung berücksichtigen. Nachfolgend werden schwerpunktmäßig diejenigen diskutiert, die mögliche *Propagationseffekte von Verspätungen* untersuchen bzw. solche Effekte bereits in der Planung verhindern wollen. Dabei werden insbesondere die Ansätze aufgegriffen, die eine *Stabilität* von Plänen verfolgen. Zum anderen werden wesentliche Ansätze zur *Flexibilität* betrachtet, die mögliche Tauschoperationen zur Reduktion von Verspätungspropagation während der Ausführung in die Planung einbeziehen. Für weiterführende Informationen über Ansätze, die beim Auftreten von (größeren) Störungen in der Ausführung zur Umplanung der bestehenden Pläne eingesetzt werden können (z.B. bei Flughafenschließungen über mehrere Stunden), sei auf [Petersen et al., 2012] und [Maher, 2016] verwiesen. [Clausen et al., 2010] geben einen ausführlichen Überblick über Forschung zu Störungsmanagement im Flugverkehr. Über die Modellierungen und Lösungsmethoden hinaus sind häufig Analysen von Verspätungsdaten Inhalt der Veröffentlichungen. Im Gegensatz zum ÖPNV stehen für den Luftverkehr mit den von Eurocontrol[1] bzw. dem Bureau of Transportation Statistics des US Transportministerium[2] erhobenen Daten auch umfangreiche, öffentlich zugängliche Analysequellen zur Verfügung.

Fokus: Flugplananpassung und Umlaufplanung

[Ahmadbeygi et al., 2010] konzentrieren sich auf das Retiming des Flugplanes. Unter Berücksichtigung von bereits getroffenen Entscheidungen zur Flugzeugflotte und zu Crews wird durch Verschieben der Abflugzeiten für die Ausführung vorhandene Puffer bzw. Standzeiten so verteilt, dass eine Propagation von Verspätungen verringert werden kann ohne die operativen Plankosten von Flugzeug- und Dienstplan zu verändern. Dazu stellen die Autoren ein Optimierungsmodell auf, bei dem ausschließlich die erwartete Verspätungspropagation minimiert werden soll. Um die Propagation von Verspätungen über Crews und Flugzeug abzubilden, werden vor dem Retiming-Prozess mögliche Propagations-Bäume erstellt, die aus Echtdaten abgeleitete Verspätungsszenarios für einzelne Flüge beinhalten und in denen Verspätungen solange propagiert werden bis sie absorbiert sind. Für den Aufbau der Propagations-Bäume werden jeweils die in Bezug auf eine mögliche Verspätungspropagation schlimmstmöglichen Flugverschiebungen zwischen zwei aufeinanderfolgenden Flügen berücksichtigt. In be-

[1] siehe http://www.eurocontrol.int/, letzter Abruf 30.3.2017
[2] siehe http://www.rita.dot.gov/bts/, letzter Abruf 30.3.2017

stimmten Fällen werden die propagierten Verspätungen dabei unter- bzw. überschätzt. Der Ansatz wird mit Datensätzen zweier Tage einer großen U.S. Flugline aus dem Jahr 2007 mit über 500 Flügen getestet. Dabei wird eine Propagation über Flugzeuge und Cockpit Crews abgebildet sowie eine maximale Verschiebung um 15 Minuten zugelassen. Um primäre Verspätungen zu generieren, werden empirsche Verteilungen aus Verspätungsdaten aus einem Zeitraum von 12 Monaten abgeleitet und nach Herkunftsstation und Länge aufgeteilt. Die entsprechenden Wahrscheinlichkeitsverteilungen werden sowohl im Lösungsverfahren als auch in einer diskret-ereignisbasierten Simlution benutzt, mit der die erstellten Pläne zusätzlich validiert werden. Es ***zeigt*** sich, das nachträgliche Anpassungen des Flugplanes ohne Veränderung der Ressourceneinsatzpläne Potenzial besitzt, die Ausführung robuster zu gestalten. Allerdings ist dabei zu berücksichtigen, dass möglicherweise Dienstregelverletzungen auftreten und bereinigt werden müssen.

[Lan et al., 2006] betrachten das Routing von Flugzeugen mit dem alleinigen Ziel die gesamten erwarteten propagierten Verspätungen auf den einzelnen Flugzeug-Routen zu minimieren. Dabei werden die einzelnen Routen isoliert voneinander betrachtet und es findet keine Verknüpfung über zusätzliche Ressourcen wie Crews oder Passagiere statt. Die Autoren formulieren ein deterministisches Optimierungsmodell mit stochastischem Input. Die Verspätungsverteilung der Ankunftsverspätungen von Flügen wird aus der öffentlich verfügbaren *Airline Service Quality Performance (ASQP) Database* hergeleitet und folgt einer logarithmischen Normalverteilung. Der Ansatz wird auf vier kleinen Flugnetzwerken (20 bis 102 Flüge) getestet und mit dem existierenden Plan einer US-Airline (aus August 2000) bezüglich propagierter Verspätungen und On-Time-Performance verglichen. In einem zweiten beschriebenen Ansatz verringern die Autoren durch Retiming des Flugplanes die Anzahl der möglicherweise verpassten Passagierverbindungen. Es ***zeigt*** sich, dass durch Berücksichtigen von Verspätungspropagation die Umlaufpläne robuster gestaltet werden können. Allerdings sollten hierbei auch andere Ebenen wie z.B. die Dienstplanebene einbezogen werden, und es bleibt fraglich, inwieweit dafür zusätzliche Kosten verursacht werden bzw. verursacht werden dürfen. Darüber hinaus wird wiederrum das Potenzial nachträglicher Flugplanänderungen deutlich, die es ermöglichen bessere Verbindungen der Flüge und somit eine verringerte Verspätungspropagation für Passagiere zu erreichen. Allerdings finden sich keine Hinweise, wie groß der Einfluss dieser Flugänderungen auf die benötigten Kosten für den Ressourceneinsatz ist.

[Burke et al., 2010] beschäftigen sich mit dem gleichzeitigen Retiming von Flügen und Rerouting von Flugzeugen. Ausgehend von einem bereits erstellten Umlaufplan für Flugzeuge, fokussieren sie auf eine Erhöhung der Robustheit ohne weitere Kos-

tenbetrachtung. Dabei werden zwei Zielgrößen verfolgt mit denen Robustheit abgebildet wird. Zum einen wird die Wahrscheinlichkeit für eine Verspätungspropagation zwischen Flügen betrachtet (*Stabilität*). Diese Punktlichkeitswahrscheinlichkeit soll in der Summe über alle Flüge erhöht werden. Zum anderen wird die Anzahl von Flugzeug-Tauschmöglichkeiten (*Flexibilität*) betrachtet. Der heuristische Ansatz basiert auf einem evolutionärem Algorithmus und arbeitet inkrementell mit schrittweisen Verbesserungen der Pläne. Als Testdaten werden Flugpläne der KLM Royal Dutch Airline aus dem Jahr 2006 verwendet, die jeweils ungefähr fünf bis 15 Flugzeugumläufe umfassen. Die Ergebnisse werden mittels Simulation validiert. Dazu wird das KLM eigene Simulationsmodell genutzt, mit welchem auch Recovery-Strategien berücksichtigt werden. Diese betreffen allerdings nur den Flugzeugeinsatz. Die Autoren kommen zu dem ***Schluss***, dass beide Zielgrößen verbessert werden können, bezogen auf den Einsatz von Flugzeugen der Einfluss von *Stabilität* bei der Planung gegenüber *Flexibilität* allerdings dominant ist, so dass primär auf Stabilität geachtet werden sollte - ergänzt um Flexibilitätsüberlegungen. Offen bleiben die Fragen, zu welchen Kosten die errechneten Pläne robuster werden und welche Veränderungen der Dienstpläne dafür nötig sein würden.

Fokus: Dienstplanung

[Shebalov und Klabjan, 2006] betrachten im Rahmen der Dienstplanung für Crews gleichzeitig zwei Zielsetzungen: die Minimierung der Crew Kosten und die Maximierung der Anzahl von sogenannten Move-up Crews zur Erhöhung der Robustheit (Flexibilität) in der Ausführung. Move-Up Crews stellen Besatzungen dar, die durch einen Crewtausch Aufgaben von anderen Besatzungen im Fall von Verspätungen übernehmen können. Durch die Maximierung der Anzahl solcher Crews soll die Anzahl von Tauschmöglichkeiten erhöht werden. Um den Trade-Off zwischen Plankosten und Move-up Crews zu steuern, können in der Modellformulierung die Kosten einer rein kostenminimalen Dienstplanlösung als Restriktion formuliert einen Kostenrahmen für die Planung von Move-up Crews vorgegeben. Um die Move-up Crews darüber hinaus geschickt zu verteilen, können Vorgaben für die benötigten Häufigkeiten einer solchen Crew pro Flug gemacht, sowie Gewichtungen für die Bedeutung der einzelnen Flüge getroffen werden. Der Lösungsansatz wird auf drei Instanzen mit 80 bis 228 Flügen getestet und durch die Simulation von Recovery-Maßnahmen während der Ausführung evaluiert. Dazu werden Verspätungsszenarien generiert, die jeweils (mehrere) 80-minütige Flughafenschließungen repräsentieren. Die Arbeit ***zeigt***, dass die operativen Kosten durch bereits vorgeplante Reaktionsmöglichkeiten deutlich gesenkt werden können. Allerdings weisen die Autoren selbst darauf hin, dass Vorsicht beim

Erkaufen von Robustheit durch zusätzliche Crew Kosten angebracht ist. So nutzen die eingeplanten Reaktionsmöglichkeiten nur, wenn Verspätungen im betrachteten Umfang auch tatsächlich eintreffen. Eine zu starke Erhöhung der Dienstplankosten zur Erhöhung der Flexibilität kann somit schnell zu Lösungen führen, die keinen betrieblichen Vorteil besitzen.

Der Ansatz wird von [Ionescu und Kliewer, 2011] aufgegriffen. Im Unterschied zu [Shebalov und Klabjan, 2006] werden zur Reduktion der Problemgröße in einem Preprocessing-Schritt Tauschmöglichkeiten basierend auf primären Verspätungen und möglicher Verspätungspropagation zwischen jeweils zwei Flügen vorgeneriert. Die auszuführenden Flugzeugumläufe sind dabei vorgegeben. Durch die identifizierten riskanten Verbindungen bei potenziellen Flugzeugwechseln und die Beschränkung auf eine geringe Anzahl pro Wechsel sollen in der Ausführung tatsächlich nutzbare Tauschmöglichkeiten passender eingeplant werden. Die Zielfunktion bildet Crewkosten und einen Bonus für jede eingeplante Tauschmöglichkeit ab. Der Lösungsansatz wird auf einer nicht näher bestimmten Menge von Flugplänen mit 396 bis 427 Flügen getestet. Die erstellten Pläne werden durch eine anschließende Simulation mit einfachen regelbasierten Recovery-Maßnahmen evaluiert. Die Arbeit ***zeigt***, dass sich durch eine Erhöhung der Flexibilität eine Steigerung der Pünktlichkeit (gemessen in einem Toleranzbereich von 15 Minuten) bis zu einem gewissen Grad (hier um 1-2%) erreichen lässt. Die aggregierten Ergebnisse bieten jedoch keine genauen Angaben über die Anzahl der jeweils verwendeten Dienste. Die Experimente zeigen darüber hinaus auch, dass eine starke Erhöhung der Flexibilität – einhergehend mit einer erhöhten Dienstanzahl – kaum zusätzlichen Einfluss auf die tatsächliche Verspätungspropagation nehmen kann. Um Pläne bei der Erstellung realistisch bewerten zu können wird von den Autoren vorgeschlagen, die potenzielle Propagation von Verspätungen in die Zielfunktion zur Lösungsbewertung einzubeziehen. Da das Verfahren Tauschvorgänge während der Ausführung bewusst forciert, sollte überdies eine Kombination mit Aspekten zur Erhöhung der Stabilität untersucht werden. So könnten bei der Dienstbildung auch vorgeplante Pufferzeiten für riskante Verbindungen statt nötige Tauschvorgänge zur Vermeidung von Verspätungspropagation beitragen.

[Ehrgott und Ryan, 2002] bilden die Crew Planung als bi-kriterielles Optimierungsproblem ab. Neben Kostenminimierung verfolgt der vorgeschlagene Ansatz auch die Minimierung eines Nicht-Robustheitsmaßes und somit eine Erhöhung der Robustheit des Dienstplanes. Dienste werden danach bewertet, ob zwischen Flugzeugwechseln ausreichend Puffer zur Verfügung steht, um erwartete primäre Verspätungen des vorherigen Fluges aufzufangen. Um beide Zielsetzungen berücksichtigen zu können, wird ein Modell basierend auf dem Elastic-Cost Constraint-Verfahrens aufgestellt (vgl. Ab-

schnitt 3.1.2) und es werden zwei Schritte durchgeführt. Im ersten Schritt wird ein kostenoptimaler Dienstplan berechnet. Im zweiten Schritt wird die Robustheit erhöht vor dem Hintergrund, dass eine Erhöhung der Kosten um $x\%$ möglich ist. Die Arbeit ***zeigt***, dass die Zielsetzungen Kostenminimierung und Robustheitsmaximierung in einem Modell so abbildbar sind, dass sie gegeneinander abgewogen werden können. Das Vorgehen erlaubt es, gemäß des abgebildeten Robustheitsmaßes, wesentlich robustere Dienstpläne zu geringen Mehrkosten zu erstellen. Die Robustheit wird dafür maßgeblich durch zusätzliche nächtliche Ruhepausen erreicht, die die Anzahl der Flugzeugwechsel einer Crew verringern. Die Lösungen werden allerdings nicht im Hinblick auf eine potenzielle Verspätungspropagation zwischen Flugzeugen und Crews validiert.

[Yen und Birge, 2006] formulieren die Dienstplanung als ein stochastisches Optimierungsproblem mit zwei Stufen. Ziel ist es, die Kosten der Crew zu minimieren und die Robustheit der Dienstpläne zu erhöhen. Dafür werden zusätzlich die Kosten von erwarteten Verspätungen aufgrund von Flugzeugwechseln der Crew minimiert. In der ersten Stufe wird ein deterministisches Crew Pairing Problem gelöst. In der zweiten Stufe werden die durch Flugzeugwechsel induzierten Verspätungen betrachtet. Der vorgeschlagene Algorithmus arbeitet auf Flugpaaren, wozu vorher eine Enumeration aller möglichen Pairings nötig ist. Die durchgeführten Tests basieren auf Daten des Air New Zealand Domestic Netzwerk aus dem Jahr 1997 und umfassen drei Testinstanzen mit 61 bis 79 Flügen (bzw. 9-11 Flugzeugen). Es werden 100 gleich wahrscheinliche Szenarien für Verspätungen gebildet. Die Länge der Verspätungen folgen dabei einer abgeschnittenen Gammaverteilung oder logarithmischer Normalverteilung und entsprechen damit den Erkenntnissen der Analyse von Air New Zealand Daten. Die Autoren kommen zu dem ***Schluss***, dass eine Reduktion der Flugzeugwechsel von Crews und ein Erhöhen der Anschlusszeiten zu in der Ausführung robusteren Plänen führen kann. Darüber hinaus zeigt sich, dass ein stochastisches Optimierungsproblem zum Berechnen eines robusten Ressourceneinsatzplanes verwendet werden kann. Allerdings nur für eine kleine, auf eine Ressourcenebene beschränkte Problemstellung unter der Voraussetzung, dass alle Lösungsmöglichkeiten vorher aufgezählt werden.

[Tam et al., 2011] vergleichen den bi-kriteriellen Ansatz von [Ehrgott und Ryan, 2002] mit einer von den Autoren modifizierten Version des stochastischen Ansatz von [Yen und Birge, 2006] im Rahmen eines kommerziellen Dienstplanungs-Systems. Als Vergleichsproblem dient ein Flugplan über sieben Tage mit insgesamt 442 Flügen von Air New Zealand und entsprechende Verspätungsdaten. Die Lösungen werden hinsichtlich mehrerer Robustheitsindikatoren beurteilt. Die Autoren kommen zu dem ***Schluss***, dass beide Ansätze robustere Crew Pläne mit geringfügig höheren Kosten als bei kostenoptimaler Planung erstellen können. Ohne detaillierte Laufzeiten an-

zugeben wird die Laufzeit des bi-kriteriellen Ansatzes als deutlich besser beurteilt. Dieser Ansatz mit Nutzen eines (Un-)Robustheitsmaß kann mögliche Auswirkungen von Störungen geeignet mitberücksichtigen und liefert gute Lösungen für moderate Verspätungskosten. Der stochastische Ansatz ist von der Lösungsqualität bevorzugbar, wenn die Verspätungskosten in ihrer Höhe als substanziell angesehen werden. Allerdings zeigen sich bereits für den Einsatz auf nur einer Ressourcenebene technische Schwierigkeiten bei der Umsetzung des stochastischem Ansatzes.

Fokus: kombinierte Umlauf- und Dienstplanung

[Weide et al., 2010] betrachten Aircraft Routing und Crew Scheduling gemeinsam mit der Zielsetzung die Gesamtzahl der Flugzeugwechsel im Dienstplan zu reduzieren und die Crew-Kosten zu minimieren. Die Kosten für Aircraft Routings werden dabei als fix angenommen. Das integriert formulierte Problem wird mit einem iterativen heuristischen Ansatz gelöst, bei dem abwechselnd zunächst ein Aircraft-Routing-Problem gelöst wird und anschließend das Crew-Scheduling-Problem basierend auf der Aircraft-Routing-Lösung. Pro Iteration wird dabei ein Bestrafungsfaktor erhöht, um Flugzeugwechsel der Crews in den Lösungen immer stärker zu vermeiden. Der Ansatz wird für Pläne mit ca. 750 Flüge und 14 Flugzeugen getestet. Es ***zeigt*** sich, dass sich Flugzeugwechsel als betrachtetes Unrobustheitsmaß verringern lassen, bei gleichzeitiger Erhöhung der Plankosten für Crews. Dabei bleibt der Einfluss auf die Kosten der Aircraft Routings, sowie auf eine tatsächliche Reduzierung der Verspätungspropagation offen. Es stellt sich die Frage, ob die Flugzeugwechsel an den richtigen Stellen eingespart werden, da mögliche Verspätungen der einzelnen Flüge nicht unterschieden werden und unabhängig von ihrer Lage im Ressourceinsatzplan und der davon ausgehenden Propagationsgefahr geplant wird.

[Dueck et al., 2012] betrachten Aircraft Routing und Crew Scheduling mit dem Ziel die Robustheit (Stabilität) von sowohl Aircraft Schedule als auch Crew Schedule zu erhöhen und gleichzeitig die Kosten für den Diensteinsatz zu minimieren. Die Kosten für Aircraft Routings werden dabei als gegeben angenommen. Die Autoren formulieren ein integriertes, stochastisches Modell mit Rekurs. Dieses wird heuristisch gelöst, indem es durch Dekomposition in zwei separate Teilprobleme zerlegt wird, die nach dem Verfahren von [Weide et al., 2010] iterativ gelöst werden. Im Lösungsverlauf werden die durch Flugzeugwechseln von Crews induzierten zusätzlichen Verspätungen bestraft. Als Vergleich wird ein iteratives, deterministisches Verfahren umgesetzt, dass ähnlich wie bei [Weide et al., 2010] keine Verspätungspropagation betrachtet, sondern zu kurze Pufferzeiten bei Flugzeugwechsel von Crews bestraft. Die Lösungen beider Ansätze werden durch eine Monte-Carlo Simulation mit einem genaueren

Modell zum Abbilden der Propagation von Verspätungen über Flugzeuge und Crews evaluiert. Die betrachteten Problemstellungen umfassen 220 bis 272 Flüge. Zur Generierung von Szenarien über primäre Verspätungen werden Log-Logistik- und Log-Normalverteilungen für verschiedene Verspätungsgründe genutzt. Die Verteilungen wurden aus historischen Daten einer europäischen Fluglinie aus dem Zeitraum 2003 bis 2006 abgeleitet. Eine primäre Verspätung ergibt sich pro Flug aus der Überlagerung der verschiedenen Verspätungsursachen. Während Optimierung und Simulation werden unterschiedliche Szenarien genutzt, die den gleichen Verspätungsverteilungen folgen. Es ***zeigt*** sich, beide Verfahren Robustheit zu vergleichbaren Kosten erreichen können, es für den deterministischen Ansatz allerdings einer intensiven Kalibrierung des betrachteten Robustheitsindikators bedarf. Da in beiden Ansätzen nur die Kosten der Crews berücksichtigt werden, werden allerdings die Ausprägungen der entstehenden Flugzeugumläufe vernachlässigt. Darüber hinaus wird allein die zusätzliche Propagation bei Flugzeugwechseln von Crews verringert. Um die insgesamt erreichbare Robustheit zu erhöhen, schlagen die Autoren vor, diesen Lösungsansatz um eine weitere Phase pro Iteration zu ergänzen. In dieser Phase soll die Verspätungspropagation gezielt innerhalb der einzelnen entstehenden Routen bzw. Pairings minimiert werden.

[Dunbar et al., 2012] präsentieren einen Ansatz, um die erwartete Verspätungspropagation von Aircraft Routings und Crews gemeinsam zu minimieren. Um die Interaktion zwischen Flugzeug und Crew abzubilden, nutzen sie ein iteratives Lösungsschema aufbauend auf dem iterativem Verfahren von [Weide et al., 2010]. Ausgehend von einer Initiallösung mit minimaler Flugzeug- und Crewanzahl, wird abwechselnd ein Optimierungsproblem zur Minimierung von Verspätungen auf Flugzeugebene und ein Problem zur Minimierung von Verspätungen auf Crew-Ebene gelöst. Dabei wird jeweils die Verspätungspropagation einer Ressourcenebene dynamisch berechnet bei fixen, vorberechneten propagierten Verspätungen der anderen Ressourcenebene. Die Autoren benutzen einen Label-Setting-Ansatz, um bei der Umlauf- bzw. Diensterstellung die Weitergabe von Verspätungen zwischen Flugsequenzen exakt zu berechnen. Das alleinige Ziel besteht darin, die gesamte Propagation über Flugzeuge und Crews zu minimieren. Verspätungen werden vor Ausführung des Verfahrens von Wahrscheinlichkeitsverteilungen abgeleitet und als Durchschnittswert für jede Verbindung gebildet. Die Autoren nutzen entweder eine Exponentialverteilung oder eine abgeschnittene Normalverteilung zur Verspätungsgenerierung. Der Ansatz wird für einen sehr kleinen Flugplan eines konkreten Tages mit 54 Flügen und 128 zulässigen Verbindungen zwischen den Flügen getestet, der mit minimal 10 Flugzeugen und 16 Crews bedient werden kann. Es ***zeigt*** sich, dass das gewählte Verfahren ermöglicht, eine exakte Verspätungspropagationen über Ressourcenebenen bei der Planerstellung ab-

zuschätzen und diese gemessen am Zielfunktionswert deutlich zu reduzieren. Da nur die mit Verspätungspropagation assoziierten Kosten minimiert werden, ist es allerdings nicht ersichtlich welche zusätzlichen Plankosten dafür entstehen bzw. in welchem Kostenrahmen die Plankosten erhöht werden können. Darüber hinaus wurde das iterative Vorgehen nur für extrem kleine Problemgrößen mit geringen Freiheitsgraden zur Planung getestet.

[Dunbar et al., 2014] erweitern den zuvor vorgestellten, eigenen Ansatz (vgl. [Dunbar et al., 2012]) um eine Retiming-Komponente. Im Lösungsansatz werden Flüge durch Netzwerkknoten repräsentiert, die durch Kanten miteinander verbunden sind, falls sie hintereinander ausführbar sind. Um das Verschieben von Flugzeiten abzubilden, werden allein auf Ebene des Unterproblems zur Generierung von Aircraft Routings zusätzliche Netzwerkknoten eingeführt. Es werden ergänzend zu den ursprünglichen Flügen jeweils vier weitere Knoten angelegt, die Verschiebungen von ± 5 Minuten bzw. ± 10 Minuten repräsentieren. Beim Einfügen der neuen Knoten werden keine neuen Verknüpfungen zwischen den Flügen geschaffen, sondern es bleiben lediglich die ursprünglichen Flugverknüpfungen bestehen. Darüber hinaus bleibt das gemeinsame Aircraft Routing und Crew Scheduling Hauptproblem unangetastet. Der Lösungsansatz wird ebenfalls für die sehr kleine Instanzen mit 54 Flügen mit wenigen möglichen Verknüpfungen von Flügen zur Bildung von Aircraft Routings bzw. Crew Pairings getestet. Die Autoren kommen zu dem ***Schluss***, dass durch die gemeinsame Betrachtung der drei Planungsschritte die Verspätungspropagation im Vergleich deutlich gesenkt werden kann. Es ist allerdings nicht ersichtlich, inwiefern sich dabei die ursprünglichen Plankosten ändern. Ebenso werden keinerlei Angaben zu dem Ausmaß der geänderten Flugzeiten gemacht. Der vorgeschlagene Lösungsansatz wird überdies in den Tests nur mit zwei bis maximal vier Iterationen ausgeführt. Im Vergleich zu den Laufzeiten im Sekundenbereich des in [Dunbar et al., 2012] beschriebenen Lösungsansatzes betragen die Laufzeiten beim Einbeziehen von Retiming selbst für die sehr kleine Problemstellung zwischen zwei und drei Stunden. Darüber hinaus werden aufgrund der Beschränkung des Retimings auf das Unterproblem zur Umlaufplanung durch das Vefahren Vorgaben aus der Umlaufebene einfach auf die Dienstebene übertragen, ohne diese Entscheidungen in einem gemeinsamen Hauptproblem abzubilden.

Fokus: Datenanalyse

In den zuvor diskutierten Veröffentlichungen wurden teilweise Untersuchungen historischer Verspätungsdaten zur Ableitung von Verspätungszenarien oder Herleitung von Wahrscheinlichkeitsverteilungen für das Auftreten und die Dauer von Störungen

durchgeführt. Daneben existieren bereits vereinzelt umfangreiche, empirische Untersuchungen über die Auswirkungen von Planungspraktiken im Flugverkehr und die Effekte von Verspätungen im gesamten Flugnetzwerk. Im Folgenden werden diese kurz vorgestellt, um mögliche Anhaltspunkte für die Analyse von Plänen im ÖPNV herauszustellen.

[Atkinson et al., 2012] analysieren ausführlich Daten über fünf Jahre aus dem Zeitraum 1997 bis 2009. Die Daten über Inlandsflüge in den U.S.A. basieren auf vier unterschiedlichen Quellen, die hauptsächlich vom us-amerikanischen Bureau of Transportation Statistics (BTS) gesammelt und öffentlich zur Verfügung gestellt werden. Die Verknüpfung von geplanten und tatsächlichen Ankunfts- und Abflugzeiten mit detaillierten Daten über zugehörige Flugzeuge, sowie Daten über Fracht und Passagiere, verfügbaren Plätze, und Informationen über die Airport-Kapazität (speziell den Zugang und die Verfügbarkeit von Boarding-Gates einzelner Fluglinien) ermöglicht es den Autoren detaillierte Informationen über die Performance von 81 Flughäfen und zehn U.S. Airlines zusammenzutragen und insbesondere in Bezug auf den Umgang mit Verspätungen miteinander zu vergleichen. Als ***Erkenntnisse*** stellen die Autoren fest, dass die Anpassungsfähigkeit während der Ausführung – dazu werden Tauschmöglichkeiten für Flugzeuge und Crews sowie Wechselmöglichkeiten der Boarding Gates für Ankünfte und Abflüge gezählt – für alle betrachteten Airlines besonders lohnend ist. Im Vergleich dazu nutzt das Erhöhen der Stabilität des Ressourceneinsatzes durch Pufferzeiten insbesondere Billigfluglinien, da diese generell mit weniger Pufferzeiten planen. Es ist anzumerken, dass die Stabilität allein aufgrund der in den Umlaufplänen eingeplanten Gesamtpufferzeit beurteilt wird. Für dienliche Hinweise zum Umgang mit Pufferzeiten in der Planung sollte im Rahmen dieser Arbeit die durch Einpflegen von Pufferzeiten erreichbare Robustheit und Kosteneffizienz des geplanten Ressourceneinsatzes differenzierter mit Blick auf die jeweils vorgenommene Verteilung der Pufferzeiten zwischen der Ausführung von Fahrten bewertet werden.

[Arikan et al., 2013] führen eine umfassende Analyse empirischer Daten aus den Jahren 2005 bis 2007 der Inlandsflüge amerikanischer Airlines durch, wobei auch Effekte von internationalen Flügen in die Untersuchung mit einfließen. Die Autoren analysieren (frei verfügbare) Beobachtungen zu 21 Millionen Inlandsflügen und bewerten 77 Flughäfen und 18 Airlines bezüglich ihrer Performance. Dazu entwickeln sie zwei stochastische Modelle, die zum einen die Zufälligkeit in der eigentlichen Reisezeit eines Fluges (primäre Verspätungen) und zum anderen eine Verspätungspropagation über Aircraft Rotations im Netzwerk abbilden. Darüber hinaus werden Robustheitsindikatoren für die Bewertung der Flugnetzwerke vorgeschlagen. Die Autoren ***folgern***, dass einzelne Fluglinien ihre verursachten propagierten Verspätung

identifizieren und kontrollieren sollten, um die Leistungsfähigkeit während der Planausführung zu steigern. So ist beim Planen ein Trade-Off zwischen dem Verteilen von Pufferzeiten am Boden und Pufferzeiten innerhalb der für die Flüge eingeplanten Reisezeiten zu berücksichtigen. Pufferzeit sollte überdies insbesondere um Flüge geplant werden, die zu einem Engpass für die weitere Ausführung werden. Das Auftreten solcher „Flaschenhals"-Flüge wird von den Autoren bezogen auf die Lage innerhalb einer Rotation (eher unter den ersten vier Flügen) und die Uhrzeit (verteilt über den ganzen Tag, nicht nur morgens auch mittags) identifiziert. Das generell im Tagesverlauf die Verspätungspropagation zunimmt und somit auch die Pünktlichkeit der Flüge abnimmt wird neben den Plänen auch an der vorhandenen Airport Infrastruktur festgemacht. Die Untersuchung zeigt, dass die für den Ressourceneinsatz erstellten Planstrukturen detailliert auf Flug- bzw. im Rahmen dieser Arbeit auf Servicefahrtenebene zu betrachten sind, um planimmanente Ursachen für Verspätungspropagation genau identifizieren zu können. Es sollte geprüft werden, inwiefern ebenfalls kritische Fahrten innerhalb eingeplanter Aufgabensequenzen erkennbar sind. Zusätzlich sollte die Weitergabe von Verspätungen in Abhängigkeit des Tagesverlaufs beachtet werden.

3.4.2 Lösungsansätze im Bahnverkehr

Auch die Planung des Ressourceneinsatz im Bahn(fern)verkehr ist in einem mehrstufigen Planungsprozess eingebettet. Ausgehend von der *Linienplanung* in welcher Routen, mögliche Fahrzeugtypen und Frequenzen unter Berücksichtigung der vorhandenen Gleiskapazitäten zur Bedienung festgelegt werden, werden im *Train Timetabling Problem* Fahrpläne für (anonyme) Züge erstellt. Daran anschließend müssen die vorhandenen, lokale Infrastrukturen konkret betrachtet werden und im *Train Platforming Problem* den Fahrten Plattformen an Stationen zugewiesen werden. Im *Rolling Stock Circulation Problem* werden den auszuführenden Fahrten dann Triebfahrzeuge und Waggons bzw. aggregierte Zugfolgen (wie z.B. durch einen ICE oder TGV repräsentiert) zugewiesen. Auch in diesem Schritt bedarf es der Berücksichtigung von möglichen Gleisbelegungen und der Wagenreihungen an einzelnen Stationen. Für das Abstellen und Parken von ungenutzten Zugteilen oder zur Wartung außerhalb der Stationen ist das *Train Unit Shunting Problem* zu lösen, bei dem zur vorhandenen Infrastruktur passende Rangierfolgen erstellt werden. Die *Dienstplanung* widmet sich der (anonymen) Diensterstellung für Fahrer und Schaffner. Vertiefende Informationen zu einzelnen Planungsschritten sowie eine Literaturübersicht werden in [Caprara et al., 2007] gegeben.

Die Anzahl von Veröffentlichungen über Forschung zum Bahnverkehr insgesamt und insbesondere mit Berücksichtigung von Robustheitsaspekten ist in den letz-

ten Jahren stark gestiegen. Dies ist maßgeblich auf das ARRIVAL-Projekt zurückzuführen, welches zu großen Teilen im Rahmen eines Zukunftsprogrammes der EU gefördert wurde und an welchem in den Jahren 2006 bis 2009 europaweit zwölf Universitäten und Forschungsreinrichtungen, sowie das französische Bahnunternehmen SCNF beteiligt waren. Gemäß der Bezeichnung (ARRIVAL = Algorithms for Robust and online Railway optimization: Improving the Validity and reliAbility of Large scale systems) fokussiert das Projekt auf proaktive und reaktive Ansätze, um Störungen während der Planausführung bewältigen zu können. Im Rahmen des Projektes wurden rund 240 Dokumente veröffentlicht, von denen mittlerweile viele auch in wissenschaftlichen Journals publiziert sind. Schwerpunkte der Veröffentlichungen zu Robustheit im Bahnverkehr liegen auf Gestaltung und Analyse von Fahrplänen inklusive Infrastrukturbenutzung (siehe Literaturüberblick von [Cacchiani und Toth, 2012] zur Fahrplangestaltung) und im Online Störungsmanagement bzw. Rescheduling von Plänen. So verweisen auch [Caprara et al., 2007] auf einen Trend in der Bahnindustrie, von der Planung im Detail zu effektiver Echt-Zeit Kontrolle zu wechseln, da aufgrund der unvermeidbaren Störungen im komplexen System Bahn und den damit verbundenen Auswirkungen auf die Infrastrukturbelegung, große Teile operationaler Pläne nie ausgeführt werden.

Nachfolgend werden diejenigen Ansätze aus der Literatur diskutiert, die sich proaktiv mit der möglichen Propagation von Verspätungen durch den geplanten Ressourceneinsatz auseinandersetzen bzw. die proaktiv auf eine *Stabilität* bzw. *Flexibilität* der Pläne in der Ausführung fokussieren. Für weiterführende Informationen zu Lösungsansätzen für Störungsmanagement sei an dieser Stelle auf folgende Beiträge verwiesen: [Jespersen-Groth et al., 2009] bietet einen Einstieg in das Störungsmanagement im Bahnverkehr und beschreibt neben möglichen Akteuren auch die nötigen Aufgaben Fahrplananpassung, Anpassung der Einsatzpläne des Rollmaterials sowie Crew Rescheduling. Ergänzend dazu stellen [Cacchiani et al., 2014] eine Übersicht gegenwärtiger Forschung über Echtzeit-Recovery-Ansätze im Bahnverkehr zusammen. [Veelenturf et al., 2016] bieten eine aktuelle Diskussion zum Rescheduling von Crews im Fall von schwerwiegenden Störungen und stellen einen Lösungsansatz vor, der auf dem Konzept der Recoverable Robustness aufbaut (siehe Abschnitt 3.1.1).

Fokus: Umlaufplanung mit Fahrplanungsaspekten

[Cadarso und Marin, 2012] verfolgen eine integrierte Planung des Fahrplans und des Fahrzeugumlaufplanes. Dafür erweitern sie das in [Cadarso und Marin, 2011] vorgestellte Mehrgüter-Flussmodell zur Fahrzeugplanung um Fahrplanentscheidungen. Um größere Freiheitsgrade bei der Planung des Fahrzeugeinsatzes nutzen zu können,

erlauben die Autoren Frequenzverschiebungen auf einzelnen Linien bzw. Verschiebungen der Abfahrtszeiten unter Berücksichtigung des vermuteten Passagieraufkommens und Beibehaltung von Sicherheitsabständen in der Fahrtenfolge. Neben Kosteniminierung des Ressourceneinsatzes soll Robustheit der Pläne dadurch erreicht werden, dass sowohl Rangieroperationen zur Änderung der Fahrzeugzusammenstellung als auch Leerfahrten (insbesondere während der Rush-Hour-Zeiten) in der Zielfunktion bestraft werden. Um mehr Möglichkeiten für den Tausch von Fahrzeugen im Fall einer Störung bieten zu können, wird pro Linie nur ein Fahrzeugtyp eingesetzt. Die Autoren betrachten zwei Testfälle aus dem regionalen Bahnnetzwerk um Madrid. Eine Linie mit insgesamt 370 geplanten Fahrten und zwei Linien gemeinsam mit insgesamt 410 Fahrten. In beiden Fällen wird jeweils nur ein Fahrzeugtyp als Einzelverbund bzw. Doppelverbund zugelassen. Für die Planerstellung werden verschiedene Parametereinstellungen untersucht. Es ***zeigt*** sich, dass das Anpassen des Fahrplans sowohl eine größere Robustheit – gemessen an den Indikatoren – als auch eine höhere Kosteneffizienz der Zugzusammenstellungen erlaubt. Darüber hinaus wird deutlich, dass in der betrachteten Problemstellung Flexibilität in der Ausführung zu Lasten von Kosteneffizienz und Stabilität erkauft werden muss. Es können zwar mehr Tauschmöglichkeiten in den Plänen erzeugt werden, wofür aber auch zusätzliche kritische Leerfahrten eingefügt werden müssen, so dass der Einbau von potenziellen Tauschoperationen in die Pläne abzuwägen ist. Es ist allerdings anzumerken, dass die Robustheit der entstehenden Pläne allein anhand der gewählten Indikatoren beurteilt wird. Eine Validierung der Ergebnisse durch das Einbeziehen von z.B. Störungs- bzw. Verspätungsszenarien wird nicht durchgeführt.

Fokus: Umlaufplanung

[DeAlmeida et al., 2008] schlagen ein Vorgehen vor, um neben der Minimierung von Kosten auch Robustheitsaspekte in die Planung des Rollmaterials mit einzubeziehen. Die Autoren identifizieren angelehnt an Planungsverfahren im Flugverkehr drei Vorgehensschritte, um sowohl Kosteneffizienz als auch Robustheit bei den Zugumläufen zu gewährleisten: Erst wird ein Plan mit minimalen Kosten berechnet. In einem zweiten Schritt wird die Verspätungspropagation minimiert unter der Bedingung das etwas mehr Kosten für den Ressourceneinsatz als im kostenoptimalen Plan erlaubt sind, so dass Pufferzeiten im Plan erhöht werden können. In einem optionalen dritten Schritt könnten durch Retiming von auszuführenden Fahrten die Fahrzeugpläne so gestaltet werden, dass die Anzahl der Tauschmöglichkeiten zur Bedienung der Fahrten maximiert wird. Zur Umsetzung der ersten beiden Schritte konzentrieren sich die Autoren insbesondere auf die Nutzung eines Mehrgüter-Flussmodells zur

Umlaufplanung, welches auch beim französischen Bahnunternehmen SCNF eingesetzt wird und auf dem von [Ahuja et al., 2005] vorgestellten Flussmodell zur Umlaufplanung im Bahnverkehr basiert. Um eine mögliche Propagation von Verspätungen zu minimieren, werden der Einsatz zusätzlicher Ressourcen, sowie Leerfahrten und Rangierfahrten der Triebfahrzeuge für die Planung zugelassen. Aufgrund der Flussformulierung auf einem Time-Space-Netzwerk sind keine exakten Propagationseffekte einzelner Fahrzeugumläufe abbildbar, da die zuvor ausgeführten Fahrten nicht genau bekannt sind. Deshalb schlagen die Autoren eine Erweiterung der Formulierung vor. Mit dieser Erweiterung wird die Verspätungspropagation zu Beginn einer Fahrt basierend auf mutmaßlichen Vorgängerfahrten abgeschätzt. Um in der vorgeschlagenen Form Propagationseffekte abbilden zu können, sind in der Modellformulierung pro zu betrachtender Fahrt insgesamt sieben Entscheidungsvariablen und 24 zusätzliche Restriktionen zu berücksichtigen. Der Ansatz wird beispielhaft mit Daten einer (langen) Linie des Schnellbahnnetzes im Großraum Paris getestet. Die beiden betrachteten Tage umfassen 147 und 150 Fahrten. Für die Fahrten werden in Bezug zu ihrer Größe unabhängig voneinander hohe primäre Verspätungen generiert. Mit den generierten Verspätungen als Vorgabe wird das Modell gelöst. Es ***zeigt*** sich, dass zusätzliche Ressourcen eine drastische Reduktion der entstehenden propagierten Verspätungen ermöglichen. Statt zusätzlicher Fahrzeuge ermöglichen auch zusätzliche Leerfahrten eine (etwas kleinere) Verringerung der gesamten propagierten Verspätungen. Aus den Angaben wird jedoch nicht deutlich, ob es sich um den jeweils berechneten Zielfunktionswert zur Verspätungspropagation oder der tatsächlichen auf den konkret generierten Umläufen basierten Verspätungspropagation handelt. Darüber hinaus erscheint der Umgang mit der Weitergabe von Verspätungen im Mehrgüter-Flussmodell in dieser Form fraglich. Die Anzahl benötigter Entscheidungsvariablen und Restriktionen ist beträchtlich. Im Gegensatz zum Busverkehr mit nahezu unbeschränkten Leerfahrtmöglichkeiten sind in den betrachteten Problemstellungen nur wenige Leerfahrtmöglichkeiten vorgesehen, eine Weitergabe von Verspätungen über Leerfahrten wird aber erst gar nicht in der Formulierung abgebildet. Überdies werden bei der Propagation von Verspätungen in den Knoten des Netzwerks nur lokale Entscheidungen basierend auf Durchschnittswerten und der Minimierung der lokalen Verspätungsweitergabe getroffen werden, so dass die entstehende Propagation in den einzelnen Umläufen unter bzw. überschätzt wird.

[Cacchiani et al., 2012] beschreiben einen Ansatz, um Fahrzeugumläufe gegenüber den Auswirkungen großer Störungen robust zu gestalten, indem bereits im Vorfeld Recovery-Möglichkeiten im Umlaufplan berücksichtigt werden. Zielsetzung ist neben der Minimierung der geplanten Kosten, eine Minimierung der maximal benötigten

Recovery-Kosten. Dazu wird das mathematische Modell von [Fioole et al., 2006] zur kostenminimalen Umlaufplanung etwas vereinfacht und erweitert um das Einbeziehen von Störungsszenarien, die beim Lösen zu betrachten sind. Die Recovery-Kosten der einzelnen Szenarien werden gemessen in Bezug auf zusätzlich abzusagende Fahrten, neue Rangieroperationen, sowie Abweichungen der Planung des Rollmaterials am Tagesende. Für die Planung stehen Ersatzfahrzeuge zur Verfügung, deren Verteilung an den Stationen auch im nominalen Planfall berechnet wird. Zum Einbeziehen der Szenarien muss jede Restriktion des nominalen Teils des Modells zusätzlich für jedes Störungsszenario abgebildet werden sowie weitere Restriktionen und Entscheidungsvariablen eingeführt werden, um Recovery-Maßnahmen und -Kosten abbilden zu können. Um das entstehende mathematische Modell lösen zu können, wird ein heuristisches Lösungsverfahren vorgeschlagen. Die Anwendbarkeit des Lösungsansatzes wird am Beispiel einer Instanz getestet (eine Intercity-Linie mit acht Stationen und insgesamt 400 Fahrten zwischen den einzelnen Stationen). Als Szenarien werden beim Lösen maximal 28 ausgesuchte Störungsfälle betrachtet. In dem Szenarienset sind die potenziellen Störungen gleichmäßig über den Tag und die sieben Teilabschnitte der Linie verteilt und haben eine genau definierte Dauer im Bereich von einer bis sechs Stunden. Beim Lösen werden Pläne für das Betrachten von 0 bis 28 Szenarien erstellt. Die erstellten Pläne werden zusätzlich anhand von 3.500 generierten Szenarien evaluiert. Die Arbeit ***zeigt***, dass mit geringfügig zusätzlichen geplanten Kosten, eine günstigere Wiederherstellung des funktionierenden Betriebs erkauft werden kann, bei der weniger Fahrten zusätzlich abgesagt und die Recovery-Kosten geringer als im nominalen Fall liegen. Bei der Planerstellung wird bevorzugt eine Ausbalancierung des über die einzelnen Stationen verteilbaren Rollmaterials vorgeschlagen, um im Störungsfall ausreichend Tauschmöglichkeiten zu besitzen. Die Autoren schließen, dass das Einbeziehen eines in der Anzahl limitierten, aber repräsentativen Szenariensets bei der Planerstellung auch für größere Mengen an Szenarien geeignete Lösungen produziert. Dabei ist anzumerken, dass die möglichen Störungen in den 3.500 Szenarien zur Evaluation des Ansatzes jeweils gleichverteilt über Orte und Länge und Uhrzeit erstellt wurden. Vielmehr weist diese Arbeit auf technische Schwierigkeiten beim Bewerten und Einbeziehen einzelner Szenarien in den Lösungsprozess hin. Durch die zum Lösen benötigte Heuristik werden in einigen Fällen Lösungen produziert, die zu deutlichen Fehleinschätzungen der erforderlichen Recovery-Kosten führen können. Die Festlegung auf maximal 28 Szenarien für das in der Größe eingeschränkte Fallbeispiel ist überdies bedingt durch die für den Lösungsansatz beanspruchte Rechen- und Speicherkapazität.

Fokus: Dienstplanung

[Kroon und Fischetti, 2000] beschäftigen sich mit Crew Scheduling und implementieren mit Hilfe des TURNI System neue Dienstregeln für das niederländische Bahnunternehmen NS Reizigers, um Konsequenzen für die zukünftige – mit Methoden des Operations Research unterstützte – Planung abzuschätzen. Ziel der neuen Regelungen ist es, die Zuverlässigkeit im Bahnverkehr zu erhöhen, wobei die Dienste verschiedenen Qualitätsansprüchen wie Robustheit und Variabilität genügen sollen. Um speziell die Robustheit in der Ausführung zu erhöhen, werden zwei Maßnahmen umgesetzt. Statt wie bisher Zugführer und Schaffner getrennt zu planen, werden nun Teams betrachtet, die aus (anonymen) Schaffner und Zugführer bestehen. Somit soll die Wahrscheinlichkeit erhöht werden, dass alle benötigen Ressourcen zur Ausführung einer Fahrt pünktlich verfügbar sind. Da die Transfer- bzw. Umstiegszeit von Personal zwischen Zügen als kritischer Faktor für die Robustheit gesehen wird, ist zusätzlich eine minimale Zeit zur Verbindung zu beachten, und es werden die Anzahl der Fahrzeugwechsel von Personal begrenzt, die an Haltepunkten im Fahrtenverlauf eines Zuges (und nicht zu Beginn oder am Ende) stattfinden. Die Arbeit ***zeigt***, dass mit der Beschränkung auf eine Planungsebene mit dem vorgestellten Lösungsansatz Dienstpläne für große Planungsprobleme (ca. 2.000 Fahrten) in Variation der jeweiligen Zielsetzungen erstellt werden können. Die mit den vorgeschlagenen, grundlegenden Prinzipien zur Dienstbildung erreichbare Robustheit wird in den Tests nicht näher ausgeführt. Die Autoren verweisen auf den im Praxiseinsatz zu erbringenden Nachweis und schlagen den Aufbau eines Simulationsmoduls vor, um zukünftig die entstehenden Pläne zuvor genauer beurteilen zu können. Für den in der Praxis benötigen Anpassungsprozess der Regeln zur Diensterstellung sei auf [Abbink et al., 2005] verwiesen.

3.5 Zusammenfassung und Implikationen

Nachdem zu Beginn dieses Kapitels generelle Konzepte für die Berücksichtigung von Unsicherheit in Optimierungsansätzen in Abschnitt 3.1.1 und für die Berücksichtigung mehrerer Zielkriterien in Abschnitt 3.1.2 identifiziert wurden, wurde ein Überblick über existierende Forschung zur Umlauf- und Dienstplanung im ÖPNV mit der Zielsetzung Kosteneffizienz in Abschnitt 3.2 und dem Einbeziehen von Robustheitsaspekten in Abschnitt 3.3 sowie verwandter Forschung im Flug- bzw. Bahnverkehr in Abschnitt 3.4 gegeben. Bestimmendes Thema aktueller Forschung ist die Integration von Planungsaufgaben erweitert um die Suche nach Möglichkeiten, die enstehenden Pläne für ihre Ausführung robust zu gestalten.

Tabelle 3.1 fasst die betrachteten Ansätze aus den unterschiedlichen Problemdomänen Busverkehr im ÖPNV (Ö), Flugverkehr (F) und Bahnverkehr (B) kategorisierend zusammen. Dabei werden die für die Ressourceneinsatzplanung einbezogenen Planungsebenen, die Zielsetzung der Ansätze, das Einbeziehen der Auswirkungen von Verspätungsprogagation und die genutzte Lösungsmethodik gegenübergestellt. Als Planungsebenen werden Fahrplanung (F), Umlaufplanung (U) und Dienstplanung (D) aufgeführt. Im Flugverkehr werden in den Ansätzen auf der Umlaufplanungsebene entweder Fleet Assignment oder Aircraft Routing Entscheidungen getroffen, wodurch der Umfang im Vergleich zur Umlaufplanung im ÖPNV eingeschränkter ist. Die Fahrplanung im Bahnverkehr ist weiter umfassend als im ÖPNV, da dort auch die Belegung von Infrastruktur zu beachten ist (siehe Abschnitt 3.4). Bei der Zielsetzung nach Kostenminimierung wird unterschieden, ob Umlaufkosten (K_U) oder bzw. und Dienstkosten (K_D) berücksichtigt werden. In der Regel sind in den Ansätzen die entsprechenden Kostensätze so gewählt, dass die Anzahl der jeweils benötigten Ressourcen minimiert wird. In die Kosten werden je nach Ansatz auch beispielsweise Arbeitszeiten oder der Einsatz von Leerfahrten einbezogen. Für die Zielsetzung nach Robustheit werden folgende Einträge unterschieden: Mit fix werden diejenigen Ansätze gekennzeichnet, die durch eine feste Vorgabe von Pufferzeiten zwischen der Verknüpfung zweier Fahrten Robustheit berücksichtigen. Darüber hinaus wird unterschieden, inwiefern fiktive Verspätungskosten bzw. Strafkosten für das jeweils eingesetzte Unrobustheitsmaß (S_K), sowie konkrete Möglichkeiten zum Ressourcentausch (T) oder das Einbeziehen von Propagationseffekten (P) in der Zielsetzung nach Robustheit abgebildet werden. Besonders gekennzeichnet werden diejenigen Ansätze, die es erlauben einen Trade-Off zwischen der Minimierung von Kosten und Verspätungspropagation zu berücksichtigen (K/P). Dabei kann ein solcher Trade-Off auf vielfältige Weise abgebildet sein (siehe Abschnitt 3.1.2). Zusätzlich wird unterschieden, in welchem Umfang Verspätungspropagation während des Lösungsverfahrens berechnet und bei der Dienst- bzw. Umlauferstellung berücksichtigt wird. Verspätungspropagation kann für eine direkt folgende Fahrt bzw. einen Flug berücksichtigt werden (P_{f_1}), oder für eine gesamte Aufgabenkette bis die Propagation absorbiert wird (P_{f_k}). Darüber hinaus werden Ansätze gekennzeichnet, die bei der Berechnung von Propagationseffekten verschiedene Ressourcenebenen gemeinsam einbeziehen (P_{r_m}). Auf welche Weise die unterschiedlichen Planungsebenen in die Lösungsmethodik einbezogen werden, beschreiben die Einträge einzeln (ein), sequenziell (seq), teilintegriert ($pint$), iterativ ($iter$) oder integriert (int).

Die Übersicht in Tabelle 3.1 zeigt die Potenziale und Grenzen existierender Forschung zu einer robusten Planung des Ressourceneinsatzes. So zeichnen sich mehrere

Tabelle 3.1: Kategorisierte Übersicht existierender Ansätze aus der Literatur mit Bezug zu robuster Planung des Ressourceneinsatzes

Referenz	Domäne	Planungsaufgabe			Zielsetzung: Kosten		Zielsetzung: Robustheit				Trade-Off	Abbilden von Verspätungspropagation			Lösungsmethodik				
		F	U	D	K_U	K_D	fix	S_K	T	P	K/P	P_{f_1}	P_{f_n}	P_{r_m}	ein	seq	$pint$	$iter$	int
[Huisman et al., 2004][1]	Ö	–	x	–	x	–	x	–	–	–	–	–	–	–	x	–	–	–	–
[Naumann et al., 2011][1]	Ö	–	x	–	x	–	x	–	–	–	–	–	–	–	x	–	–	–	–
[Huisman et al., 2004][2]	Ö	–	x	–	x	–	–	x	–	–	–	–	–	–	x	–	–	–	–
[Kramkowski et al., 2009]	Ö	–	x	–	x	–	–	x	–	x	x	–	x	–	x	–	–	–	–
[Naumann et al., 2011]	Ö	–	x	–	x	–	–	x	–	x	x	x	–	–	x	–	–	–	–
[Huisman und Wagelmans, 2006][1]	Ö	–	x	x	x	x	x	–	–	–	–	–	–	–	–	–	–	–	x
[Huisman und Wagelmans, 2006][2]	Ö	–	x	x	x	x	–	x	–	–	–	–	–	–	–	x	–	–	x
[Ahmadbeygi et al., 2010][3]	F	x	–	–	–	–	–	x	–	x	–	–	x	x	x	–	–	–	–
[Burke et al., 2010][3]	F	x	x	–	–	–	–	x	x	x	–	x	–	–	–	–	–	x	–
[Lan et al., 2006]	F	–	x	–	–	–	–	x	–	x	–	–	x	–	x	–	–	–	–
[Weide et al., 2010][4]	F	–	x	x	–	x	–	x	–	–	–	–	–	–	–	–	–	x	–
[Dueck et al., 2012][4]	F	–	x	x	–	x	–	x	–	x	x	–	x	–	–	–	–	x	–
[Dunbar et al., 2012]	F	–	x	x	–	–	–	x	–	x	–	–	x	x	–	–	–	x	–
[Dunbar et al., 2014]	F	x	x	x	–	–	–	x	–	x	–	–	x	x	–	–	–	x	–
[Shebalov und Klabjan, 2006][5]	F	–	–	x	–	x	–	x	x	–	–	–	–	–	x	–	–	–	–
[Ionescu und Kliewer, 2011][6]	F	–	–	x	–	x	–	x	x	–	–	–	–	–	x	–	–	–	–
[Ehrgott und Ryan, 2002]	F	–	–	x	–	x	–	x	–	x	x	x	–	–	x	–	–	–	–
[Yen und Birge, 2006]	F	–	–	x	–	x	–	x	–	x	–	x	–	–	x	–	–	–	–
[Cadarso und Marin, 2012]	B	x	x	–	x	–	–	x	x	–	–	–	–	–	–	–	–	–	x
[DeAlmeida et al., 2008][7]	B	–	x	–	x	–	–	x	–	x	x	x	–	–	x	–	–	–	–
[Cacchiani et al., 2012]	B	–	x	–	x	–	–	x	–	–	–	–	–	–	x	–	–	–	–
[Kroon und Fischetti, 2000][8]	B	–	–	x	–	x	–	x	–	–	–	–	–	–	x	–	–	–	–
Kapitel 7	Ö	–	x	x	x	x	–	x	–	x	x	–	x	x	x	x	x	–	–
Kapitel 8	Ö	–	x	x	x	x	–	x	–	x	x	–	x	x	–	–	–	–	x
Kapitel 9	Ö	x	x	x	x	x	–	x	–	x	x	–	x	x	–	–	–	–	x

[1] Zu Vergleichszwecken eingesetzter Lösungsansatz, der einer rein kosteneffizienten Planung entspricht. Robustheit wird implizit dadurch berücksichtigt, dass feste Pufferzeiten zwischen Fahrten vorgegeben sind.
[2] Online-Ansatz, der die Ressourceneinsatzpläne schrittweise während der Ausführung erstellt.
[3] Retiming- und/ oder Rerouting-Ansatz. Die Ressourceneinsatzpläne sind als Eingabe gegeben.
[4] Keine Angaben über Dienstanzahl. Kosten verallgemeinert über betrachtete Instanzen.
[5] Maximal erlaubte Dienstkosten als Kostenrestriktion bei Zielsetzung nach Move-Up Crews.
[6] Bonus für eingeplante Tauschmöglichkeiten in Zielfunktion. Verwendbare Tauschmöglicheiten werden durch Preprocessing bestimmt.
[7] Maximal erlaubte Umlaufkosten als Kostenrestriktion bei Zielsetzung nach Minimierung von Verspätungspropagation.
[8] Keine Angaben über Zusammensetzung der betrachteten Dienstkosten. Robustheit durch Restriktionen bei der Diensterstellung.

Ansätze dadurch aus, dass sie in der tatsächlichen Planausführung auftretende Effekte – wie die mögliche Propagation von Verspätungen – abbilden und berücksichtigen. In einigen Ansätzen werden dabei auch beide Ressourcenebenen einbezogen. Im Flug- und Bahnverkehr existieren darüber hinaus Ansätze, die explizite Tauschmöglichkeiten von Ressourcen einplanen, um die Flexibilität des Ressourceneinsatzes in der Ausführung zu erhöhen. So sollen zumeist gegenüber größeren Störungen Reaktionsmöglichkeiten geschaffen werden. Dabei wird häufig die Erhöhung der Robustheit durch einen Mehreinsatz von Ressourcen erkauft.

Überdies lassen sich Grenzen insbesondere bei einer gemeinsamen Planung von Umläufen und Diensten identifizieren. Um die gemeinsamen Abhängigkeiten zwischen den beiden Planungsebenen zu berücksichtigen, werden zwar integrierte Modelle aufgestellt, die Robustheitsaspekte abbilden. Allerdings werden als Lösungsverfahren für diese integrierten Modelle iterative Vorgehensweisen gewählt. Dabei werden die Freiheitsgrade für die Planung des Ressourceneinsatzes einer Ebene jeweils immer von einer konkreten Lösung der anderen Ebene begrenzt. Darüber hinaus wird in den existierenden Ansätzen eine gleichzeitige Bestimmung von kosteneffizienten und robusten Ressourceneinsätzen nur eingeschränkt verfolgt. Zwar werden Robustheit und Kosteneffizienz als Zielgröße definiert, aber bei einer gemeinsamen Planung von Umläufen und Diensten (wie es beispielsweise bei einigen Ansätzen im Flugverkehr der Fall ist) werden die Kosten auf der Umlaufebene nicht weiter betrachtet. In den Ansätzen, welche eine exakte Propagation von Verspätungen über mehr als eine Ressourcenenbene abbilden, werden keinerlei Plankosten für den Ressourceneinsatz betrachtet.

Die Ausführungen dieses Kapitels haben gezeigt, dass die Ressourceneinsatzplanung im ÖPNV viele Gemeinsamkeiten mit der Ressourceneinsatzplanung anderer Domänen des öffentlichen Personenverkehrs besitzt und es Anknüpfungspunkte zwischen den verschiedenen Lösungsansätzen gibt. Allerdings existieren auch signifikante Unterschiede in den jeweils zu berücksichtigenden Rahmenbedingungen. Insbesondere die nahezu unbeschränkten Leerfahrtmöglichkeiten für Fahrzeuge und Transfermöglichkeiten für Fahrer sowie spezifische Dienst- und Pausenregelungen führen zu besonderen Problemstellungen im ÖPNV, für die Lösungsansätze anderer Problemdomänen in der Regel nicht adaptierbar sind. Um eine robuste und kosteneffiziente Planung zu ermöglichen, sollten die gezeigten Potenziale der Lösungsansätze zu robuster Ressourceneinsatzplanung anderer Domänen adäquat in Ansätze für die Umlauf- und Dienstplanung im ÖPNV einbezogen werden. Gleichzeitig sollten für eine umfassende Untersuchung die in der aktuellen Forschung vorhandenen Grenzen bei der Planung eines mehrschichtigen, d.h. eines mehrere Ressourcenebenen umfassenden Ressourceneinsatzes überwunden werden.

Kapitel 4

Forschungsbedarf und Zielsetzung

Die Übersicht über den aktuellen Stand der Forschung in Kapitel 3 zeigt den Fortschritt der letzten Jahre im Bereich Optimierung für die Ressourceneinsatzplanung im Personenverkehr. Neben dem Lösen großer realer Problemstellungen ist mittlerweile auch die Integration einzelner Planungsaufgaben fortgeschritten. Die Forschungsergebnisse sind darüber hinaus teilweise in kommerzielle Planungssysteme eingebaut und sind zur Unterstützung einer kosteneffizienten Planung in der Praxis nutzbar. Auf der anderen Seite werden Möglichkeiten und Notwendigkeiten für eine vertiefende Forschung zu robuster Planung im ÖPNV im Speziellen, sowie zu robuster und kosteneffizienter Planung im Personenverkehr im Allgemeinen ersichtlich. Diese betreffen insbesondere die gemeinsame Betrachtung des Einsatzes von Fahrzeugen und Fahrer und das Zusammenspiel der Ziele Kosteneffizienz und Robustheit. Tabelle 4.1 fasst die in Abschnitt 3.5 herausgearbeiteten Potenziale und Grenzen aktueller Lösungsansätze zusammen.

	Potenziale		**Grenzen**
(+)	Berücksichtigung von in der Praxis auftretenden Effekten (*Verspätungspropagation*)	(−)	Iterative Vorgehensweisen bei der Betrachtung mehrerer Ressourcenebenen (*Einschränkung der Freiheitsgrade*)
(+)	Betrachtung der auftretenden Effekte auf mehr als einer Ressourcenebene (*gemeinsame Umlauf- und Dienstplanung*)	(−)	Begrenzte gleichzeitige Betrachtung der Ziele Kosteneffizienz und Robustheit: Kosten nur auf einer Ebene oder sogar gar nicht (*Einschränkung der Ziele*)
(+)	Ergänzung von Stabilitätsaspekten um Flexibilitätsaspekte (*Abbilden von Tauschmöglichkeiten*)		

Tabelle 4.1: Zusammenfassung der in Abschnitt 3.5 beschriebenen Potenziale und Grenzen existierender Ansätze zu robuster Ressourceneinsatzplanung in Bus-, Flug- und Bahnverkehr

Die existierenden Ansätze und Untersuchungen zu robuster Ressourceneinsatzplanung im ÖPNV konzentrieren sich hauptsächlich auf die Umlaufplanung. Der Einsatz der Ressourcen Fahrzeuge und Fahrer steht jedoch durch die gemeinsame Bedienung der gleichen Fahrtenmenge in einem unmittelbaren Zusammenhang. Wie robust und kosteneffizient kann also der gesamte Ressourceneinsatz werden, wenn die Dienstplanung auf einem robusten Umlaufplan basiert? Und welche Auswirkungen hat ein kosteneffizienter Umlaufplan auf die robuste Planung des gesamten Ressourceneinsatzes? Welche Freiheitsgrade bieten demgegenüber Planungsmethoden, die Umläufe und Dienste gemeinsam erstellen können? Es ist insgesamt zu untersuchen, in welchem Ausmaß eine *Berücksichtigung der gegenseitigen Abhängigkeiten zwischen Umläufen und Diensten* bei der Planung zu einem robusten und kosteneffizienten Gesamtressourceneinsatz führen kann.

Darüber hinaus legen die Zeiten und Frequenzen der im Fahrplan vorgesehen Fahrten die Randbedingungen für die Umlauf- und Dienstplanung fest. Retiming-Ansätze, die bestehende Lösungen für den Ressourceneinsatz durch leichte Modifikation des zugrunde liegende Fahr- bzw. Flugplans verändern, versprechen durch das Umverteilen von bereits im Netzwerk eingeplanten Puffer gleichbleibende Plankosten bei höherer Robustheit. Zusätzlich zeigt sich, dass die Nutzung von Zeitfenstern bzw. Verschiebeintervallen für Fahrten während der eigentlichen Planerstellung einen großen Einfluss auf die Effizienz der berechneten Pläne haben kann. Es ist allerdings zu untersuchen, in welchem Ausmaß das *Einbeziehen solcher Fahrplanentscheidungen* während der Ressourceneinsatzplanung die konkurrierenden Zielsetzungen nach zusätzlicher Robustheit und Kosteneffizienz erfüllen kann. Die vorangegangen Überlegungen zusammenfassend ergibt sich die erste Forschungsfrage dieser Arbeit.

Forschungsfrage F1: *Welchen Einfluss haben unterschiedliche Planungsmethoden (sequenzielle, teilintegrierte oder integrierte Planung sowie Einbeziehen von Zeitfenstern) auf Robustheit und Effizienz der erstellten Pläne?*

Eine Zielsetzung, die gleichzeitig sowohl die Kosten des gesamten Ressourceneinsatzes minimiert als auch seine Robustheit erhöht, wurde in den bisherigen Forschungsarbeiten nicht verfolgt. Entweder ist der Zielkonflikt zwischen Kosten und Robustheit bei der Suche nur für eine Ressourcenebene steuerbar. Oder Kosten werden nicht mehr berücksichtigt, wenn robuste Pläne berechnet werden. Können kostenoptimale Pläne auch explizit robuster gestaltet werden? Es ist also zu evaluieren, ob und in welchem Maße sowohl kostenoptimale als auch robuste Pläne erstellt werden können. Daraus ergibt sich die nächste Forschungsfrage.

Forschungsfrage F2: *Wieviel Robustheit kann in kostenoptimalen Plänen realisiert werden?*

Um bei den Untersuchungen das Zusammenspiel der Zielkriterien Kosteneffizienz und Robustheit in der Umlauf- und Dienstplanung berücksichtigen und auch beeinflussen zu können, um also sowohl kosteneffiziente als auch robuste Pläne erstellen zu können, bedarf es der Entwicklung neuer Optimierungsmodelle und -methoden. Abbildung 4.1 stellt diese Zielsetzung der Arbeit grafisch dar. Die Suche nach Lösungsansätzen, welche die beiden Zieldimensionen gleichzeitig verfolgen, wird dabei im Rahmen dieser Arbeit unter dem Leitbild *Robuste Effizienz* zusammengefasst.

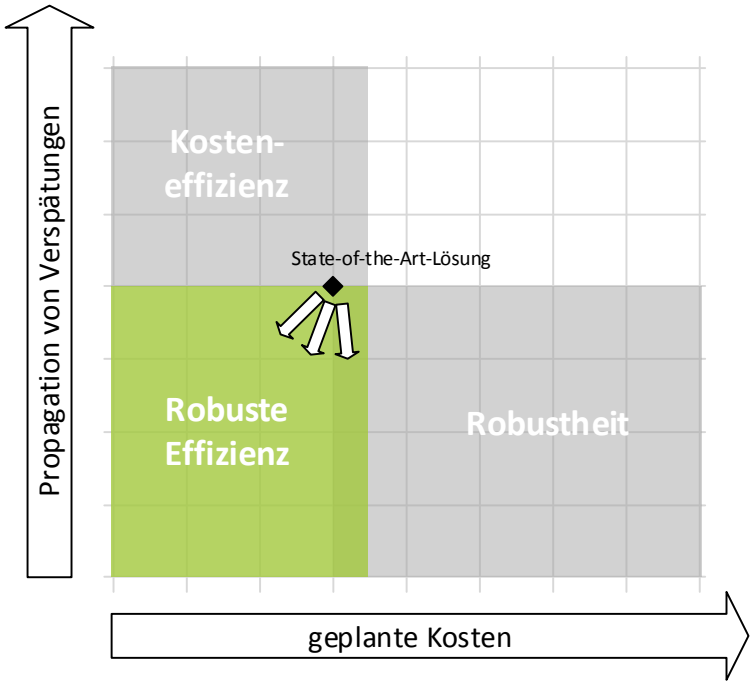

Abbildung 4.1: Robuste Effizienz der zu erstellenden Pläne als Zielsetzung für die Weiterentwicklung aktueller Lösungsansätze

Im folgenden Teil II dieser Arbeit wird die Realisierung einer ebenenübergreifenden, kosteneffizienten und robusten Planung von Fahrzeugumläufen und Fahrerdiensten schrittweise beschrieben. Die dabei vorgestellten Ansätze berücksichtigen in verschiedenem Maße die Verknüpfung der Planungsebenen Fahr-, Umlauf- und Dienstplanung. Durch die unterschiedlichen Grade der Verknüpfung kann dabei der Untersuchung nach einem Einfluss der Planungsmethoden auf Robustheit und Kosteneffizienz Rechnung getragen werden. Diese Untersuchung wird im anschließenden Teil III anhand aufeinander aufbauender Studien mit Blick auf die entstehenden Ressourceneinsatzpläne zur Beantwortung der folgenden Forschungsfrage konkretisiert.

Forschungsfrage F3: *Welche Planungsempfehlungen können für den robusten und kosteneffizienten Fahrzeug- und Personaleinsatz hergeleitet werden?*

Neben der Variation der *Planungsmethoden* sowie der Untersuchung der *Planstrukturen* der verschiedenartig erstellten Ressourceneinsatzpläne werden zur Beantwortung der Forschungsfragen auf einer zusätzlichen Untersuchungsebene die einzubeziehenden *Robustheitsaspekte* variiert. So umfasst die Evaluation der Robustheit des geplanten Ressourceneinsatzes zum einen den Aspekt der Stabilität. Zum anderen wird als zusätzlicher Aspekt die potenzielle Flexibilität während der Ausführung in die Bewertung einbezogen.

Teil II

Neue Lösungsansätze für Robuste Effizienz

Kapitel 5

Bewertung von Effizienz und Robustheit

Dieses Kapitel legt die für diese Arbeit wesentlichen Überlegungen zur Bewertung der operationalen Robustheit und Kosteneffizienz des geplanten Ressourceneinsatzes dar. Dazu werden zum einen Zielgrößen der zu entwickelnden Lösungsansätze diskutiert, die geeignet sind, die im vorangegangenem Kapitel 4 als Forschungsbedarf herausgearbeitete Zielsetzung nach Robuster Effizienz bei der Planerstellung forcieren zu können. Abbildung 5.1 gibt einen Überblick über mögliche zur Bewertung verwendbare Kennzahlen, gegliedert nach den beiden Zieldimensionen und unterteilt nach sich direkt aus der Planstruktur ergebenden Indikatoren einzelner Planeigenschaften sowie aggregierenden Maßen. Zum anderen werden in diesem Kapitel mit der Ausgestaltung einer Monte-Carlo basierten Simulation, sowie der Verwendung einer agentenbasierten Simulation Mittel zur Evaluation der erstellten Ressourceneinsatzpläne vorgeschlagen und ihr jeweiliger Einsatzzweck erörtert.

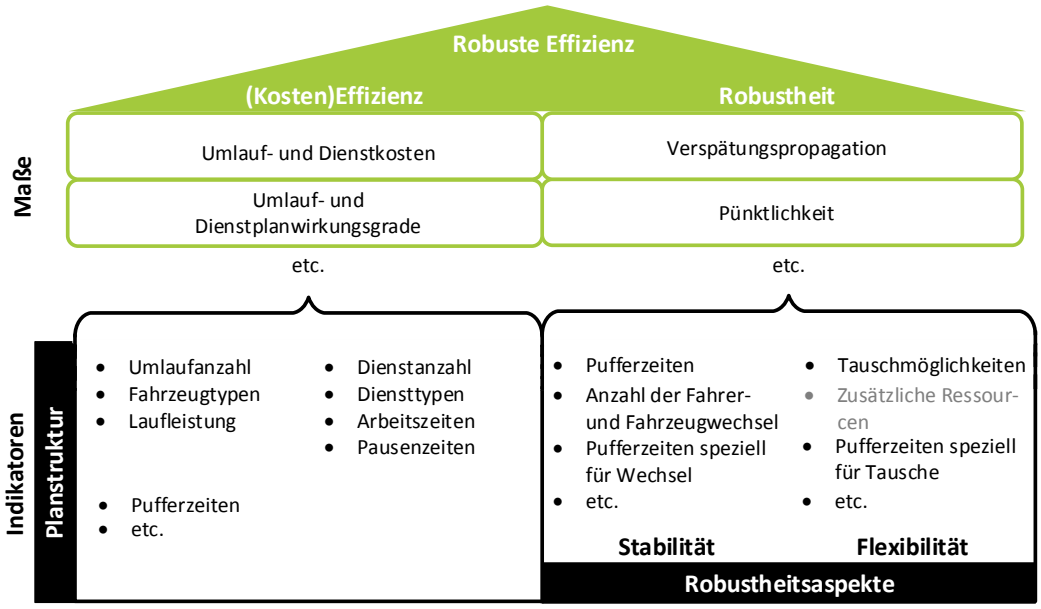

Abbildung 5.1: Kennzahlen zur Bewertung von Robustheit und Kosteneffizienz

Aufbau des Kapitels

In Abschnitt 5.1 wird zunächst kurz auf Kennzahlen zur Bewertung von Kosteneffizienz eingegangen. Anschließend werden in Abschnitt 5.2 Kennzahlen zur Bewertung von Robustheit diskutiert. Abschnitt 5.3 arbeitet zwei verschiedene Arten von Simulation zur Evaluation von Ressourceneinsatzplänen unter Einbeziehen der zuvor vorgestellten Indikatoren und Maße heraus. Eine Zusammenfassung und Implikationen für die weitere Arbeit werden in Abschnitt 5.4 gegeben.

5.1 Kennzahlen für Kosteneffizienz

Der Forschung zu rein kostenminimierender Planung folgend (siehe Abschnitt 3.2) werden im Rahmen dieser Arbeit zur Bewertung der geplanten Kosteneffizienz die als Ressourceneinsatz benötigte *Umlauf- und Dienstanzahl*, sowie von Art und Dauer des Einsatzes ableitbare Betriebskosten, die *Umlauf- und Dienstkosten*, als Kennzahlen genutzt.

Umlauf- und Dienstanzahl

Die Anzahl der eingeplanten Umläufe (#v) und damit einhergehend die Anzahl einzusetzender Fahrzeuge, sowie die Anzahl der eingeplanten Dienste (#d) und damit einhergehend die Anzahl der einzusetzenden Fahrer bilden in nahezu allen Lösungsansätzen zur Kostenminimierung des Ressourceneinsatzes die primäre Zielsetzung. Um auch die jeweiligen Arten der Ressourcen in die Bewertung einzubeziehen, werden die Anzahlen häufig mit Fixkostensätzen in Abhängigkeit verwendeter Fahrzeug- oder Diensttypen verrechnet. Für eine umfassende Bewertung des Ressourceneinsatzes können diese durch variable Betriebskosten ergänzt werden (siehe *Umlauf- und Dienstkosten*).

Umlauf- und Dienstkosten

Mit den (geplanten) Umlauf- und Dienstkosten werden einzelne Indikatoren bzw. Planausprägungen wie Reise- und Standzeiten oder Arbeits- und Pausenzeiten, sowie die eingeplante Umlauf- und Dienstanzahl, ebenso wie die verwendeten Fahrzeug- und Diensttypen, aggregiert und zusammenfassend bewertet.

Die *Umlaufkosten* setzen sich gewöhnlicherweise aus fixen Kosten für jedes Fahrzeug und variablen Kosten für die zurückzulegende Fahrtstrecke, sowie Reise- und Standzeiten außerhalb eines Depots zusammen. Dabei werden durch die Fixkosten (anteilige) Anschaffungskosten repräsentiert, die vom tatsächlichen Einsatzumfang

unabhängig anfallen. Die *Dienstkosten* setzen sich im Allgemeinen aus fixen Kosten für Löhne in Abhängigkeit möglicher Diensttypen und variablen Kosten für Arbeitszeit oder Überstundenzeit zusammen.

Die Summe aus Umlauf- und Dienstkosten wird in dieser Arbeit zur Bewertung der geplanten Kosteneffizienz des gesamten Ressourceneinsatzes verwendet. Ergänzt um die Angabe der benötigten Umlauf- und Dienstanzahl ist eine klare Unterscheidung bezüglich der Effizienz zwischen den Plänen möglich.

Wirkungsgrade

In der Praxis werden oft weitere Kennzahlen zur Bewertung des Ressourceneinsatzes verwendet. So gibt der *Umlaufplanwirkungsgrad* (bzw. auch Fahrplanwirkungsgrad genannt) in der zeitbezogenen Variante den Anteil der durch den Fahrplan vorgegebenen Einsatzzeit zur Personenbeförderung an der gesamten zu leisteten Einsatzzeit der Umläufe wieder. Analog wird in der wegbezogenen Variante der Anteil der vorgegebenen Nutzstreckenlänge an der gesamten zu leistenden Streckenlänge der Umläufe berechnet. Der *Dienstplanwirkungsgrad* setzt die Summe aller Umlaufdauern ins Verhältnis zur Summe aller Dienstdauern. Beim Vergleich unterschiedlicher Ressourceneinsatzpläne liefern beide Kennzahlen jeweils für eine Ressourcenebene nachvollziehbare Aussagen über die Auslastung des Ressourceneinsatzes. Für die Bewertung des gesamten Ressourceneinsatzes eignet sich die separate Betrachtung allerdings nur bedingt, da Verbesserungen eines Grades nicht zwangsläufig eine Verbesserung des Gesamtergebnisses darstellen.[1] [Borndörfer et al., 2002] bilden deshalb mit dem *Planbauwirkungsgrad* durch Multiplikation ein Produkt aus dem Umlaufplan- und Dienstplanwirkungsgrad. So wird als Gesamtwirkungsgrad des Ressourceneinsatzes der Anteil der durch den Fahrplan vorgegebenen Nutzzeit an der Summe aller Dienstdauern gemessen.

Die genannten Wirkungsgrade können allerdings auch falsche Fährten für die Bewertung des Ressourceneinsatzes legen. Da z.B. die Fahrzeug- oder Fahreranzahlen bzw. entsprechende Fixkosten nicht berücksichtigt werden, müssen hohe Wirkungsgrade nicht zwangsläufig hoher Kosteneffizienz entsprechen. Vielmehr können diese auch durch einen Mehreinsatz an Ressourcen erreicht worden sein. Deshalb liegt in dieser Arbeit der Fokus zur Bewertung von Kosteneffizienz auf den zuvor genannten, besser handhabbaren Kennzahlen.

[1] Beispielsweise können für einen hohen Dienstplanwirkungsgrad die Standzeiten in den Umläufen so erhöht werden, dass vollständige Dienstpausen auf den Fahrzeugen ermöglicht werden, was allerdings zu einer Verschlechterung des Umlaufplanwirkungsgrades führt (vgl. [Borndörfer et al., 2002]).

5.2 Kennzahlen für Robustheit

Die Diskussion des Standes der Forschung zur robusten Ressourceneinsatzplanung im ÖPNV in Abschnitt 3.3 und zu weiteren Verkehrsdomänen in Abschnitt 3.4 hat gezeigt, dass zur Bewertung von Robustheit viele unterschiedliche (quantitative) Indikatoren in den Zielsetzungen der einzelnen Lösungsansätze und Maße bei der Beurteilung der erstellten Lösungen verwendet werden. Dabei werden die in Abschnitt 2.5 beschriebenen Robustheitsaspekte der Stabilität und Flexibilität des Ressourceneinsatzes unterschieden und in den Ansätzen gesondert adressiert. Der Überblick in der zu Beginn dieses Kapitels aufgeführten Abbildung 5.1 fasst wesentliche Kennzahlen zur Robustheitsbewertung zusammen.

Die Kennzahlen lassen sich gruppieren in unmittelbar aus der Planstruktur ablesbare Indikatoren und ermittelbaren, aggregierenden Maßen, die zur Berechnung zusätzliche Informationen wie beispielsweise mögliche Verspätungsdauern einzelner Aufgaben bzw. Fahrten benötigen. Während sich durch die Berücksichtigung von Indikatoren einzelne Planausprägungen steuern lassen, z.B. Pufferzeiten für Fahrer- und Fahrzeugwechsel, kann ein Robustheitsmaß auch dazu benutzt werden, die operationalen Kosten, die durch Verspätungen entstehen, abzuschätzen bzw. in der Zielfunktion des Lösungssansatzes abzubilden.

Als grundlegend für die planimanente Robustheit des (mehrschichtigen) Ressourceneinsatzes können sicherlich die vorhandene *Pufferzeit* sowie die Vernetzung der Ressourcen durch eingeplante *Dienst- bzw. Umlaufwechsel* angesehen werden. Wie zu Beginn dieser Arbeit in Kapitel 1 ausgeführt, ist allerdings bei der Planerstellung zu klären, wie Pufferzeiten genau verteilt und in welcher Form Planaktivitäten der unterschiedlichen Ressourcen vernetzt werden können, um Pläne robust gegenüber Störungen und daraus resultierenden Verspätungen zu gestalten. Dabei liegt die Erhöhung von Robustheit in der Vermeidung von Verspätungspropagation während der Planausführung.

Im Folgenden werden mit *Verspätungspropagation* und *Pünktlichkeit* Robustheitsmaße, sowie als spezieller Indikator für Flexibilität *Tauschmöglichkeiten* diskutiert. Diese Kennzahlen umfassen (und beurteilen daher auch) zur Bewertung der Robustheit die passende Verteilung von Pufferzeiten bzw. Vernetzung der Planaktivitäten unterschiedlicher Ressourcen.

5.2.1 Verspätungspropagation

Die ursprünglichen, primären Verspätungen liegen grundsätzlich außerhalb der Kontrolle der ÖPNV-Unternehmen. Der Umgang mit diesen Verspätungen wird allerdings

durch die erstellten Planstrukturen festgelegt und vorgegeben (siehe Abschnitt 2.4). Um einen Plan bezüglich seiner Robustheit beurteilen zu können, bedarf es also einer Kennzahl, die angibt, inwiefern die primären Verspätungen abgebaut werden können bzw. inwiefern diese Verspätungen Auswirkungen auf die weitere Planausführung haben. Indem die Propagation von Verspätungen quantitativ erfasst wird, ist es möglich, den direkten Einfluss von Störungen bzw. primären Verspätungen basierend auf den vorhandenen Planstrukturen abzuleiten.

Das Abbilden von Verspätungspropagation wird in der Literatur bereits in einigen Lösungsansätzen der Ressourceneinsatzplanung zur Bewertung von Robustheit eingesetzt. Für die Grenzen einzelner Ansätze beim Umgang mit möglicher Verspätungspropagation (z.B. Abbilden nur zwischen zwei Fahrten bzw. Flügen) sei auf die Übersicht in Tabelle 3.1 sowie die zugehörige Diskussion der Literatur in den Abschnitten 3.3 und 3.4 verwiesen.

In dieser Arbeit wird zur Bewertung der Robustheit die Summe der (potenziellen) Verspätungspropagation des geplanten Ressourceneinsatzes verwendet, bzw. davon ableitbar die durchschnittliche Verspätungspropagation pro Servicefahrt. Zu beachten ist dabei eine Unterscheidung der Bewertung bei der Planerstellung und bei der Evaluation eines erstellten Ressourceneinsatzplanes. In der Regel sind erst im letztgenannten Fall im mehrschichtigen Ressourceneinsatz alle Abhängigkeiten zwischen den Ressourcen bekannt und die potenzielle Verspätungspropagation kann vollständig gemessen werden (siehe *Propagationsmodell* in Abschnitt 5.3.1). Daher wird bei der einzelnen Umlauf- oder Dienstbildung die potenzielle Verspätungspropagation in Abhängigkeit der auszuführenden Fahrten bzw. Aufgaben anhand einer *Propagationsformel* bestimmt, die im Folgenden erläutert wird.

Bei der Planerstellung wird in dieser Arbeit ähnlich wie in Ansätzen zur Planung im Flugverkehr (z.B. [Ehrgott und Ryan, 2002]) eine Kennzahl über die mögliche Propagation von Verspätungen einzelner Ressourceneinsätze verwendet, die nachfolgend auch als (Nicht)-Robustheitsmaß bezeichnet wird. Dabei gilt, je geringer das Maß und somit die gemessene Verspätungspropagation, desto höher ist die Verspätungstoleranz der berechneten Umläufe oder Dienste. Im Vergleich zum von [Ehrgott und Ryan, 2002] verwendeten Nicht-Robustheitsmaß beinhaltet das vorgeschlagene Maß Propagationseffekte über mehrere Fahrten bzw. Aufgaben. Darüber hinaus werden nicht allein Propagationseffekte einbezogen, die bei Fahrer- bzw. Fahrzeugwechseln entstehen. Da im Rahmen dieser Arbeit sowohl Umlauf- als auch Dienstpläne unter Robustheitsaspekten erstellt werden sollen und grundsätzlich nicht als gegeben angenommen werden, sind die gemeinsam von einem Umlauf- bzw. Dienst auszuführenden Fahrten nicht vorbestimmt. Vielmehr bestehen die Freiheitsgrade der Planung darin,

auch diese Fahrtensequenzen (Dienststücke) robust zu gestalten und so die gesamten Propagationseffekte im Ressourceneinsatz zu minimieren, statt allein die durch Wechsel induzierten Verspätungen.

Sei M die Menge der zulässigen Fahrzeugumläufe oder Fahrerdienste und sei $\mathcal{T}' = \{1, 2, .., m\}$ die Menge der Servicefahrten oder die Menge der auszuführenden Aufgaben. Dann repräsentiert das (Nicht)-Robustheits-Maß r_m die Verspätungspropagation für jeden Umlauf bzw. Dienst $m \in M$, der eine Menge T_m von Fahrten/Aufgaben ausführt, für die $T_m \subseteq \mathcal{T}'$ gilt. Für die Berechnung von r_m gilt:

$$r_m = \sum_{i=1}^{T^m-1} p(t_i, t_{i+1}) \tag{5.1}$$

Mit $p(t_i, t_{i+1})$ definiert als:

$$p(t_i, t_{i+1}) = max\{0, PAT(t_i) + PD(t_i) + p(t_{i-1}, t_i) - PDT(t_{i+1})\} \tag{5.2}$$

wobei $p(t_0, t_1) = 0$, da dies die Verspätungspropagation der ersten Fahrt beschreibt. Falls die erste Fahrt bereits verspätet ist, ist $p(t_0, t_1)$ größer als null. $PAT(t_i)$ bezeichnet die geplante Ankunftszeit von Fahrt t_i und $PDT(t_{i+1})$ bezeichnet die geplante Abfahrtszeit der Fahrt t_{i+1}. Die Differenz entspricht der zwischen den Fahrten zur Verfügung stehenden Pufferzeit. Diese wird eingesetzt, um vor Beginn der Fahrt t_{i+1} die erwartete primäre Verspätung $PD(t_i)$ von Fahrt t_i sowie eine möglicherweise bereits erfolgte propagierte Verspätung $p(t_{i-1}, t_i)$ für Fahrt t_i zu absorbieren. Die erwartete primäre Verspätung einer Fahrt ist dabei abhängig von (historischen) Verspätungsinformationen, die z.B. durch eine Menge von Verspätungsszenarien gegeben sein kann. Die Szenarien können dabei echten Verkehrstagen entsprechen oder aufgrund von stochastischen Verteilungen generiert worden sein, die mögliche Verspätungsverteilungen repräsentieren. $p(t_i, t_{i+1})$ entspricht insgesamt der propagierten Verspätung für Fahrt t_{i+1}. Konsequenterweise entspricht r_m somit der Summe der pro Ressource m auftretenden sekundären Verspätungen.

5.2.2 Pünktlichkeit

Eng mit der Messung von Verspätungen bzw. Verspätungspropagation verbunden sind Aussagen zur *(erwarteten) Pünktlichkeit* des eingeplanten Ressourceneinsatzes. Dabei werden in der Regel geringe Verspätungen der geplanten Abfahrts- bzw. Ankunftszeiten unterhalb eines bestimmten Schwellenwertes toleriert und die davon betroffenen Fahrten weiterhin als pünktlich angesehen. Die Messung der Pünktlich-

keit ist abhängig vom definierten Toleranzbereich und liefert daher im Gegensatz zur reinen Messung von Verspätungspropagation eine weniger präzise Aussage über die Robustheit des Ressourceneinsatz. Es sei denn, bei der Messung wird keinerlei Verspätung toleriert. Die Bewertung der Pünktlichkeit wird dann allerdings bereits durch die gemessene Verspätungspropagation wiedergegeben. Da im Rahmen dieser Arbeit die sich aus dem geplanten Ressourceneinsatz ergebenenden, planimmanenten Verspätungen im Vordergrund stehen und nur eine präzise Bewertung der Robustheit des Ressourceneinsatzes einen genauen Vergleich der vorgeschlagenen Lösungsansätze ermöglicht, liegt der Fokus der Untersuchungen auf Analysen der genauen Verspätungspropagation statt auf Aussagen zur Pünktlichkeit.[2]

5.2.3 Tauschmöglichkeiten von Ressourcen

Durch Eingriffe der Disponenten in die Planausführung kann auf Störungen bzw. primäre Verspätungen reagiert werden, um Propagationseffekte zu verhindern (siehe Abschnitt 2.4). Der Umgang mit Verspätungen während der Planausführung kann dabei auch durch (kostenneutrale) planimmanente Änderungsmöglichkeiten beeinflusst werden. Diese dienen dazu, vorhandene Pufferzeit zur Absorption von Verspätungen umzuverteilen bzw. einzelne geplante Aufgabensequenzen anzupassen (siehe Abschnitt 2.5). Das Ausmaß einer solchen Flexibilität des Ressourceneinsatzes kann durch Indikatoren beschrieben werden, die die Anzahl der Tauschmöglichkeiten von Aufgabenzuordnungen zu Umläufen und Diensten wiedergeben, ebenso wie die Anzahl der Möglichkeiten zum Tausch der Zuordnungen von Umläufen zu Diensten (und umgekehrt). Die Schaffung von Tauschmöglichkeiten wird in der Literatur insbesondere bei Lösungsansätzen im Flugverkehr untersucht, siehe z.B. [Shebalov und Klabjan, 2006] für die Dienstplanung bzw. [Burke et al., 2010] für die Flugzeugumlaufplanung. Zielsetzung dabei ist es, zur Erhöhung der Robustheit eine möglichst hohe Anzahl Tausche (engl. Swaps) für unterschiedliche Aufgaben während der Planausführung zu ermöglichen.

Im Rahmen dieser Arbeit werden *Tauschpunkte* und *Tauschmöglichkeiten* unterschieden. An Tauschpunkten ist es räumlich und zeitlich gesehen möglich, Ressourcen

[2] Ein Beispiel für eine umfassende Pünktlichkeitsdefinition liefert die Qualitätsmessung für die Berliner Verkehrsbetriebe: So liegt der Toleranzbereich für eine pünktliche Fahrt im Busverkehr in Berlin bei bis zu dreieinhalb Minuten nach der geplanten Abfahrtszeit (und sogar eine Verfrühung um eineinhalb Minuten ist erlaubt). Zusätzlich kommt es seit 2014 für die Messung nicht mehr auf das aus Sicht des Betriebes korrekt eingesetzte Fahrzeug, sondern auf das generelle Erscheinen eines Fahrzeuges an (siehe [Center Nahverkehr Berlin, 2017]). Dahingegen gilt beispielsweise für die Münchner Verkehrsgesellschaft eine Fahrt nur als pünktlich, wenn sie mit einer Verspätung kleiner gleich zwei Minuten beginnt (siehe [Münchner Verkehrsgesellschaft mbH, 2017].

bzw. Fahrtenzuordnungen zu Ressourcen zu tauschen. Voraussetzung dafür ist, dass sich die eingeplanten Pufferzeiten zwischen den zu tauschenden Fahrten- bzw. Aufgabensequenzen zweier Ressourcen (einer Ressourcenebene) an einem Ort zeitlich überschneiden. Jeder Tauschpunkt bietet mindestens eine Tauschmöglichkeit. Die gesamte Anzahl der Tauschmöglichkeiten des Ressourceneinsatzes ergibt sich aus den Kombinationsmöglichkeiten zur Verknüpfung der Aufgabensequenzen unterschiedlicher Ressourcen (einer Ressourcenebene) an den potenziellen Tauschpunkten. Abbildung 5.2 gibt ein Beispiel für einen Tauschpunkt zur Vermeidung von Verspätungspropagation bei unterschiedlich geplantem Ressourceneinsatz. In Bild I kann die Zuordnung der Fahrten- bzw. Aufgabensequenzen A-X in A-Y getauscht werden. In Bild II kann die Zuordnung B-X in B-Y getauscht werden. In beiden Fällen existiert genau eine Tauschmöglichkeit.

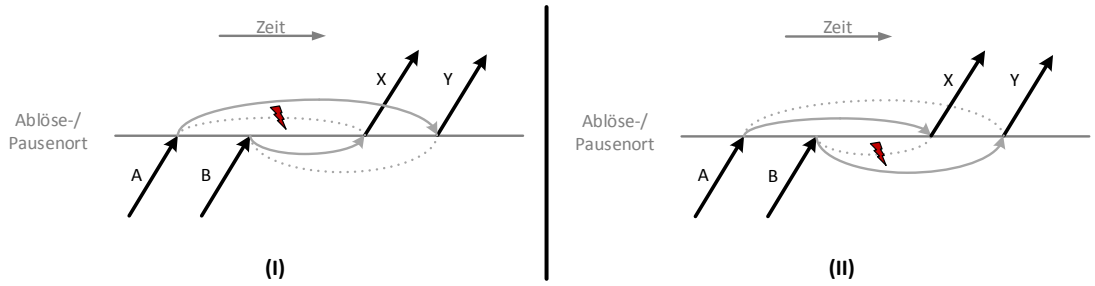

Abbildung 5.2: Tauschpunkt zur Umverteilung von Aufgabensequenzen und Pufferzeiten zur Vermeidung von Verspätungspropagation bei unterschiedlich geplanten Aufgabenzuordnungen (I) und (II)

Bei der gemeinsamen Betrachtung von Umläufen und Diensten sind potenzielle Tauschpunkte eng mit im Plan vorgesehenen Ablösepunkten verknüpft, da diese aus Planungssicht Möglichkeiten zum Umlauf- bzw. Dienstwechsel darstellen (siehe Abschnitt 2.3). Für kostenneutrale Änderungsmöglichkeiten bieten sich mögliche Tauschpunkte insbesondere nach Ende bzw. vor Beginn eines Dienststückes an. Da das Ende eines Dienststückes den Abschluss einer zusammengehörigen Aufgabensequenz eines Umlaufs und Dienstes repräsentiert, in der Regel gefolgt von einer Pause auf Dienstebene und/oder einem eingeplanten Dienstwechsel auf Umlaufebene, lassen sich an diesen Stellen leicht Zuordnungen von Dienststücken und damit Diensten zu Umläufen tauschen. Dabei müssen keine Änderungen der in den Dienststücken zusammengefassten Aufgaben vorgenommen werden. Insbesondere für den Tausch ganzer Fahrtensequenzen auf Umlaufebene (und damit auch für den Tausch von Dienstzuordnungen zu Umläufen) können durch vorgeplante Aufenthalte bzw. Parken im Depot Tauschpunkte vorliegen.

Die Bedingungen für das Vorliegen eines Tauschpunktes zum Umverteilen von Pufferzeiten zur Absorption von Verspätungspropagation werden durch die folgenden Ungleichungen (5.3) formal beschrieben. Die Definition erfolgt aus Sicht der Umlaufplanungsebene und eine Sequenz von Fahrten bzw. Aufgaben bezeichnet dabei die Fahrten bzw. Aufgaben mindestens eines Dienststückes. Eine Definition auf Dienstplanungsebene erfolgt analog, dabei sind jedoch zusätzlich Regeln zur Dienstbildung wie Mindestpausenzeiten und Dienstdauern zu berücksichtigen. Sei M die Menge der zulässigen Fahrzeugumläufe (oder Fahrerdienste) und sei H die Menge aller Haltestellen, Pausenorte und/oder Depots. Sei zusätzlich A^h die Menge aller Ankünfte von Fahrtensequenzen und D^h die Menge aller Abfahrtssequenzen an einem Ort $h \in H$. Ein Tauschpunkt für einzelne Ressourceneinsätze $m \in M$ an Ort h liegt vor, wenn für die Aufgabensequenzen $a_m^h \in A^h$ und $d_m^h \in D^h$, Aufgabensequenzen a_n^h und d_n^h von mindestens einem anderen Ressourceneinsatz $n \in M \setminus \{m\}$ existieren, so dass gilt:

$$PAT(a_n^h) < PDT(d_m^h) < PDT(d_n^h) \tag{5.3}$$

Dabei repräsentiert $PAT(a^h)$ die geplante Ankunftszeit und $PDT(d^h)$ die geplante Abfahrtszeit der jeweils betrachteten Fahrten- bzw. Aufgabensequenzen $a^h \in A^h$ und $d^h \in D^h$ am Ort $h \in H$. Anzumerken ist, dass die geplante Ankunftszeit von $PAT(a_m^h)$ größer, kleiner oder gleich $PAT(a_n^h)$ sein kann. Entscheidend für die Möglichkeit Pufferzeiten umzuverteilen, ist die zeitliche Differenz zwischen den geplanten Abfahrtszeiten $PDT(d_m^h)$ und $PDT(d_n^h)$. Bei gleichen Zeiten ist keine Umverteilung möglich und daher kein echter Tauschpunkt zur Absorption von Verspätungspropagation vorhanden.

5.3 Simulation zur Evaluation der Umlauf- und Dienstpläne

Im Rahmen dieser Arbeit wurden zwei Simulationsarten unterschiedlichen Umfangs getestet, um Umlauf- und Dienstpläne bezüglich ihrer Robustheit während der Ausführung zu evaluieren. Die eine basiert auf ereignisorientierter *Monte-Carlo Simulation*, die andere entspricht einer *agentenbasierten Simulation*. Für genaue Begriffsbestimmungen sei jeweils auf die folgenden Abschnitte verwiesen (sowie auf Abschnitt 3.1). Beide Simulationsvarianten dienen dazu, Störungen bzw. primäre Verspätungen des Ressourceneinsatz und den Umgang damit zu simulieren. Sie unterscheiden sich

jedoch fundamental hinsichtlich der Generierung von Verspätungen und des modellierten Abstraktionsgrades.

Während sich die Monte-Carlo basierte Simulation auf der wie in den Abschnitten 2.2 und 2.3 zur Ressourceneinsatzplanung beschriebenen Abstraktionsebene bewegt, erfordert die agentenbasierte Simulation eine deutlich detailliertere Abbildung der Realität. Neben dem Ressourceneinsatz wird explizit die Umwelt (d.h. das Straßennetz, andere Verkehrsteilnehmer, sowie die Passagiere) modelliert, in der sich dieser bewegt. Verspätungen entstehen dabei durch das Verhalten und die Interaktion der abgebildeten Akteure. Bei der Monte-Carlo basierten Simulation wird eine solche Umwelt weitgehend abstrahiert. Störungen bzw. primäre Verspätungen während der Ausführung des Ressourceneinsatzes werden anhand von Wahrscheinlichkeitsverteilungen über ihr potenzielles Auftreten erzeugt. Mit der Monte-Carlo basierten Simulation wird passgenau für diese Arbeit der Fokus auf die grundlegende Evaluation der Robustheit und das Herausarbeiten zentraler Effekte von Verspätungspropagation der Umlauf- und Dienstpläne gelegt. Die agentenbasierte Simulation dahingegen ist nicht allein für die Evaluation der Umlauf- und Dienstpläne ausgelegt. Sie ermöglicht bei Bedarf das explizite Einbeziehen der Passagiersicht in die Evaluation und kann damit eine weitergehende Perspektive auch auf angebotsorientierte Aspekte der Steuerung (z.B. Anschlusssicherung) sowie der Planung (z.B. auch Fahrplanung) eröffnen.

Im Folgenden wird die für die Experimente verwendete Monte-Carlo basierte Simulation ausführlich in Abschnitt 5.3.1 und der Umgang mit der agentenbasierten Simulation kurz in Abschnitt 5.3.2 vorgestellt.

5.3.1 Monte-Carlo-basierte Simulation

Dieser Abschnitt beschreibt die zur Evaluation des geplanten Ressourceneinsatzes entwickelte diskrete, ereignisorientierte Monte-Carlo basierte Simulation. Sie ermöglicht es, potenzielle Verspätungspropagation beim Auftreten von Störungen abzubilden und darüber hinaus Dispositionsentscheidungen bei der Ausführung des Ressourceneinsatzes zu berücksichtigen. Mit zunehmender Anzahl von Experimenten, die dem Durchspielen verschiedener Realitäten anhand von generierten Zufallszahlen entsprechen, konvergieren die gemittelten Ergebnisse gegen die theoretischen Eintrittswahrscheinlichkeiten bzw. Erwartungswerte (siehe auch „Exkurs Simulation“ in Abschnitt 3.1).

Eingabe für die Simulation bilden Analysen von Verspätungsdaten sowie daraus abgeleitete mögliche Verspätungsverteilungen, Fahrpläne, sowie die mit den Optimierungsmodellen bzw. -methoden aus den folgenden Kapiteln 6 bis 10 errechneten

Umlauf- und Dienstpläne. Abbildung 5.3 gibt einen Überblick über das entwickelte Simulationsframework und über seine Nutzung im Rahmen dieser Arbeit.

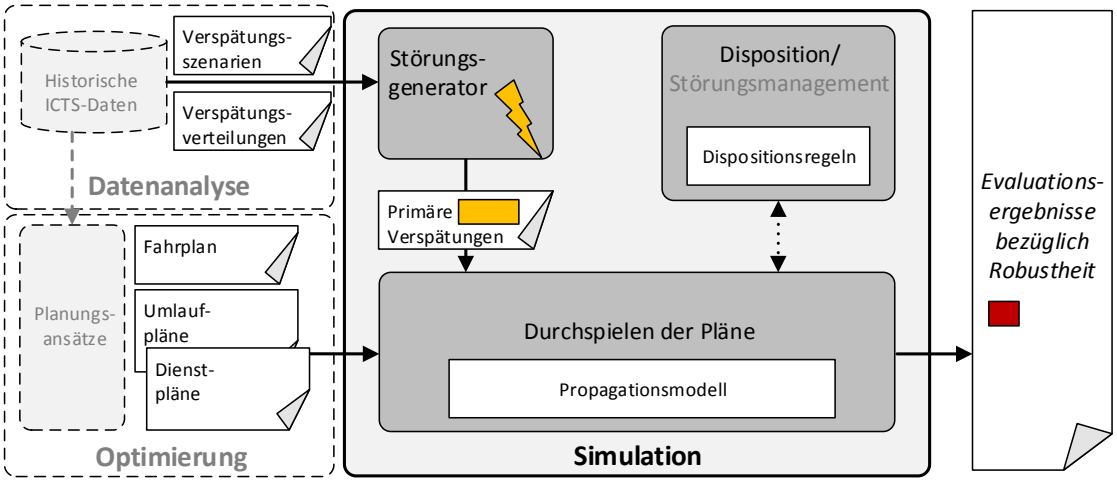

Abbildung 5.3: Simulationsframework zur Evaluation der Robustheit des geplanten Ressourceneinsatzes

Das Framework enthält drei Kernelemente: einen Störungsgenerator, das zugrunde liegende Propagationsmodell und Dispositionsregeln. Basierend auf stochastischen Verteilungen über mögliche Eintrittswahrscheinlichkeiten und Dauern primärer Verspätungen werden mit dem *Störungsgenerator* durch Störungen induzierte primäre Verspätungen erstellt. Der Umgang mit diesen Verspätungen wird mit Hilfe des *Propagationsmodells* und der *Dispositionsregeln* abgebildet. So können die Pläne evaluiert und die in Abschnitt 2.5 diskutierten Aspekte *Stabilität* und *Flexibilität* robuster Pläne für die Ausführung untersucht werden.

Propagationsmodell

Um den unverzerrten Umgang mit Verspätungen und somit die *Stabilität* der Umlauf- und Dienstpläne während ihrer Ausführung zu untersuchen, wird in diesem Abschnitt ein Modell vorgeschlagen, das die reine Verspätungspropagation über sowohl Umlauf- als auch Dienstebene gemeinsam abbildet. Eine solche konkrete mehrschichtige Betrachtung von Verspätungspropagation wird in der bisherigen Forschung zum Ressourceneinsatz im ÖPNV nicht abgebildet (vgl. Kapitel 3).

Das Modell wird in Anlehnung an das von [Dueck et al., 2012] vorgeschlagene Propagationsmodell für mehrere Ressourcen im Flugverkehr formuliert. Ein Unterschied besteht darin, dass der Zeitpunkt als Parameter für die Dauer einer primären Verspätung Einfluss finden kann. Bei [Dueck et al., 2012] werden mögliche primäre

Verspätungen als Startverspätungen aller einzelnen Flüge bereits im Vorfeld jedes Simulationslaufes festgelegt. In der vorliegenden Arbeit werden dahingegen Verspätungen erst mit Ausführung des zugeordneten Ereignisses (z.B. Abfahrt) generiert. So wird berücksichtigt, dass sich aufgrund vorheriger Verspätungen die Ausführung einer Fahrt so verzögern kann, dass sie z.B. nun (zum Teil) zu Rush-Hour Zeiten durchzuführen ist.

Sei T die Menge der betrachteten Fahrten (inklusive Servicefahrten). Für jede Fahrt $t \in T$ werden mit PDT_t die geplante Abfahrts- bzw. Startzeit und mit PAT_t die geplante Ankunfts- bzw. Endzeit bezeichnet. Die tatsächlichen Zeiten werden durch die Abfahrtsereignisse d_t und Ankunftsereignisse a_t repräsentiert. Dabei ist die Ankunftszeit für eine Fahrt t abhängig von ihrer tatsächlichen Abfahrtszeit sowie ihrer Dauer t_t und einer möglichen primären Verspätung PD_t. Die Abfahrtszeit einer Fahrt t ist abhängig von den zuvor durchgeführten Aufgaben. Diese können sich in Abhängigkeit des betrachteten Umlaufes und Dienstes unterscheiden, die jeweils Fahrt t umfassen. Mit $v(t)$ wird die zuvor durchgeführte Fahrt des Umlaufs angegeben, der auch Fahrt t ausführt. Analog wird mit $c(t)$ die zuvor durchgeführte Fahrt des Dienstes angegeben, der auch Fahrt t ausführt. Zusätzlich gibt $b_{c(t),t}$ die Pausenzeit bzw. Lenkzeitunterbrechung zwischen der Fahrt t und seiner Vorgängerfahrt auf Dienstebene an, die eingehalten werden muss, damit der Dienst Zulässigkeitskriterien erfüllen kann.

Das Modell zum Abbilden von Verspätungspropagation unter Einbeziehung der Umlauf- und Dienstebene (P-VC) kann dann wie folgt formuliert werden:

$$a_t = max\left\{PAT_t, d_t + t_t\right\} + PD_t \qquad \forall t \in T \qquad (5.4)$$

$$d_t = max\left\{PDT_t, max\left\{a_{c(t)} + \overset{a_{v(t)}}{b_{c(t),t}}\right\}\right\} \qquad \forall t \in T \qquad (5.5)$$

$$SD_t = d_t - PDT_t \qquad \forall t \in T \qquad (5.6)$$

Gleichung (5.4) stellt dabei die Ankunftsereignisse, Gleichung (5.5) die Abfahrtsereignisse dar. Gleichung (5.6) gibt für jede Fahrt $t \in T$ die zu ihrem Beginn aufgrund von Propagationseffekten verursachte sekundäre Verspätung SD_t wieder. Die stochastische Variable PD repräsentiert das Auftreten einer primären Verspätung. Die Eintrittswahrscheinlichkeit und Dauer der primären Verspätung folgt dabei der gewählten Verspätungsverteilung.

Wird bei der Simulation von Umlauf- und Dienstplänen allein das Propagationsmodell (5.4)-(5.6) eingesetzt, werden die Auswirkungen von Verspätungen während der tatsächlichen Ausführung der Pläne überschätzt und die Robustheit der Pläne somit unterschätzt. Wie im Abschnitt 2.4 beschrieben, bestehen für ÖPNV-Unternehmen

vielfältige Eingriffsmöglichkeiten während der Planausführung, um durch Änderungen der Pläne die Auswirkungen von Verspätungen einzugrenzen. Um diese Möglichkeiten in der Simulation mit einzubeziehen, werden im folgenden Abschnitt zusätzlich *Dispositionsregeln* zum Aufgabentausch von Ressourcen formuliert. Der alleinige Einsatz des Propagationsmodells dahingegen erlaubt es, eine obere Schranke für mögliche Propagationseffekte im Netzwerk zu bestimmen.

Dispositionsregeln

Neben der Entscheidung, die Propagation von Verspätungen ohne Eingriff zuzulassen, werden im entwickelten Simulationsframework weitere Reaktionsmaßnahmen auf Störungen abgebildet. Diese werden als einfache Dispositionsregeln formuliert, die einen Tausch von Aktivitäten zwischen einzelnen Ressourcen prüfen, um in den Plänen vorhandene Pufferzeit zur Absorption von Verspätungen umzuverteilen. Mit Hilfe dieser Regeln kann eine potenzielle *Flexibilität* der erstellten Umlauf- und Dienstpläne während ihrer Ausführung evaluiert werden. Eine diesbezügliche Untersuchung ist in der aktuellen Forschung zur Ressourceneinsatzplanung im ÖPNV bisher nicht vorhanden (vgl. Abschnitt 3.3). Mehr noch, in Kombination mit Ergebnissen zur Bewertung reiner Verspätungspropagation kann so das Zusammenspiel der Aspekte Stabilität und Flexiblität des Ressourceneinsatzes untersucht werden (vgl. Abschnitt 2.5).

Die umgesetzten Dispositionsregeln zielen darauf ab, die Aufgaben von Ressourcen kostenneutral zu tauschen. Voraussetzung für einen Tausch ist das Vorliegen eines Tauschpunktes (siehe Abschnitt 5.2.3). Im Allgemeinen wird in die Ausführung eingegriffen sobald abzusehen ist, dass eine Fahrt von einer Verspätungspropagation größer gleich einer vorgegebenen Toleranzgrenze betroffen sein wird. Ein Tausch zwischen Aktivitäten auf der Umlauf- und/oder Dienstebene findet allerdings nur dann statt, wenn durch ihn die betroffenen Ressourceneinsätze weiterhin zulässig bleiben ohne einen Teil des Einsatzes neuplanen zu müssen. Ansonsten wird weiterhin eine Propagation von Verspätungen zugelassen. Im Rahmen der Simulation findet keine Absage von Fahrten oder der Einsatz von Ersatzfahrer bzw. -fahrzeugen oder eine Neuplanung während der Ausführung statt, damit die Ergebnisse innerhalb dieser Arbeit im Hinblick auf die gemessene Verspätungspropagation und den ausgeführten Ressourceneinsatz vergleichbar bleiben.

Die beiden Abbildungen 5.4 und 5.5 zeigen schematisch vier grundlegende Regeln zum Tausch ohne zusätzlichen Kosteneinsatz. Konkret zeigt Abbildung 5.4 zwei Regeln, die den Tausch von Aktivitäten auf Umlaufebene betreffen. Dabei werden die Zuordnungen von Diensten zu Umläufen getauscht. Die auf Dienstebene geplanten

Aufgabenfolgen bleiben bestehen. In Abbildung 5.5 werden zwei Regeln skizziert, die den Tausch von Aktivitäten sowohl auf Umlauf- als auch Dienstebene betreffen. In den Darstellungen werden nur die Servicefahrten abgebildet, alle anderen Aktivitäten wie Warten oder Auf- und Abrüsten werden durch schwarze Linien repräsentiert. In den Abbildungen sind diejenigen Servicefahrten grau hinterlegt, die von einem Dienst ausgeführten werden, der daran beteiligt ist, eine Verspätung zu erzeugen. Rote Striche kennzeichnen eine potenzielle Weitergabe von Verspätungen. Grüne Pfeile kennzeichnen dahingegen erfolgreiche Tauschoperation zur Reduktion der Verspätungsweitergabe.

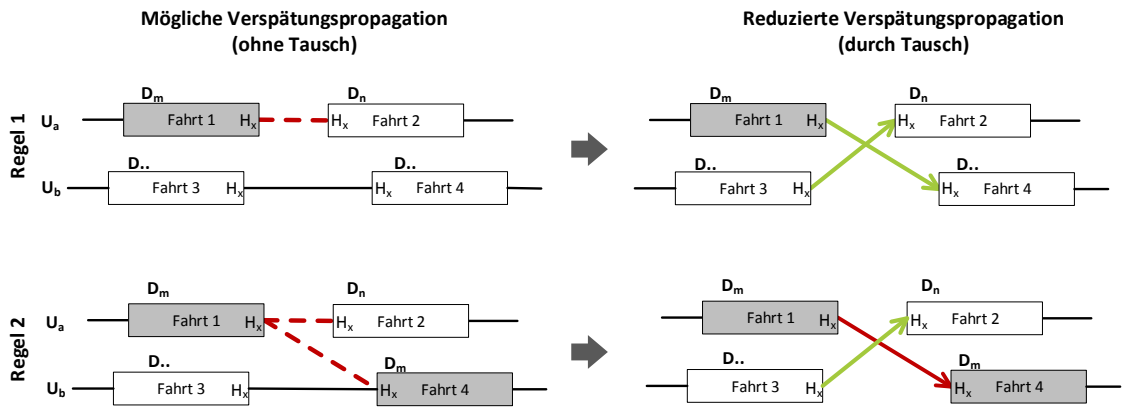

Abbildung 5.4: Dispositionsregeln zum Fahrzeugtausch

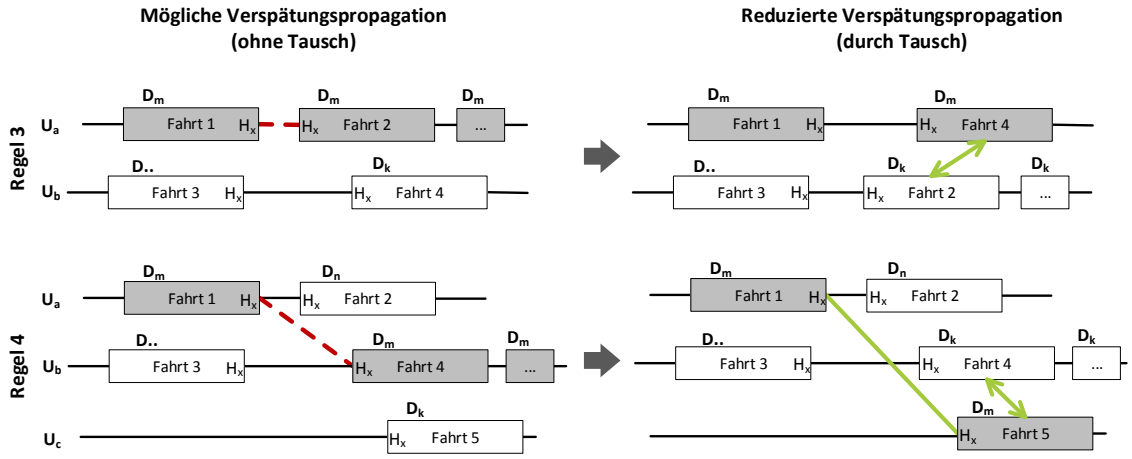

Abbildung 5.5: Dispositionsregeln für Dienständerungen

Die abgebildeten Regeln werden im Folgenden genauer beschrieben. Dabei ist angegeben, über welche Ressourcen Verspätungen ohne einen Eingriff weitergegeben würden. Zusätzlich wird jeweils die Eintrittsbedingung zur Prüfung eines Tausches

formuliert. Ergänzend zur im vorherigen Abschnitt definierten Ankunftszeit a_t einer Fahrt t (siehe Gleichung 5.4), sowie der geplanten Abfahrtszeit PDT, bezeichne dabei $s_c(t)$ die nach einer Fahrt t folgende Servicefahrt, die von dem gleichen Dienst ausgeführt wird, der auch die Fahrt t bedient. Analog bezeichne $s_v(t)$ die zu t nachfolgende Servicefahrt auf Umlaufebene. Mit der *Verspätungspropagationstoleranz* θ wird der Toleranzwert für Verspätungspropagation in Minuten bzw. Sekunden gekennzeichnet, bevor in die Ausführung eingegriffen wird.

Dispositionsregel 1 (von Umlauf auf Dienst): Um bei einem Dienstwechsel die Weitergabe von Verspätungen innerhalb eines Umlaufes auf einen anderen Dienst zu vermeiden, werden die Aktivitäten zweier Umläufe miteinander getauscht. Voraussetzungen sind, dass unmittelbar durch den Tausch keine neuen Verspätungen entstehen und die bestehenden Verspätungen abgebaut werden können. Ein solcher Tausch wird insbesondere auch dann durchgeführt, wenn für einen Umlauf keine weiteren Aufgaben mehr eingeplant sind.

Anwendung bei $a_t \geq PDT_{s_v(t)} + \theta$ *und* $s_c(t) \neq s_v(t)$

Dispositionsregel 2 (von Umlauf auf Dienst und von Dienst auf Umlauf): Um bei einem Dienstwechsel die Weitergabe von Verspätungen innerhalb eines Umlaufes auf einen Dienst und durch einen Dienst auf einen anderen Umlauf (z.B. aufgrund einzuhaltender Pausenzeit) zu reduzieren, werden die Aktivitäten zweier Umläufe miteinander getauscht. Voraussetzungen sind, dass unmittelbar durch den Tausch keine zusätzlichen Verspätungen entstehen und die Propagation zumindest in einer Ebene abgebaut werden kann.

Anwendung bei $a_t \geq PDT_{s_v(t)} + \theta$ *und* $a_t + b_{t,s_c(t)} \geq PDT_{s_c(t)} + \theta$ *und* $s_c(t) \neq s_v(t)$

Dispositionsregel 3 (innerhalb Umlauf und Dienst): Um die Weitergabe von Verspätungen innerhalb eines Umlaufes und innerhalb eines Dienstes zu vermeiden, werden die Aktivitäten von Diensten zweier Umläufe getauscht. Voraussetzungen sind, dass unmittelbar durch den Tausch keine neuen Verspätungen entstehen und die bestehenden Verspätungen abgebaut werden können. Darüber hinaus müssen die vom Aufgabentausch betroffenen Dienste weiterhin nach den Dienstregelungen zulässig bleiben.

Anwendung bei $a_t + b_{t,s_c(t)} \geq PDT_{s_c(t)} + \theta$ *und* $s_c(t) = s_v(t)$

Dispositionsregel 4 (von Dienst auf Umlauf): Um bei einem Dienstwechsel die Weitergabe von Verspätungen durch den Dienst auf einen anderen Umlauf (z.B. aufgrund einzuhaltender Pausenzeit) zu vermeiden, werden die Aktivitäten von

Diensten zweier Umläufe getauscht. Voraussetzungen sind, dass unmittelbar durch den Tausch keine neuen Verspätungen entstehen und die bestehenden Verspätungen abgebaut werden können. Darüber hinaus müssen die vom Aufgabentausch betroffenen Dienste weiterhin nach den Dienstregelungen zulässig bleiben.

Anwendung bei $a_t + b_{t,s_c(t)} \geq PDT_{s_c(t)} + \theta$ *und* $s_c(t) \neq s_v(t)$

Durch Justierung der Verspätungspropagationstoleranz θ kann die Häufigkeit der Tauschoperationen in der Simulation gesteuert werden. Ein sehr hoher Wert entspricht dem alleinigen Einsatz des Propagationsmodells (5.4) - (5.6). Ein kleiner Wert von wenigen Minuten bzw. Sekunden erlaubt im Rahmen dieser Arbeit das Berechnen einer unteren Schranke der propagierten Verspätungen während des Ressourceneinsatzes. Dabei ist anzumerken, dass aus Toleranzwerten unterhalb weniger Minuten bzw. von wenigen Sekunden möglicherweise resultierende, hochfrequente Tauschoperationen nicht zwangsläufig in der Realität umsetzbar sind.

5.3.2 Agentenbasierte Simulation

Eine agentenbasierte Simulation bildet das Verhalten von Individuen (Agenten) ab, die innerhalb eines möglichen Handlungspielraumes eigenständige Entscheidungen treffen und autonom agieren können. Durch die Interaktion der Agenten untereinander sowie mit ihrer Umwelt können Eigenschaften des abgebildeten Systems untersucht werden. Eine kurze Einführung in agentenbasierte Simulation bietet [Macal und North, 2010]. Für eine generelle Diskussion über den Einsatz und die Möglichkeiten agentenbasierter Simulation im Rahmen des Operations Research sei auf [Siebers et al., 2010] verwiesen.

Im Rahmen dieser Arbeit wurde der Einsatz des agentenbasierten Open-Source-Simulationsframework MATSim (Multi-Agent-Transport-Simulation) am konkreten Fall der Simulation des ÖPNV einer Großstadt in Süddeutschland geprüft. Für weiterführende Informationen über MATSim sei auf [Horni et al., 2016] verwiesen. MATSim umfasst neben der Simulation des privaten Individualverkehrs durch die Implementierung eines Zusatzmoduls auch die Simulation des öffentlichen Personenverkehrs. So ist es möglich, den Einfluss von Fahrgästen und Individualverkehr auf den geplanten Ressourceneinsatz zu analysieren und Wechselwirkungen über das gesamte System zu beobachten. Primäre Verspätungen entstehen dabei insbesondere durch Ein- und Aussteigevorgänge der Fahrgäste oder durch Stau im Zusammenspiel von ÖPNV und Individualverkehr.

Für den Einsatz werden umfangreiche Daten zur Eingabe benötigt. Als Grundlage ist ein Transport- bzw. Straßennetzes als gerichtetes Netzwerk abzubilden. Für die Kanten sind Informationen zur Länge, mögliche Kapazitäten sowie Geschwindigkeiten und die Anzahl befahrbarer Spuren zu hinterlegen. Zusätzlich sind die Haltestellen und Routen der betrachteten Buslinien auf das Netzwerk zu projizieren. Als wesentlicher Aspekt der Simulation sind darüber hinaus die Agenten zu modellieren. Dabei sind für die abzubildende Bevölkerung je Individuum ein Wohnort, Aktivitäten und eine mögliche Verkehrsmittelwahl zwischen den Aktivitäten festzulegen. Um das Reiseverhalten der Individuen sowie die Gesamtheit der Bevölkerung realitätsnah abzubilden, bedarf es detaillierter Daten auch über mögliche Passagierströme im ÖPNV. Im Hinblick auf die Zielsetzung dieser Arbeit wird daher für die durchzuführenden Untersuchungen allein die passgenau zur Evaluation des Ressourceneinsatzes entwickelte ereignisorientierte Monte-Carlo basierte Simulation verwendet.

5.4 Zusammenfassung und Implikationen

In diesem Kapitel wurden neben Kennzahlen zur Kosteneffizienz des Ressourceneinsatzes das Ausmaß von Verspätungspropagation als eine wesentliche, direkte Zielgröße der Lösungsansätze zum Erreichen Robuster Effizienz definiert. Darüber hinaus wurde mit einer diskreten, ereignisorientieren Monte-Carlo basierten Simulation ein Simulationsframework zur genauen Messung von Verspätungspropagation entwickelt. Ergänzend wurde der Einsatz einer agentenbasierten Simulation geprüft. Das entwickelte Monte-Carlo basierte Simulationsframework bildet sowohl eine reine Bewertung der erwarteten Verspätungspropagation zur Prüfung der Stabilität des Ressourceneinsatzes ab, als auch eine Bewertung beim zusätzlichen Einbeziehen kostenneutraler dispositiver Maßnahmen während der Planausführung zur Prüfung der Flexibilität des Ressourceneinsatzes. So lassen sich durch Variation des Simulationsumfanges im Rahmen dieser Arbeit obere und untere Schranken möglicher Verspätungspropagation untersuchen. Diese bestimmen wiederum die untere bzw. obere Schranke der erreichbaren Robustheit des jeweils geplanten Ressourceneinsatzes.

Die Messung der Propagation von Verspätungen ermöglicht es, die gesamte planimmanente Robustheit zu quantifizieren und dabei den Einfluss einzelner Planausprägungen zu abstrahieren. Dies schließt sowohl Robustheitsaspekte für Stabilität als auch für Flexibilität ein. Mit dem Abbilden von Verspätungspropagation in Lösungsansätzen kann darüber hinaus ein einzelnen Planausprägungen übergeordnetes Robustheitsmaß eingesetzt werden, das bei der Optimierung von Umlauf- und Dienstplänen keine direkten Vorgaben für den konkreten Aufbau des Ressourceneinsatzes festlegt.

Im Hinblick auf die Beantwortung der Forschungsfrage F3 über Planungsempfehlungen bietet es sich an, die Struktur der erstellten Ressourceneinsatzpläne nachträglich zu analysieren. So kann untersucht werden, wie insbesondere mehrschichtiger Ressourceneinsatz gestaltet werden sollte, um eine bestmögliche Erfüllung der Minimierung von Verspätungspropagation aufzuweisen und im Idealfall dennoch kosteneffizient zu sein.

Mit der vorgeschlagenen Propagationsformel kann die potenzielle Verspätungspropagation der geplanten Ressourceneinsätze als Zielkriterium in Lösungsansätze zur Bildung von Umläufen und Diensten berücksichtigt werden. Um dem Zusammenspiel der Zielgrößen Robuster Effizienz gerecht werden zu können, werden dabei Modellformulierungen bzw. Lösungsverfahren benötigt, die mit mehreren Zieldimensionen umgehen können. Für das Berücksichtigen potenzieller Verspätungspropagation sollten die zu entwickelnden Lösungsverfahren den Umgang mit historischen Verspätungsinformationen ermöglichen bzw. Informationen über primäre Verspätungen einbeziehen können. Das anschließende Kapitel 6 gibt einen Überblick über die Basiskomponenten der in dieser Arbeit für Varianten der Umlauf- und Dienstplanung vorgeschlagenen Lösungsansätze, die eine Kombination der Zielsetzungen zur Minimierung geplanter Kosten und insbesondere auch zur Minimierung von Verspätungsspropagation des Ressourceneinsatzes ermöglichen.

Kapitel 6

Übersicht und Zusammenhang der Lösungskomponenten

Dieses Kapitel gibt einen Überblick über die in dieser Arbeit genutzten und vorgeschlagenen Optimierungskomponenten zur Berücksichtigung von sowohl Kosteneffizienz als auch Robustheit – speziell Verspätungspropagation – bei der Planerstellung.

Als Modellierungsansätze werden netzwerkflussbasierte Formulierungen mit pfadbasierten Formulierungen kombiniert. Die pfadbasierten Formulierungen erlauben es, eine mögliche Propagation von Verspätungen für sowohl Umläufe als auch Dienste korrekt abzubilden. Durch die netwerkflussbasierten Formulierungen und das zugrunde liegende Netzwerkmodell für die Umlaufplanung profitieren die Lösungsansätze darüber hinaus. Als Lösungen können Bündel gleichermaßen kosteneffizienter Umlaufpläne betrachtet werden, die implizit durch den Lösungsfluss vorgegeben sind. In Abhängigkeit der Lösungsmethode kann der Lösungsfluss unter gleichzeitiger Berücksichtigung möglicher Robustheitsaspekte – speziell Verspätungspropagation – und/oder zulässiger Dienste in Umläufe dekomponiert werden. So kann derjenige Umlaufplan ausgesucht werden, welcher (im Zusammenspiel mit dem Dienstplan) die höchste Robustheit besitzt. Um Informationen über initiale, primäre Verspätungen einzubeziehen, können Verspätungsszenarien für jede Servicefahrt bzw. Aufgabe während der Planung berücksichtigt werden. Diese Szenarien sind aus Echtdaten oder von typischen, künstlichen Verspätungsverteilungen hergeleitet.

Aufbau des Kapitels

Abschnitt 6.1 beschreibt kurz das den Lösungsansätzen zugrunde liegende Netzwerkmodell. Abschnitt 6.2 bietet einen Überblick über die in den vorgeschlagenen Lösungsansätzen genutzten Optimierungsmodelle. Abschnitt 6.3 beschreibt wie diese Optimierungsmodelle zu Basislösungsschemata für robuste und kosteneffiziente Umlauf- und Dienstplanung kombiniert werden.

6.1 Zugrunde liegende Time-Space-Netzwerkstruktur

Aufgrund der aufgezeigten Vorteile (vgl. auch Abschnitte 3.2.1 bzw. 3.2.2) wird in dieser Arbeit zur grundlegenden Modellierung des Umlaufplanungsproblems eine Time-Space-Netzwerkformulierung als kostenminimales Mehrgüter-Zirkulationsflussproblem nach [Kliewer et al., 2006] gewählt. Dabei wird jede Depot-Fahrzeugtypkombination als eigene Netzwerkschicht betrachtet, in der mögliche Fahrzeugaktivitäten als Kanten modelliert sind. So werden Servicefahrten, Ausrückfahrten, Einrückfahrten, Verbindungsfahrten, sowie Warten im und außerhalb eines Depot explizit als Kanten abgebildet. Die Kanten verbinden Time-Space Knoten, die wiederum mögliche Abfahrten und Ankünfte an einer Haltestelle bzw. einem Depot repräsentieren. Dabei verbinden *Servicefahrtkanten* Abfahrten an einer Haltestelle mit Ankünften an einer anderen Haltestelle. Fahrten aus bzw. in ein Depot sind über *Ausrückfahrtkanten* bzw. *Einrückfahrtkanten* modelliert. *Leerfahrtkanten* verbinden (aggregierte) Gruppen kompatibler, d.h. nacheinander ausführbarer, Servicefahrten miteinander, während *Wartekanten* Ankunfts- und Abfahrtsereignisse an einem Ort miteinander verknüpfen. In dem Time-Space-Netzwerk ist eine Leerfahrtkante für die Verbindung zweier kompatibler Servicefahrten redundant, wenn die gleiche Verbindung durch die Verwendung einer anderen Leerfahrtkante und Wartekanten erreicht werden kann. Bei der Modellierung werden nur nicht redundante Leerfahrtkanten benutzt, die Gruppen von Ankunftsereignissen an einer Haltestelle mit Gruppen von Abfahrtsereignissen einer anderen Haltestelle verbinden. Zusätzlich verbindet eine *Zirkulationsflusskante* das letzte Ankunftsereignis mit dem ersten Abfahrtsereignis am Depot und stellt so durch Flusserhaltung einen gültigen Lösungsfluss sicher, da alle Knotenbedarfe mit null definiert sind. Für ein Beispiel einer solchen Netzwerkstruktur – im Kontext der Berücksichtigung von Verspätungspropagation – sei auf Abbildung 7.3 im späteren Abschnitt 7.3 verwiesen.

6.2 Basisoptimierungsmodelle zum Einbeziehen von Verspätungspropagation

Um Umlauf- und Dienstplanungsprobleme in sequentieller (SEQ), teilintegrierter (PINT) oder integrierter (INT) Art und Weise zu lösen, werden im Rahmen dieser Arbeit verschiedene Kombinationen von Basismodellen angewendet. Diese Basismodelle erlauben es, kosteneffiziente und/oder robuste Ressourceneinsatzpläne zu erstellen. Dabei wird im Wesentlichen in Abhängigkeit des Einsatzzwecks zwischen drei Arten von Modellen unterschieden.

Die erste Art Basismodell stellt sicher, dass die betrachteten Fahrten/Aufgaben so zu Umläufen und Fahrerdiensten zugewiesen werden, dass die geplanten Kosten und/oder eine mögliche Verspätungspropagation minimiert wird. Diese Modelle werden als *Mehrgüter-Minimalkosten-Fluss-Probleme* (engl. *multi-commodity-minimal-cost-flow-problem, MCFP*) und als *Set-Partitioning-Probleme* (*SPP*) formuliert.

Die zweite Art Modell wird dafür genutzt, Flusslösungen zu dekomponieren, welche basierend auf Modellen des ersten Typs berechnet wurden. Im Rahmen dieser Arbeit wird dabei der Lösungsfluss nicht einzeln an jedem Time-Space Knoten zu Fahrten- bzw. Aufgabensequenzen verknüpft, sondern der gesamte Lösungsfluss wird in die Dekomposition einbezogen. [Kliewer, 2005] bezeichnet eine solche Vorgehensweise als *globale Flussdekomposition* (engl. *global flow decomposition, GFD*) im Gegensatz zu lokalen Flussdekompositionsschemata an einzelnen Knoten.

Die dritte Art Modell erlaubt es, mögliche Verspätungspropagation beim Erstellen von zulässigen (und kosteneffizienten) Teilumläufen, Umläufen sowie Fahrerdiensten zu berücksichtigen. Mit Hilfe dieses Modelltyps werden die *Column Generation Pricing-Probleme* (*CG-P*) formuliert, welche in den vorgeschlagenen Lösungsansätzen genutzt werden. Dabei werden die Pricing-Probleme wie in der Literatur etabliert als *ressourcenbeschränkte Kürzeste-Wege Probleme* (engl. *resource constrained shortest path problems, RCSP*) (siehe z.B. [Irnich und Desaulniers, 2004]) formuliert.

Tabelle 6.1: Basisoptimierungsmodelle zur Minimierung von Kosten (C) und/oder Verspätungspropagation (P)

Typ	Problem	Ziel	Modell	Planungselemente	Lösungsprozedur	
1	VSP	C	TSN basiertes MCFP	Fahrten	LP/MIP-Solver und Flussdekomposition	§6.1
	R-VSP	C und P	TSN bas. MCFP verknüpft mit pfadbas. SPP	Fahrten, Fahrzeugumläufe	Column Generation mit Lagrange-Relaxation und Flussdekomposition	§7.4.1, §8.2
	(R-)CSP	C (und P)	Pfadbas. SPP	Aufgaben, Dienste	Column Generation	§7.5.1, §8.2
2	GFD (R-)VSP I	P	Pfadbas. SPP	Fahrten, Umlaufblöcke	Column Generation	§7.3.2
	GFD (R-)VSP II	P	Pfadbas. SPP	Umlaufblöcke, Fahrzeugumläufe		§7.3.2
3	(R-)VSP-CG-P I	C (und P)	TSN bas. RCSP	Fahrten	Label Setting Algorithmus auf Basis	§7.4.2, §8.3.3
	(R-)VSP-CG-P II	C (und P)	TSN bas. RCSP	Umlaufblöcke /Dienststücke	Dynamischer Programmierung (mit Berücksichtigung von Verspätungspropagation	§7.4.2, §8.3.3
	(R-)CSP-CG-P	C (und P)	TSN bas. RCSP	Aufgaben		§7.5.2, §8.3.3

Die Tabelle 6.1 fasst die genutzten Basismodelle zusammen bezüglich ihrer Zielfunktion(en), der zugrunde liegenden mathematischen Formulierung, den jeweils als Basis zur Modellierung verwendeten Planungselementen, sowie einer Kurzbeschrei-

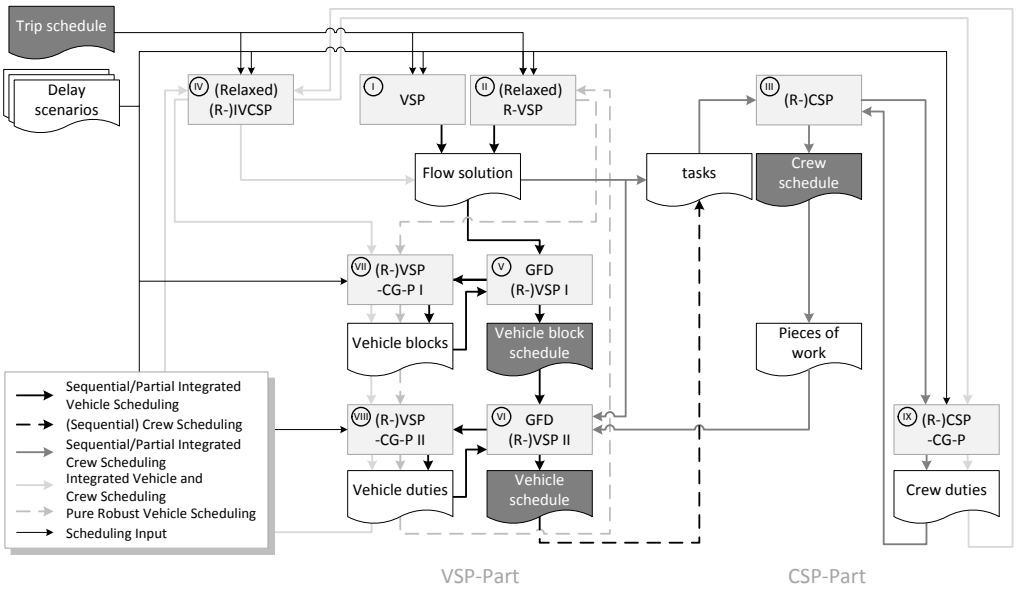

Abbildung 6.1: Überblick über Basislösungsschemata zur Berücksichtigung von Kosteneffizienz und Verspätungspropagation

bung der zugehörigen Lösungsprozedur und einen Hinweis auf die jeweiligen Kapitelabschnitte, in denen der Einsatz der Modelle genauer beschrieben wird.

6.3 Basislösungschemata zur Erstellung robuster und kosteneffizienter Pläne

Die in Abschnitt 6.2 vorgestellten Basismodelle bzw. Komponenten werden in mehreren Kombinationen genutzt, um Lösungsschemata für unterschiedlichste Planungsvarianten und Zielsetzungen etablieren zu können. Abbildung 6.1 gibt einen Überblick über die Interaktion zwischen den Komponenten. Es wird unterschieden zwischen Schemata für kosteneffiziente und/oder robuste Umlaufplanung, für sequenzielle und teilintegrierte Umlaufplanung, sequenzielle Dienstplanung, teilintegrierte Dienstplanung sowie für integrierte Umlauf- und Dienstplanung.

6.3.1 Schema zur Umlaufplanung

Zur Planung von Umläufen werden zwei Arten von Umlaufplanungsproblemen einbezogen: Ein klassisches Umlaufplanungsproblem zur Minimierung von geplanten Kosten (I), und ein Umlaufplanungsproblem, welches zusätzlich die potenzielle Verspätungspropagation minimiert (II). Letzteres wird mit Column Generation in Kom-

bination mit Lagrange-Relaxation gelöst. Dabei werden in den Pricing-Problemen in zwei Schritten zunächst (verspätungstolerante) Teilumläufe bzw. Umlaufblöcke (VII) erstellt und darauf basierend gesamte (Fahrzeug-)Umläufe (VIII). Zur Erstellung eines Umlaufplans muss für beide Planungsprobleme die Flusslösung eines zugehörigen MCFP dekomponiert werden. In dieser Arbeit wird ein globales Flussdekompositionsschema vorgeschlagen, um Verspätungspropagation zu berücksichtigen, wenn entsprechende Teilumläufe bzw. Umlaufblöcke (V) sowie der finale Umlaufplan (VI) erstellt werden. Während der globalen Flussdekomposition werden mögliche Teilumläufe und Umläufe mit einem Column Generation Ansatz ermittelt. Die Pricing-Probleme (VII) und (VIII) werden daher ebenfalls während der Flussdekomposition aufgestellt und gelöst. Details zur Vorgehensweise bei kosteneffizienter und/oder robuster Umlaufplanung werden im Zusammenhang mit der Dienstplanung in Kapitel 7 beschrieben, in welchem insbesondere Abschnitte 7.3 und 7.4 einen Einblick in die globale Flussdekomposition und die Erstellung robuster Netzwerkflüsse geben.

6.3.2 Schema zur sequenziellen Planung

Für die sequenzielle Planung werden basierend auf einem rein kosteneffizienten, robusten, oder kosteneffizient und robusten Umlaufplan, aus den eingeplanten Fahrzeugaktivitäen Aufgaben abgeleitet, die von Fahrerdiensten abgedeckt werden müssen. Das zugehörige Dienstplanungsproblem wird als SPP (III) formuliert und mit einem Column Generation Ansatz gelöst, wobei – je nach Zielsetzung – eine Berücksichtigung von Verspätungspropagation während der Diensterstellung in den Column Generation Pricing-Problemen abgebildet wird oder nicht (IX).

6.3.3 Schema zur teilintegrierten Planung

Die in der Dienstplanung von Fahrerdiensten abzudeckenden Aufgaben werden bei der teilintegrierten Planung ähnlich der sequenziellen Planung von kosteneffizienten, robusten oder kosteneffizienten und robusten Umlaufplanungslösungen abgeleitet. Der Unterschied zur sequenziellen Vorgehensweise besteht darin, dass diese Aufgaben nicht von einem fixen Umlaufplan hergeleitet werden, sondern von einer Netzwerkflusslösung, die ein Bündel zulässiger Umlaufpläne mit dem gleichen Zielfunktionswert repräsentiert. Dies bietet zusätzliche Freiheitsgrade, die beim Lösen der zugehörigen Dienstplanungsprobleme (III) genutzt werden können, um Aufgaben zu Diensten zusammenzufassen. So gibt es mehr Möglichkeiten, zulässige Dienste in den Column Generation Pricing-Problemen (IX) zu generieren. Nach der Dienstplanerstellung werden die im Plan enthaltenen Dienststücke wiederum genutzt, um mit-

tels globaler Flussdekomposition (VI) einen zum Dienstplan kompatiblen Umlaufplan zu erstellen. Dabei kann optional die potenzielle Verspätungspropagation minimiert werden. Details zum teilintegrierten Lösungsschema zur Erstellung sowohl kosteneffizienter als auch robuster Umlauf- und Dienstpläne werden in Kapitel 7 beschrieben.

6.3.4 Schema zur integrierten Planung

Das integrierte Umlauf- und Dienstplanungsproblem wird mit Column Generation in Kombination mit Lagrange-Relaxation gelöst (IV). Während der Lösungsprozedur werden sowohl Fahrzeugumläufe als auch Fahrerdienste durch Lösung der zugehörigen Pricing-Probleme (VII), (VIII) und (IX) generiert. Zulässige ganzzahlige Lösungen werden mit Hilfe des sequenziellen Lösungsschemas (siehe Abschnitt 6.3.2) oder des teilintegrierten Lösungsschemas (siehe Abschnitt 6.3.3) erstellt. Eine detailierte Beschreibung der Modelle und Verfahren, die während des integrierten Lösungsschemas genutzt werden, findet sich in Kapitel 8 sowie in Kapitel 9 für die Erweiterung des ursprünglichen Planungsproblems um zusätzliche Freiheitsgrade durch den Einbezug von Zeitfenstern für Servicefahrten.

6.4 Zusammenfassung

Mit diesem Kapitel wurde ein einleitender Überblick für die nachfolgenden Kapitel geschaffen, in denen die verschiedenen Modellierungen und Lösungsverfahren für die unterschiedlichen in diesem Kapitel skizzierten Lösungsschemata und Zielsetzungen zur kosteneffizienten und robusten Ressourceneinsatzplanung detailliert beschrieben werden. Es wurde verdeutlicht, welche wechselseitig nutzbaren Basismodelle und Basismethoden den unterschiedlichen Planungsvarianten zugrunde liegen: Kombinationen aus netzwerkfluss- und pfadbasierten Formulierungen, Methoden zur globalen Dekomposition gesamter Netzwerkflusslösungen sowie Modelle und Methoden mit denen im Rahmen von Column Generation Pricing-Problemen ressourcenbeschränkte Kürzeste-Wege-Probleme zur Umlaufgenerierung und zur Dienstgenerierung mit Berücksichtigung von potenzieller Verspätungspropagation formuliert und gelöst werden können.

Kapitel 7

Robuste Effizienz durch teilintegrierte Planung

In diesem Kapitel wird ein Lösungsansatz vorgestellt, der im Gegensatz zu einer sequenziellen Planung von Umläufen und Diensten die gegenseitigen Abhängigkeiten zwischen Fahrzeugen und Fahrern mit mehr Freiheitsgraden berücksichtigen kann und geeignet ist, neben Kostenminimierung auch Robustheitsaspekte in die Umlauf- und Dienstplanung einzubeziehen. Dieser *teilintegrierte* Ansatz basiert auf einem Ansatz für kosteneffiziente Umlauf- und Dienstplanung, der von [Gintner, 2008] vorgeschlagen wird. Er vermeidet die Beeinträchtigung einer sequenziellen Planung, bei der die Dienstplanung ausgehend von einem bereits fixierten Umlaufplan durchgeführt wird. Die verschiedenen Schritte des (erweiterten) Ansatzes und die Interaktion zwischen Umlaufplanung und Dienstplanung sind in Abbildung 7.1 schematisch dargestellt.

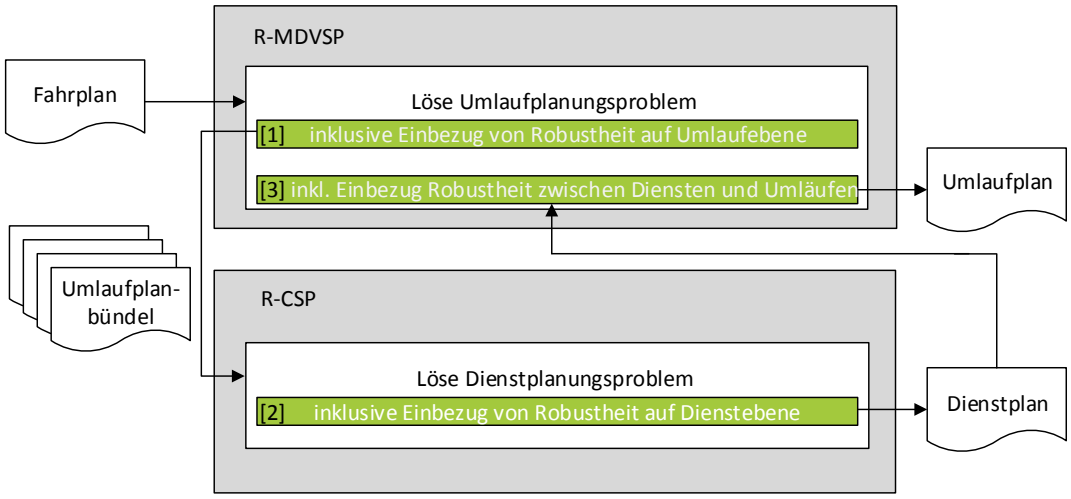

Abbildung 7.1: Schritte der teilintegrierten Umlauf- und Dienstplanung mit Einbezug von Robustheitsaspekten in die Planung

Grundidee

Ausgehend vom Lösen des Umlaufplanungsproblems erhält man aufgrund des zugrunde liegenden TSN-Netzwerkmodells nach [Kliewer et al., 2006] zunächst ein Bündel von (kostenoptimalen) Umlaufplänen, die implizit durch den Netzwerkfluss gegeben sind. Anstatt diesen Netzwerkfluss während der Umlaufplanungsphase bereits zu einem Umlaufplan zu dekomponieren, wird dieser Freiheitsgrad in die Dienstplanungsphase übernommen. Das Dienstplanungsproblem wird somit basierend auf einer Menge von (kostenoptimalen) Umlaufplänen gelöst. Basierend auf der errechneten Dienstplanlösung wird anschließend ein (kostenoptimaler) Umlaufplan erstellt. Da sich die Kosten dieses finalen Umlaufplans im Vergleich zur ursprünglichen Umlaufplanungslösung nicht weiter ändern, kann beim Erstellen des Umlaufplans neben der Kostenminimierung ein weiteres Ziel wie die Erhöhung der Robustheit im Zusammenspiel mit den eingeplanten Diensten verfolgt werden.

Das Verfahren bietet die Möglichkeit, grundlegende Untersuchungen über das Zusammenspiel der Planungsebenen Umlauf- und Dienstplanung hinsichtlich der Zielsetzungen Kosteneffizienz und Robustheit durchzuführen. Innerhalb des entwickelten Lösungsansatzes existieren dabei drei Stellschrauben um neben Kostenminimierung auch eine Erhöhung der Robustheit zu verfolgen. Diese Stellschrauben beziehen sich im Einzelnen auf die Umlaufebene (siehe Schritt 1 in Abbildung 7.1), die Dienstebene (Schritt 2), sowie die Verknüpfung der Umlauf- und Dienstebene (Schritt 3). Insbesondere durch die Planung unter Beibehaltung der Kosteneffizienz können so Aussagen zur Klärung der Forschungsfrage F2 getroffen werden, wieviel Robustheit in kosteneffizienten Plänen realisiert werden kann (vgl. Kapitel 4).

Aufbau des Kapitels

Nach der Definition des zugrunde liegenden Planungsproblems in Abschnitt 7.1 wird zunächst das vorgeschlagene Lösungsschema in Abschnitt 7.2 vertieft. Anschließend wird in Abschnitt 7.3 als Grundlage für die Entwicklung der weiteren Verfahren eine globale Netwerkflussdekompositionsstrategie für die Umlaufplanungsebene vorgestellt. Danach werden die eingangs aufgeführten Stellschrauben für die Berücksichtigung von Robustheit innerhalb bzw. zwischen der Umlauf- und Dienstplanungsebene und ihre Umsetzung im Lösungsansatz ausführlich erläutert. Die Umlaufplanungsebene wird in Abschnitt 7.4, die Dienstplanungsebene in Abschnitt 7.5 und die Rückkopplung der Dienst- zur Umlaufebene in Abschnitt 7.6 betrachtet. Abgeschlossen wird dieses Kapitel mit einer Zusammenfassung des vorgestellten Ansatzes in Abschnitt 7.7.

7.1 Problemdefinition R-MDPVCSP

Das *teilintegrierte (Mehrdepot-)Umlauf- und Dienstplanungsproblem mit der Zielsetzung nach robuster Effizienz* (engl. *robust partial integrated multiple-depot vehicle and crew scheduling problem, R-MDPVCSP*) kann wie folgt definiert werden: Zu gegebenen Servicefahrten und (betrachteten) Verspätungsszenarien werden ein kosteneffizienter und in Bezug auf die Szenarien möglichst robuster Umlaufplan und Dienstplan gesucht. Umlaufplan und Dienstplan sollen unter Berücksichtigung folgender Abhängigkeiten zueinander kompatibel und miteinander verknüpft sein:

1. Für die Menge zu bedienender Servicefahrten des Fahrplans soll ein Bündel von Umlaufplänen gefunden werden, das zulässige und kosteneffizente (sowie möglichst robuste) Umlaufpläne enthält und einen robust effizienten Dienstplan ermöglicht. Dabei sind die Menge möglicher Leerfahrten zwischen einzelnen Haltestellen sowie Depotstandorte und unterschiedliche Fahrzeugtypen zu berücksichtigen.

2. Basierend auf der aus der Menge an möglichen Umlaufplänen ableitbaren zu leistenden Aufgaben für Fahrer soll eine Menge von Diensten gefunden werden, die einen zulässigen, kosteneffizienten und möglichst robusten Dienstplan bilden. Dabei sind die Möglichkeiten zum Ablösen und Transfer von Fahrern sowie Regeln zur Dienstbildung zu berücksichtigen.

3. Für die Menge von Diensten soll derjenige Umlaufplan aus der Menge kostengleicher Umlaufpläne ausgewählt werden, der zur Dienstmenge kompatibel ist und einen möglichst robusten gesamten Ressourceneinsatz ermöglicht.

Generell soll dabei die Verspätungspropagation wie in Abschnitt 5.2.1 beschrieben als Nicht-Robustheitsmaß zur Bewertung der Ressourceneinsatzpläne herangezogen werden.

7.2 Lösungsschema für R-MDPVCSP

Abbildung 7.2 detailliert die zu Beginn dieses Kapitels beschriebenen Schritte zur Lösung des teilintegrierten Ressourceneinsatzproblems mit der Zielsetzung nach robuster Effizienz. Dabei werden die Ziele Kostenminimierung und Robustheit sowohl auf Umlauf- als auch auf Dienstebene auf verschiedenen Stufen adressiert. Im Einzelnen werden zur Lösung Teilumläufe und Umläufe, sowie Dienststücke und Dienste betrachtet.

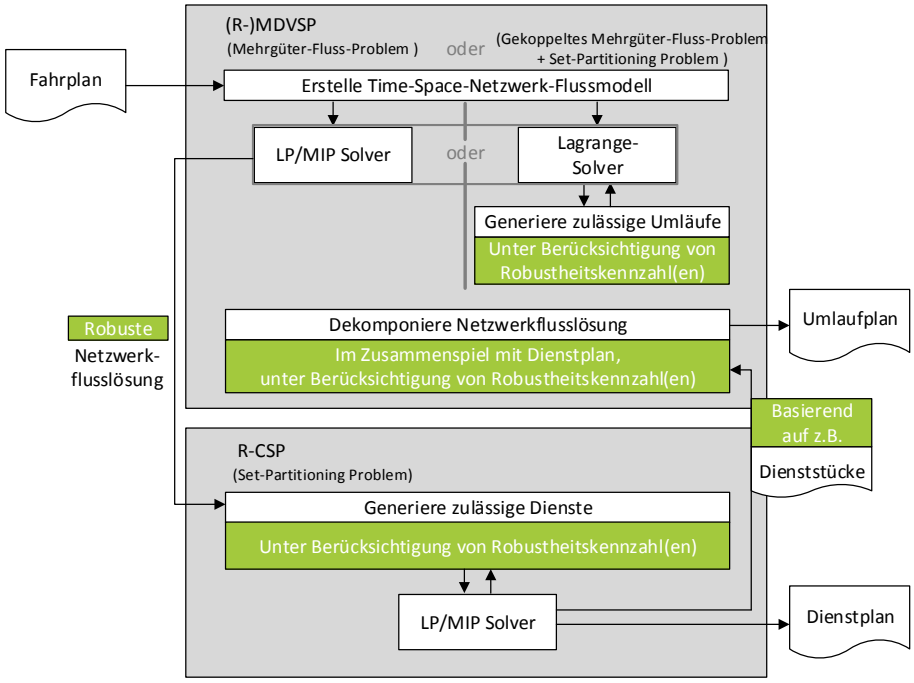

Abbildung 7.2: Teilintegrierte Umlauf- und Dienstplanung mit Berücksichtigung der Zielsetzungen Kosteneffizienz und Robustheit

Zu einem gegebenen Fahrplan eines Verkehrstages wird unter Einbeziehen aller Depot-Fahrzeugtypkombinationen ein Umlaufplanungsproblem (MDVSP) als Mehrgüter-Fluss-Problem basierend auf der TSN-Formulierung nach [Kliewer et al., 2006] modelliert. Das Flussproblem kann zum einen kostenoptimal mit einem LP/ MIP Solver gelöst werden. Die dabei resultierende Flusslösung repräsentiert dann ein Bündel kostenoptimaler Umlaufpläne und dient als Eingabe für die Dienstplanung. Als Variante wird in dieser Arbeit vorgeschlagen, die Flusslösung unter Berücksichtigung von Robustheitsaspekten mit Hilfe von Lagrange-Relaxation in Kombination mit Column Generation so zu erstellen, dass sie potenziell robustere Umlaufpläne enthalten kann als beim Lösen ohne expliziten Einbezug der Robustheitszielsetzung. Dafür wird ein Umlaufplanungsproblem (R-MDVSP) formuliert, das ein Mehrgüter-Fluss-Problem und ein Set-Partitioning-Problem miteinander koppelt. Dabei werden geringfügige Mehrkosten im Vergleich zu einer kostenoptimalen Lösung erlaubt, womit ein anderes Bündel an möglichen Umlaufplänen als Eingabe für Dienstplanung zur Verfügung stehen kann. Das Dienstplanungsproblem (R-CSP) kann dann basierend auf der Flusslösung je Depot-Fahrzeugtypkombination als Set-Partitioning-Problem formuliert werden, wobei diejenigen Aufgaben von Fahrern abzudecken sind, die durch die Flusslösung auf Umlaufplanungsebene vorgegeben sind. Das R-CSP wird mit ei-

nem Column Generation Ansatz gelöst – unter Berücksichtigung von Robustheitsaspekten in den als ressourcenbeschränkte Kürzeste-Wege-Problemen formulierten Pricing-Problemen zur Diensterstellung. Anschließend werden auf Basis der erstellten Dienste bzw. der in den Diensten enthaltenen Dienststücke Umläufe so gebildet, dass der Ressourceneinsatzplan möglichst wenig Verspätungspropagation erlaubt: Mit Hilfe der Dienststücke wird eine globale Dekomposition des Lösungsflusses für das Umlaufplanungsproblem durchgeführt.

Die einzelnen Schritte des Lösungsschemas werden in den folgenden Abschnitten näher ausgeführt. Begonnen wird im nächsten Abschnitt mit den entwickelten Modellen und Methoden zur globalen Flussdekomposition für die Umlaufebene.

7.3 Globale Flussdekomposition als Vergleichsgrundlage für Lösungen der Umlaufebene

Dieser Abschnitt entwickelt mit der globalen Dekomposition des Lösungsflusses eine grundlegende Methodik zur Berücksichtigung von Verspätungspropagation im Rahmen der TSN-basierten Ressourceneinsatzplanung. Dabei wird die Propagation von Verspätungen zunächst ohne Verknüpfung mit der Dienstplanungsebene allein auf der Umlaufebene betrachtet. Wie in Kapitel 6 in Übersichtsabbildung 6.1 gezeigt, stellen die ausgearbeiteten Modelle und Methoden wesentliche Lösungskomponenten dar, auf die im weiteren Verlauf der Arbeit in verschiedenen Lösungsschemata zurückgegriffen werden kann. Beispielsweise können die mit der globalen Netzwerkflussdekomposition erstellten Umlaufpläne auch nach dem Schema der sequenziellen Planung (vgl. Abschnitt 6.3.2) als Eingabe für eine (robuste) Dienstplanung dienen.

Bevor in den Abschnitten 7.3.1 und 7.3.2 Modellierung und Lösungsansatz vorgestellt werden, wird zunächst noch der Unterschied zwischen lokaler und globaler Netzwerkflussdekomposition genauer erläutert und ein Beispiel zur globalen Netzwerkflussdekomposition mit Einbeziehung von Robustheitsaspekten gegeben.

Lokale Netzwerkflussdekomposition und globale Netzwerkflussdekomposition

Als Lösung des auf Basis eines TSN modellierten Umlaufplanungsproblems ergibt sich ein Netzwerkfluss. Eine (kostenoptimale) Flusslösung repräsentiert dabei eine Vielzahl (kostenoptimaler) Umlaufpläne. Um einen konkreten Umlaufplan auszuwählen, d.h. um konkrete Umläufe aus der Flusslösung abzuleiten, muss die Flusslösung in Pfade aufgeteilt werden, von denen jeder einem zulässigen Umlauf entspricht. Jede in

der Flusslösung beschriebene Aufgabe bzw. Fahrt muss dabei genau einem Umlauf zugeteilt werden. [Kliewer, 2005] beschreibt verschiedene Strategien zur Flussdekomposition für TSN-Strukturen und teilt sie ein in lokale und globale Dekompositionsstrategien. *Lokale Dekompositionsstrategien* verknüpfen Flusseinheiten an einzelnen Haltestellen bzw. Time-Space-Knoten, z.B. nach dem First-In-First-Out (FIFO-) oder Last-In-First-Out (LIFO-) Prinzip. Bei einer *globalen Dekompositionsstrategie* wird die Flusslösung dekomponiert, indem gleichzeitig alle Entscheidungen zur Verknüpfung der Flusseinheiten und somit alle möglichen Umläufe bzw. Umlaufpläne hinsichtlich eines gewählten Zielkriteriums betrachtet werden. Mit Hilfe einer Formulierung der Flussdekomposition als Set-Partitioning-Problem, in welcher Umläufe einzelnen Variablen bzw. Spalten im Modell entsprechen, können dann die Umläufe ausgewählt werden, die zusammen bezüglich des gewählten Zielkriteriums die bestmögliche Lösung, d.h. die bestmögliche Dekomposition, darstellen. Eine Schwierigkeit besteht dabei darin, alle möglichen Spalten bzw. Umläufe abzubilden, da es schon bei kleinen Problemstellungen mit wenig Servicefahrten sehr viele Möglichkeiten gibt, zulässige Umläufe bzw. Pfade aus dem Netzwerkfluss zu generieren. Im Rahmen dieser Arbeit soll eine globale Netzwerkflussdekomposition entwickelt werden, mit der möglichst robuste Umläufe gefunden werden können.

Globale Netzwerkflussdekomposition mit Zielsetzung Robustheit

Um mit einer globalen Flussdekompositionsstrategie (zunächst unabhängig von der Dienstplanung) aus dem Bündel möglicher, kostengleicher Lösungen den robustesten Umlaufplan herauszufinden, muss Robustheit geeignet abgebildet werden. Basierend auf dem von [Amberg et al., 2011] bzw. [Amberg, 2014] vorgeschlagenen Ansatz zur Dekomposition eines Netzwerkflusses zur Erhöhung der Ähnlichkeit von Ressourceneinsatzplänen, wird im Folgenden ein entsprechender globaler Dekompositionsansatz entwickelt. Dabei wird eine mögliche Propagation von Verspätungen innerhalb einzelner Umläufe unter Einbeziehung des gesamten Netzwerkflusses berücksichtigt und der Umlaufplan mit der bezogen auf die betrachteten Verspätungsszenarien geringsten Verspätungspropagation erstellt.

Abbildung 7.3 veranschaulicht in einem Beispiel die Idee für eine solche globale Flussdekomposition. Dargestellt ist ein Time-Space-Netzwerk für ein Depot mit zugehörigen Kanten für vier zu bedienende Servicefahrten, Kanten für potenzielle Ausrück- und Einrückfahrten, Wartekanten, einer Leerfahrtkante und einer Zirkulationsflusskante (vgl. Abschnitt 6.1). Aufgrund der Struktur der kostenoptimalen Flusslösung sind in dem Beispiel zwei verschiedenen Umlaufpläne mit jeweils zwei Umläufen U1 und U2 möglich. Wird bei der Planerstellung berücksichtigt, dass eine

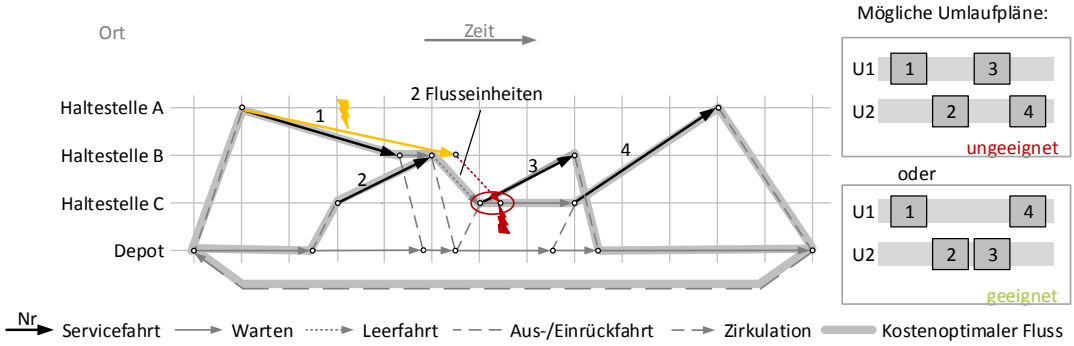

Abbildung 7.3: Flussdekomposition mit Berücksichtigung von Verspätungspropagation in einem Time-Space-Netzwerk, angelehnt an [Amberg et al., 2011]

mögliche Verspätung von Servicefahrt 1 auch Fahrt 2 beeinflussen könnte, stellen die Fahrtenfolgen 1, 4 und 2, 3 den robusteren der beiden möglichen Umlaufpläne dar, da dort die mögliche Verspätung nicht an Fahrt 2 weitergegeben wird.

7.3.1 Modellierung

Mathematisch kann das Modell zur globalen Flussdekomposition unter Berücksichtigung der Propagation von Verspätungen wie folgt beschrieben werden. Sei $\mathcal{D} = \{1, ...n\}$ die Menge aller Kombinationen aus Depot und Fahrzeugtyp. Je Netzwerkschicht $d \in \mathcal{D}$ ist ein als TSN modelliertes Umlaufplanungsnetzwerk $G^d = (N^d, A^d)$ definiert, mit der Knotenmenge N^d und der Kantenmenge A^d (siehe Abschnitt 6.1). Der Lösungfluss auf einer Kante $(i, j) \in A^d$ mit $i, j \in N^d$ sei durch b^d_{ij} beschrieben. Bei der Dekomposition des Flusses muss lediglich der Teil des Umlaufplanungsnetzwerks G^d betrachtet werden, der einen Lösungsfluss enthält, also alle Kanten $(i, j) \in A^d_B \subseteq A^d$ mit $b^d_{ij} > 0$ und alle Knoten $i \in N^d_B \subseteq N^d$ mit $\{j : (j, i) \in A^d_B\} \neq \emptyset$ bzw. $\{j : (i, j) \in A^d_B\} \neq \emptyset$.

Binäre Entscheidungsvariablen x^d_l für Fahrzeugumläufe $l \in L^d_B$ mit zugeordnetem Zielfunktionskoeffizienten w^d_l geben an, ob Umlauf l für die Lösung ausgewählt wird oder nicht. Dabei sei L^d_B die Menge aller möglichen Fahrzeugumläufe für den gegebenen Lösungsfluss einer Netzwerkschicht und der Koeffizient w^d_l die Summe der Verspätungspropagation eines Umlaufs l bezogen auf die erwarteten primären Verspätungen, berechnet nach Robustheitsmaß 5.1. Die Menge aller Umläufe, die Kante $(i, j) \in A^d_B$ enthalten, sei durch $L^d_B(i, j)$ definiert. Dann wird das *Flussdekompositionsproblem zur Erstellung von robusten, verspätungstoleranten Umläufen* für eine Netzwerkschicht $d \in \mathcal{D}$ (engl. *robust global flow decomposition for the vehicle sche-*

duling problem, GFD-R-VSP) als allgemeines Set-Partitioning-Problem wie folgt formuliert.

GFD-R-VSP:

$$\sum_{l \in L_B^d} x_l^d w_l^d \to \min \tag{7.1}$$

$$\text{s.t.} \sum_{l \in L_B^d(i,j)} x_l^d = b_{ij}^d \qquad \forall (i,j) \in A_B^d \tag{7.2}$$

$$x_l^d \in \{0,1\} \qquad \forall l \in L_B^d \tag{7.3}$$

Die Zielfunktion (7.1) minimiert die Verspätungspropagation der Umläufe. Nebenbedingungen (7.2) garantieren, dass in der dekomponierten Flusslösung jede durch eine mit Fluss belegte Kante repräsentierte Fahrt bzw. Aktivität genau so oft von den ausgewählten Umläufen bedient wird, wie benötigt, der vorhandene Fluss also vollständig abgedeckt wird. Restriktion (7.3) beschreibt den binären Charakter der Entscheidungsvariablen für Fahrzeugumläufe.

7.3.2 Lösungsverfahren

Die Anzahl der zulässigen Pfade bzw. Fahrzeugumläufe für die Formulierung (7.1)-(7.3) ist in der Regel extrem hoch, da der Arbeitstag für Fahrzeuge üblicherweise zeitlich nicht eingeschränkt ist und diese von der ersten auszuführenden Fahrt bis zur letzten Fahrt des betrachteten Planungszeitraumes eingesetzt werden können. Bereits für kleine Problemstellungen entstehen so extrem viele Kombinationsmöglichkeiten von Fahrtensequenzen zu zulässigen Umläufen. Um die betrachtete Anzahl möglicher Pfade bei der Dekomposition der Flusslösung zu reduzieren, werden zwei Techniken eingesetzt: Zum einen werden, statt explizit alle möglichen Pfade beim Lösen des Modells zu berücksichten, mit Hilfe eines Column Generation Ansatzes alle möglichen Pfade implizit in den Lösungsprozess einbezogen. Da das Lösen von Umlaufplanungsproblem mit Hilfe eines Column Generation Ansatzes aber zu einer schlechten Konvergenz des Verfahrens führen kann, wenn in den Pricing-Problemen lange Pfade gebildet werden (vgl. [Benchimol et al., 2012]), wird der Column Generation Ansatz ähnlich zu [Amberg, 2014] in zwei Phasen ausgeführt. In der ersten Phase werden Teilumläufe erstellt und daran anschließend werden diese Teilumläufe zu Tagesumläufen verknüpft (vgl. Definitionen in Abschnitt 2.2). Algorithmus 7.3.1 fasst das gesamte Vorgehen der globalen Flussdekomposition auf Umlaufebene unter Berücksichtigung der Propagation von Verspätungen zusammen. Die beiden Phasen des Algorithmus werden im Folgenden genauer beschrieben.

Algorithmus 7.3.1 : Globale Flussdekomposition zur Minimierung der Verspätungspropagation durch Einsatz eines zweiphasigen Column Generation Schemas

(Schritt 1) **Initialisierung**
Löse MDVSP und speichere kostenoptimale Flusslösung x^*.
Dekomponiere Flusslösung x^* lokal in Teilumläufe.
Nimm Teilumläufe als initiale Spaltenmenge $\widetilde{L}'_B$.
Setze i = 0.

Führe für jede Netzwerkschicht $d \in \mathcal{D}$ folgende Phase I und Phase II durch:

Phase I: Bestimme robuste Teilumläufe U_B^d auf zugehöriger Flusslösung x^{*d}:

(Schritt 2) **Löse (beschränktes) Master-Problem**
Löse LP-Relaxation von (7.4)-(7.6) mit aktueller Spaltenmenge $\widetilde{L}'^d_B$.
Speichere Zielfunktionswert w^d_{LP} und Dualwerte μ^d.

(Schritt 3) **Löse Pricing-Problem**
Erstelle neue Spalten/ Teilumläufe $\widetilde{L}''^d_B$ mit negativen reduzierten Kosten auf Teilumlaufgenerierungsnetzwerk $\widetilde{G}^d_{B_{\widetilde{V}}}$.
Falls $|\widetilde{L}''^d_B| = 0$ gehe zu Schritt 6.

(Schritt 4) **Führe Spaltenmanagement durch**
Füge neue Spalten zum Master-Problem hinzu: $\widetilde{L}'^d_B := \widetilde{L}'^d_B \cup \widetilde{L}''^d_B$.
Falls $|\widetilde{L}'^d_B| \geq L_{\max}$ lösche Spalten mit hohen positiven reduzierten Kosten von $\widetilde{L}'^d_B$.

(Schritt 5) **Prüfe Abbruchkriterien**
Gehe zu Schritt 6, falls eins folgender Kriterien erfüllt ist:
$\text{i} \geq \text{i}_{\max}$
Keine signifikante Verbesserung von w^d_{LP}
Andernfalls setze i = i + 1 und gehe zu Schritt 2.

(Schritt 6) **Bestimme eine ganzzahlige Lösung**
Löse (7.4)-(7.6) mit aktueller Spaltenmenge $\widetilde{L}'^d_B$.
Speichere Teilumläufe U_B^d und Zielfunktionswert w^d

Phase II: Bestimme robuste Umläufe aus Teilumläufen U_B^d:

(Schritt 7-11) **Führe Schritte analog zu Phase I aus**
Betrachte dabei Formulierung (7.11)-(7.14).
Nutze als Initiallösung Umlaufplan mit minimaler Fahrzeuganzahl basierend auf U_B^d.
Verwende Spaltenmengen L'^d_B und L''^d_B, sowie Umlaufgenerierungsnetzwerke $G^d_{B_V}$.

Phase I: Erstellen von Teilumläufen mit minimaler Verspätungspropagation

Da in der ersten Phase der globalen Flussdekomposition zunächst nur Teilumläufe bzw. Umlaufblöcke betrachtet werden, sind für diese Phase nur Aktivitäten außerhalb des Depots relevant. Dabei sei angemerkt, dass eine Minimierung der Verspätungspropagation in den Teilumläufen dann eine untere Schranke für die gesamte Verspätungspropagation innerhalb des Umlaufplanungsnetzwerks liefert.

Für die Formulierung (7.1) - (7.3) werden statt kompletten Tagesumläufen $l \in L_B^d$ Teilumläufe l betrachtet mit $l \in \widetilde{L}_B^d \subseteq L_B^d$. Da beim Erstellen von Teilumläufen das Warten innerhalb der Depots und der Fluss auf der Zirkulationsflusskante nicht einbezogen werden muss, wird während der ersten Phase der Dekomposition die zu betrachtende Kantenmenge des Umlaufplanungsnetzwerkes G^d entsprechend angepasst. Die Menge aller Kanten, die Fahrzeugaktivitäten außerhalb des Depots repräsentieren, sei durch $\widetilde{A}^d \subset A^d$ beschrieben. Dann sei mit $\widetilde{A}_B^d = \widetilde{A}^d \cap A_B^d$ die Menge aller Kanten außerhalb des Depots definiert, deren entsprechende Fahrzeugaktivitäten mit Lösungsfluss belegt sind. Die Menge aller Teilumläufe, die Kante $(i,j) \in \widetilde{A}_B^d$ enthalten, sei durch $\widetilde{L}_B^d(i,j)$ definiert. In der erste Phase der globalen Flussdekomposition einer Netzwerkschicht $d \in \mathcal{D}$ wird dann der Teilgraph $\widetilde{G}_B^d = (N_B^d, \widetilde{A}_B^d)$ betrachtet und es ergibt sich pro Netzwerkschicht eine modifizierte Version des globalen Flussdekompositionsmodells (GFD-R-VSP I).

GFD-R-VSP I:

$$\sum_{l \in \widetilde{L}_B^d} x_l^d w_l^d \to \min \tag{7.4}$$

$$\text{s.t.} \sum_{l \in \widetilde{L}_B^d(i,j)} x_l^d = b_{ij}^d \qquad \forall (i,j) \in \widetilde{A}_B^d \tag{7.5}$$

$$x_l^d \in \{0,1\} \qquad \forall l \in \widetilde{L}_B^d \tag{7.6}$$

Die Zielfunktion (7.4) minimiert die Verspätungspropagation der Teilumläufe (und damit auch die der Umläufe) während Nebenbedingungen (7.5) garantieren, dass der vorhandene Fluss vollständig in Teilumläufe zerlegt wird. Bedingungen (7.6) beschreiben den binären Charakter der Entscheidungsvariablen für (Teil)Umläufe.

Modell R-GFD-VSP I wird je Netzwerkschicht mit einem Column Generation Ansatz gelöst. *Initiallösung* für die globale Flussdekomposition ist eine (kostenoptimale) Flusslösung x^*, die sich durch Lösen eines MDVSP mit einem LP/ MIP Solver auf der in Abschnitt 6.1 vorgestellten Netzwerkstruktur ergibt. Um eine initiale Spaltenmenge $\widetilde{L}_B'^d \subset L_B^d$ für das Master-Problem jeder Netzwerkschicht $d \in \mathcal{D}$ zu erhalten, wird die Flusslösung x^* lokal in Umläufe bzw. Teilumläufe dekomponiert (Schritt 1).

Für die sich ergebenden Teilumläufe wird basierend auf den darin auszuführenden Fahrzeugaktivitäten der exakte Wert der Verspätungspropagation w berechnet.

Das *(beschränkte) Master-Problem* jeder Netzwerkschicht wird als LP-Relaxation mit einem LP-Solver gelöst (Schritt 2). Mit Hilfe der Dualwerte μ^d der Flussabdeckungsbedingungen (7.5) aus der jeweiligen Lösung des Master-Problems werden im Column Generation Pricing, die Teilumläufe gesucht, mit denen der aktuelle Zielfunktionswert, also die Propagation von Verspätungen innerhalb der Teilumläufe, verringert werden kann.

Beim Lösen der *Pricing-Probleme* hat sich im Rahmen von Ressourceneinsatzplanungsproblemen etabliert, je Column Generation Iteration nicht nur einen Ressourceneinsatz bzw. in diesem Fall einen Teilumlauf zu generieren, der mit den besten reduzierten Kosten den Zielfunktionswert verbessern kann, sondern mehrere, um den Lösungsprozess zu beschleunigen (Schritt 3). In den Pricing-Problemen werden daher für jede Netzwerkschicht mehrere Teilumläufe mit negativen reduzierten Kosten gesucht. Wenn $\widetilde{A}_B^d(l)$ die Menge aller Kanten bezeichnet, die Teilumlauf l von der Kantenmenge $\widetilde{A}_B^d$ abdeckt, sind also Teilumläufe $l \in \widetilde{L}_B^d$ gesucht, für die gilt $\bar{c}_l < 0$, wobei die reduzierten Kosten $\bar{c}_l$ definiert sind als:

$$\bar{c}_l = w_l - \sum_{\{(i,j) \in \widetilde{A}_B(l)\}} \mu_{ij}^d \tag{7.7}$$

Dabei setzen sich die reduzierten Kosten zum einen aus den kantenabhängigen (reduzierten) Kosten μ_{ij}^d zusammen und aus pfadabhängigen (reduzierten) Kosten w_l, die die Verspätungspropagation des entsprechenden Teilumlaufs abbilden und nicht direkt auf den Kanten abgebildet werden können, da sie von den zuvor bedienten, aufeinanderfolgenden Fahrtensequenzen abhängen.

Da die Verspätungspropagation nicht einfach der Summe der Verspätungen entlang eines Pfades entspricht, kann das Pricing-Problem also beispielsweise nicht als einfaches Kürzeste-Wege-Problem formuliert und mit einem LP/ MIP-Solver gelöst werden. Als Pricing-Probleme werden deshalb ressourcenbeschränkte Kürzeste-Wege-Probleme aufgestellt, in welchen auf jedem Pfad mindestens eine Servicefahrt bedient werden muss und eine mögliche Propagation von Verspätungen geeignet berücksichtigt werden kann. Gelöst werden die Pricing-Probleme mit einem Label Setting Algorithmus auf Basis dynamischer Programmierung (vgl. [Desrosiers et al., 1995]).

Die *ressourcenbeschränkten Kürzeste-Wege-Probleme* zur Erzeugung (zulässiger) Teilumläufe werden je Netzwerkschicht $d \in \mathcal{D}$ auf Teilumlaufgenerierungsnetzwerken $\widetilde{G}_{B_{\widetilde{V}}}^d$ formuliert, die aus dem mit der Flusslösung belegten Teil $\widetilde{G}_B^d$ des Umlaufplanungsnetzwerkes G^d abgeleitet werden. Um Teilumläufe generieren zu können, wird

das Basisnetzwerk $\widetilde{G}_B^d$ mit einem Quellknoten s und einem Senkeknoten t ergänzt, sowie mit zusätzlichen Kanten vom Quellknoten zu allen Time-Space-Knoten, die den Start einer Ausrückfahrt repräsentieren, und zusätzlichen Kanten zum Senkeknoten von allen Time-Space-Knoten ausgehend, die das Ende einer Einrückfahrt abbilden. Die zusätzlichen Kanten stellen somit Beginn bzw. Ende eines Teilumlaufs dar und jeder Pfad von Quellknoten zu Senkeknoten beschreibt dann einen Teilumlauf, beginnend mit einer Ausrückfahrt und endend mit einer Einrückfahrt in ein Depot.

Mit Hilfe des *Label Setting Algorithmus* kann die Verspätungspropagation in Abhängigkeit der bedienten Fahrtensequenzen abgebildet und berücksichtigt werden. Bezeichne L_i die Menge aller Label an Knoten i im Netzwerk $\widetilde{G}_{B_{\bar{V}}}^d$. Jedes Label $l \in L_i$ repräsentiert dabei einen konkreten, zulässigen (Teil-)Pfad von Quelle s zu Knoten i und besitzt Informationen über Kostenbewertungen sowie Ressourcenverbrauch pro Ressource $r \in R$ auf diesem Pfad. Sei (i, j) eine Kante von Knoten i zu Knoten j, dann wird ein neues Label m an Knoten j gebildet und somit der betrachtete (Teil-)Pfad um diese Kante verlängert, indem die Informationen des vorangegangenen Labels l um die Informationen auf der Kante (i, j) erweitert werden. Dabei beschreibe $\bar{c}(l)$ die kumulierten reduzierten Kosten, $w(l)$ die kumulierte Summe der Verspätungspropagation auf den Servicefahrten des bisherigen (Teil-)Pfads, $p(l)$ die aktuelle Verspätungspropagation, sowie $d^r(l)$ im Allgemeinen den kumulierten Ressourcenverbrauch pro Ressource r eines Labels $l \in L_i$. Die Informationen auf der Kante (i, j) seien generell wie folgt definiert: $\bar{c}_{ij}$ bezeichne die (reduzierten) Kosten, sowie im Allgemeinen d_{ij}^r pro Ressource $r \in R$ einen Ressourcenverbrauch auf der Kante. Die Verspätungspropagation $p(m)$ an Knoten j ergibt sich als Funktion in Abhängigkeit der bisherigen, aktuellen Verspätungspropagation $p(l)$ und einer möglichen primären Verspätung $PD(i, j)$ auf Kante (i, j) abzüglich möglicher Pufferzeit $B(i, j)$ als::

$$p(m) = max\,\{p(l) + PD(i, j) - B(i, j), 0\} \tag{7.8}$$

Dabei beträgt die Pufferzeit $B(i, j)$ im einfachsten Fall für Wartekanten die zeitliche Differenz $t(j) - t(i)$ zwischen den betrachteten Time-Space-Knoten. In Abhängigkeit der Art der Kante werden neue Label dann wie folgt gebildet: Sei $T(i, j) : A \to T$ eine Funktion, die für jede Kante (i, j) die Fahrt $t \in T$ liefert, die durch Kante (i, j) im Netzwerk repräsentiert wird. Falls die Kante zwischen Knoten i und Knoten j eine Servicefahrt repräsentiert, dann ergibt sich als neues Label m an Knoten j

$$L_j^m = (l, p(m), w(l) + p(m), \bar{c}(l) + \bar{c}_{ij} + p(m), ..., d^{|R|}(l) + d_{ij}^{|R|}). \tag{7.9}$$

Andernfalls ergibt sich als neues Label

$$L_j^m = (l, p(m), w(l), \bar{c}(l) + \bar{c}_{ij}, ..., d^{|R|}(l) + d_{ij}^{|R|}). \tag{7.10}$$

Algorithmus 7.3.2 fasst das beschriebene Vorgehen zur Bestimmung von Teilumläufen durch Lösen der entsprechenden ressourcenbeschränkten Kürzeste-Wege-Probleme mit Hilfe eines Label Setting Verfahrens zusammen. Zusätzlich zu den oben beschriebenen Definitionen sei u^r die obere Schranke für den Verbrauch der Ressource $r \in R$ und $\bar{L}_j$ sei die Menge effizienter, nicht dominierter permanenter Label an Knoten j, also die Menge der Label, die im weiteren Algorithmusverlauf nicht weiter verändert wird. Mit diesem Verfahren erstellte Teilumläufe werden dann zur aktuellen Spaltenmenge des (beschränkten) Master-Problems hinzugefügt (Schritt 4 in Algorithmus 7.3.1).

Algorithmus 7.3.2 : Label Setting mit Berücksichtigung der Verspätungspropagation

```
// Schritt 1: Initialisierung
Setze L_s = L̄_s = {(nil, 0, ..., 0)}
Setze L_i = L̄_i = ∅ ∀i ∈ N \ {s}
// Schritt 2: Pfaderweiterungen und Dominanztests
// Wähle Knoten j
foreach j ∈ N \ {s} do
    // Betrachte alle Vorgängerknoten i von Knoten j
    foreach i ∈ {i : (i, j) ∈ A} do
        // Betrachte alle Label an Knoten i
        foreach l ∈ L̄_i do
            // Prüfe Ressourcenverbräuche
            if ∃r ∈ R : d^r(l) > u^r then
                next l
            // Prüfe Aufgabenfolge und erstelle neues Label m
            if T(i, j) ∈ Service then
                L_j^m = (l, p(m), w(l) + p(m), c̄(l) + c̄_ij + p(m), ..., d^|R|(l) + d_ij^|R|)
            else
                L_j^m = (l, p(m), w(l), c̄(l) + c̄_ij, ..., d^|R|(l) + d_ij^|R|)
            L_j = L_j ∪ L_j^m
    // Entferne dominierte Label an Knoten j
    L̄_j = Effiziente_Label(L_j ∪ L̄_j)
```

Werden während der Lösungsprozedur keine neuen Teilumläufe mehr mit negativen reduzierten Kosten gefunden oder greifen andere Terminierungskriterien (Schritt 5),

wird eine ganzzahlige Lösung bestimmt. Zur Ermittlung einer *ganzzahligen Lösung* wird die Formulierung (7.4)-(7.6) für die aktuelle Spaltenmenge mit einem MIP-Solver gelöst (Schritt 6). Als Lösung ergibt sich ein Teilumlaufplan mit geringster Verspätungspropagation bezüglich der betrachteten Verspätungsszenarien.

Phase II: Erstellen von Umläufen mit minimaler Verspätungspropagation

In der zweiten Phase werden die erstellten Teilumläufe zu Tagesumläufen zusammengesetzt. Beim Bilden der Umläufe werden mögliche Propagationen von Verspätungen berücksichtigt, so dass Umläufe mit minimaler Verspätungspropagation entstehen. Dabei ist die Güte der Lösung an der durch Phase I als untere Schranke berechneten Verspätungspropagation bewertbar.

Sei $U_B^d = \{1, ..., n\}$ die Menge aller in Phase I berechneten Teilumläufe der entprechenden Depot-Fahrzeugtypkombination $d \in \mathcal{D}$ und bezeichne L_{max}^d die in der Flusslösung bestimmte maximal erlaubte Anzahl von Fahrzeugumläufen jeder Kombination. Jeder Tagesumlauf $l \in L^d$ besitzt einen Parameter w_l^d, der sich in Abhängigkeit der ausgeführten Teilumläufe ergibt und die gesamte Verspätungspropagation innerhalb des Umlaufes abbildet. Sei zusätzlich $L^d(u)$ die Menge aller Umläufe, die Teilumlauf $u \in U_B^d$ enthalten. Dann ergibt sich pro Netzwerkschicht $d \in \mathcal{D}$ eine zweite modifizierte Version des globalen Flussdekompositionsmodells (GFD-R-VSP II).

GFD-R-VSP II:

$$\sum_{l \in L^d} x_l^d w_l^d \rightarrow \min \tag{7.11}$$

$$\text{s.t.} \sum_{l \in L^d(u)} x_l^d = 1 \qquad \forall u \in U_B^d \tag{7.12}$$

$$\sum_{l \in L^d} x_l^d \leq L_{max}^d \tag{7.13}$$

$$x_l^d \in \{0, 1\} \qquad \forall l \in L^d \tag{7.14}$$

Mit der Zielfunktion (7.11) wird die gesamte Verspätungspropagation aller Umläufe einer Netzwerkschicht minimiert. Nebenbedingungen (7.12) stellen sicher, dass jeder aus Phase I gegebene Teilumlauf genau einmal innerhalb der erstellten Umläufe ausgeführt wird. Restriktion (7.13) begrenzt die maximale Anzahl erlaubter Tagesumläufe je Depot bzw. Netzwerkschicht, so dass die Gesamtanzahl der genutzten Umläufe der Umlaufanzahl in der zugrunde liegenden Flusslösung entspricht. Die Bedingung (7.14) bezeichnet die binäre Eigenschaft der Entscheidungsvariablen, die Umläufe respräsentieren.

Schon für kleine Probleminstanzen existieren zu viele Kombinationsmöglichkeiten von Teilumläufen zu Umläufen, um diese effizient zu enumerieren und dann das Modell lösen zu können. Daher werden die entstandenen Optimierungsprobleme R-GFD-VSP II analog zu Phase I mit einem Column Generation Ansatz gelöst (siehe Algorithmus 7.3.1).

Als *Initiallösung* werden die vorhandenen Teilumläufe so zusammengesetzt, dass die Gesamtanzahl der entstehenden Umläufe der Vorgabe aus der Flusslösung entspricht. Dies kann beispielsweise durch eine Modellierung als einfaches Netzwerkflussproblem mit Zielsetzung der Umlaufanzahlminimierung erreicht werden. Dabei werden die Teilumläufe als Knoten modelliert und über Kanten diejenigen Teilumläufe miteinander verbunden, die zeitlich kompatibel zueinander sind. Zusätzlich werden ein Quellknoten und Senkeknoten eingefügt, die durch Kanten zu bzw. von allen Teilumläufen mit diesen verbunden sind. Die vom Quellknoten ausgehenden Verbindungskanten werden mit Kosten von 1 belegt. Alle anderen Kanten haben einen Kostenwert von 0. Auf einer Zirkulationsflusskante zwischen Senke- und Quellknoten ist als Kapazität die in der ursprünglichen Flusslösung berechnete Fahrzeuganzahl (für das betrachtete Depot) abgebildet. Außerdem gilt, dass jeder Teilumlauf(knoten) genau einmal in der Lösung besucht werden muss. Um bereits Robustheitsaspekte zu betrachten und so für eine schnellere Konvergenz des Column Generation Verfahrens zu sorgen, können auch bereits die propagierten Verspätungen zwischen direkt aufeinanderfolgenden Teilumläufen berücksichtigt werden. Dazu kann eine Netzwerkfluss-Formulierung analog der im späteren Abschnitt 7.6.2 beschriebenen Formulierung (7.57) - (7.60) gelöst werden. Eine mögliche Propagation von Verspätungen wird dabei zunächst unterschätzt. Die als Initiallösung erstellten Umläufe werden anschließend anhand ihrer ausgeführten Fahrten mit der jeweils korrekt berechneten Verspätungspropagation w gemäß Formel 5.1 bewertet und bilden dann die initiale Spaltenmenge des (beschränkten) Master-Problems.

Als *(beschränktes) Master-Problem* wird die Formulierung (7.11) - (7.14) LP relaxiert und mit einem LP Solver gelöst. Mit den Dualwerten μ^d für Teilumlaufabdeckungsbedingungen (7.12) und den Dualwerten ν^d für die Fahrzeuganzahlrestriktion (7.13) werden im Column Generation Pricing den aktuellen Zielfunktionswert verbessernde Umläufe gesucht.

In den *Pricing-Problemen* werden dann je Netzwerkschicht $d \in \mathcal{D}$ Umläufe mit negativen reduzierten Kosten gesucht. Dabei sind die reduzierten Kosten $\bar{c}_l$ eines

Umlaufs $l \in L^d$ definiert durch

$$\bar{c}_l = w_l - \sum_{\{u \in U^d(l)\}} \mu_u^d - \nu^d \tag{7.15}$$

wobei w_l die Verspätungspropagation des Umlaufs ist und $U^d(l)$ die Menge der in Umlauf l enthaltenden Teilumläufe von Depot d. Zur Bestimmung entsprechender Umläufe mit negativen reduzierten Kosten werden Umlaufgenerierungsnetzwerke $G^d_{B_V}$ erstellt, die die zuvor erstellen Teilumläufe U^d_B enthalten und mit deren Hilfe die Teilumläufe zu Umläufen so zusammengesetzt werden, dass die Propagation von Verspätungen im gesamten Umlaufplan minimiert wird.

Dafür werden analog zu Phase I *ressourcenbeschränkte Kürzeste-Wege-Probleme* formuliert und mit einem Label Setting Algorithmus gelöst. Im Netzwerk werden Teilumläufe als Kanten dargestellt, die zusammengehörige Start- und Endknoten miteinander verbinden. Die Knoten repräsentieren jeweils Start- und Endzeitpunkt eines Teilumlaufs. Die Knoten, die den Endzeitpunkt eines Teilumlaufs repräsentieren, werden durch Verbindungskanten mit den Startknoten aller zeitlich kompatiblen Teilumläufe verknüpft. Zusätzlich werden ein Knoten s als Quelle und ein weiterer Knoten t als Senke hinzugefügt. Von der Quelle werden Kanten zu allen Startknoten hinzugefügt, die Beginn eines Umlaufs repräsentieren. Zusätzlich verknüpfen Kanten, die das Ende eines Umlaufs darstellen, die Endknoten der einzelnen Teilumläufe mit der Senke. Jeder Pfad von Quellknoten zu Senkeknoten beschreibt dann einen kompletten Umlauf, der aus mindestens einem Teilumlauf besteht.

Mit Hilfe des zuvor beschriebenen *Label Setting Algorithmus* 7.3.2 kann analog zu den Teilumläufen in Phase I die Verspätungspropagation eines Umlaufs in Abhängigkeit der bedienten Teilumläufe und somit der bedienten Fahrtensequenzen berücksichtigt werden. Wie in Phase I bezeichne $\bar{c}(l)$ dabei die kumulierten reduzierten Kosten, $w(l)$ die kumulierte Summe der Verspätungspropagation auf den Servicefahrten des bisherigen (Teil-)Pfads, sowie $p(l)$ die aktuelle Verspätungspropagation eines Labels $l \in L_i$ an Knoten i. Wird der durch Label l beschriebene (Teil-)Pfad über Kante (i, j) des Umlaufgenerierungsnetzwerkes erweitert, ergibt sich die Verspätungspropagation in Abhängigkeit der bisher aktuellen Verspätungspropagation $p(l)$ und der auszuführenden Fahrtenmenge T^{ij} des durch Kante (i, j) repräsentierten Teilumlaufs als Funktion $P(p(l), T^{ij})$. Dabei sei die auszuführende Fahrtenmenge (zeitlich) sortiert, so dass für die aktualisierte Verspätungspropagation gemäß Formel 5.1 gilt

$$P(p(l), T^{ij}) = \sum_{h=1}^{T^{ij}-1} p(t_h, t_{h+1}) \tag{7.16}$$

mit $p(t_h, t_{h+1})$ definiert als

$$p(t_h, t_{h+1}) = p(t_{h-1}, t_h) + PD(t_h, t_{h+1}) - B(t_h, t_{h+1}) \tag{7.17}$$

wobei gilt $p(t_0, t_1) = p(l)$. Ebenso berechnet sich die kumulierte Verspätungspropagation von Servicefahrten in Abhängigkeit der durch Kante (i, j) repräsentierten Fahrten und der bisher aktuellen Verspätungspropagation $p(l)$ als Funktion $W(p(l), T^{ij})$. Falls die Kante (i, j) einen Teilumlauf repräsentiert, wird also ein neues Label m an Knoten j gebildet als:

$$L_j^m = (l, P(p(l), T^{ij}), w(l) + W(p(l), T^{ij}), \bar{c}(l) + \bar{c}_{ij} + W(p(l), T^{ij}), ..., d^{|R|}(l) + d_{ij}^{|R|}). \tag{7.18}$$

Andernfalls ergibt sich als neues Label:

$$L_j^m = (l, p(m), w(l), \bar{c}(l) + \bar{c}_{ij}, ..., d^{|R|}(l) + d_{ij}^{|R|}). \tag{7.19}$$

Ganzzahlige Lösungen werden nach Abschluss des Column Generation Verfahrens wie in Phase I pro Netzwerkschicht d mit einem MIP-Solver berechnet. Als Lösung ergibt sich der (kostenminimale) Umlaufplan mit geringster Verspätungspropagation.

7.4 Umlaufebene – Generierung einer robusten Netzwerkflusslösung

Mit dem im vorangegangenem Abschnitt 7.3 vorgestellten Verfahren der globalen Flussdekomposition lässt sich der in Bezug auf eine mögliche Propagation von Verspätungen robusteste Umlaufplan als Grundlage für die weitere Ressourceneinsatzplanung finden. Wie später in Kapitel 11 gezeigt wird, führt ein solcher Umlaufplan tendenziell zu einer robusteren Gesamtlösung. Die Festlegung auf einen solchen, bereits ausgewählten Umlaufplan schränkt allerdings die Dienstplanung aufgrund fixer Vorgaben von abzudeckenden Fahrtensequenzen stark ein und führt ebenso zu vergleichsweise teureren Dienstplänen (siehe Abschnitt 11.2.2). Deshalb wird in diesem Abschnitt untersucht, wie zum einen mögliche Umlaufpläne (noch) robuster gestaltet werden können und zum anderen weiterhin das Potenzial einer Flusslösung benutzt werden kann, um die bei der Dienstplanung abzudeckenden Umlaufaufgaben nicht als Sequenzen im Vorfeld fixieren zu müssen, sondern mehr Freiheitsgrade für die Dienstbildung zu überlassen. Eine Schwierigkeit eines solchen Ansatzes besteht darin, dass in einer Modellierung als (effizient lösbares) Netzwerkfluss-Problem basierend auf

(Service-)Fahrten bei der Propagation von Verspätungen keine unmittelbare Reihenfolgebeziehung zwischen den Fahrten abbildbar ist. Um exakte Propagationseffekte auf Umläufen abbilden zu können, sollten allerdings vorher ausgeführte Fahrten eines Umlaufes bekannt sein. Trotz dieser Schwierigkeit muss identifiziert werden können, wann ein Netzwerkfluss robuster als ein anderer Netzwerkfluss ist, d.h. wann eine Flusslösung potenziell robustere Umläufe enthält als eine andere Flusslösung.

In den folgenden Abschnitten 7.4.1 und 7.4.2 werden Modell und Lösungsverfahren für einen Ansatz entwickelt, mit dem bereits bei Berechnung der Flusslösung für das Umlaufplanungsproblem die mögliche Propagation von Verspätungen bzw. Robustheit mit einbezogen werden kann, um potenziell ein Bündel robusterer Umlaufpläne erstellen zu können. Dabei ist die Güte der Umlauflösung steuerbar zwischen Kosteneffizienz und/oder Robustheit. Kernidee ist es dabei, die netzwerkflussbasierte Modellierung des Umlaufplanungsproblems, die den „Umlauffluss“ abbildet, mit einer pfadbasierten Modellierung zu verknüpfen, die einzelne Umläufe mit ihren potenziell propagierten Verspätungen repräsentiert. Zur Lösung des Modells wird ein Lösungsschema vorgeschlagen, welches sich am Verfahren von [Huisman et al., 2005] bzw. [Steinzen et al., 2010] für eine kosteneffiziente integrierte Umlauf- und Dienstplanung orientiert, da dort ebenfalls fluss- und pfadbasierte Formulierungen erfolgreich verknüpft und gelöst werden – dort allerdings getrennt auf den unterschiedlichen Ressourcenebenen in eine flussbasierte Beschreibung von Umläufen und eine pfadbasierte Beschreibung von Diensten. Das vorgestellte Verfahren ermöglicht es, eine Flusslösung zu berechnen, die mindestens einen Umlaufplan enthält, der robuster als der robusteste dekomponierte Umlaufplan einer kostenoptimalen Lösung ist, falls ein solcher existiert.

Abbildung 7.4 verdeutlicht das Potenzial eines solchen Vorgehens anhand eines Beispiels. Im Beispielnetzwerk sind fünf Servicefahrten zu bedienen. Der kostenoptimale Fluss erlaubt die Dekomposition in zwei Umlaufpläne mit jeweils zwei Umläufen. In beiden möglichen Plänen führt allerdings eine primäre Verspätung zwangsläufig zu einer Nachfolgeverspätung (aufgrund der Verbindung von Fahrt 4 und 5). Bei einem leicht modifizierten Fluss, der statt des Einrückens ins Depot nach Fahrt 3 eine Standzeit außerhalb des Depots erlaubt, ergeben sich neue Möglichkeiten zur Dekomposition des Lösungsflussen in Umlaufpläne. Unter diesen Möglichkeiten finden sich zwei Pläne, die eine Propagation von Verspätungen zwischen den Servicefahrten verhindern. Im ersten Plan werden Fahrt 1 und 4 von einem Umlauf bedient, und Fahrt 2, 3 sowie 5 von einem zweiten Umlauf. Im zweiten möglichen Plan werden Fahrt 1, 3 sowie 5 von einem Umlauf bedient, und Fahrt 2 und 4 von einem zweiten Umlauf.

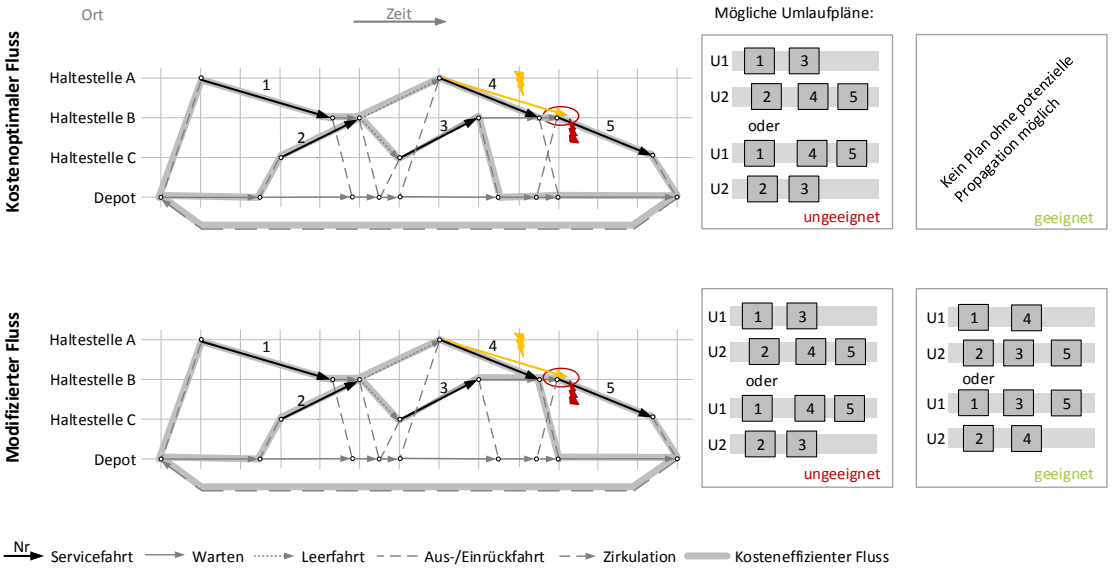

Abbildung 7.4: Flussdekomposition mit Berücksichtigung von Verspätungspropagation in zwei Time-Space-Netzwerken bei unterschiedlich robusten Flusslösungen

7.4.1 Modellierung

Das Umlaufplanungsproblem wird weiterhin als Mehrgüter-Fluss-Problem auf einem Time-Space-Netzwerk wie in Abschnitt 6.1 beschrieben modelliert, bei welchem jede Servicefahrt in der Flusslösung genau einmal bedient werden muss. Zusätzlich wird die Formulierung mit einer Formulierung als Set-Partitioning-Problem ergänzt, bei welchem jede in der Flusslösung bediente Fahrzeugaktivität von einem Umlauf abgedeckt werden muss.

Mathematisch kann das gekoppelte Modell zur robusten Umlaufplanung wie folgt beschrieben werden. Mit $\mathcal{D} = \{1, 2, .., m\}$ sei die Menge aller Kombinationen aus Depotstandort und Fahrzeugtyp bezeichnet und mit $\mathcal{T} = \{1, 2, .., n\}$ die Menge aller zu bedienenden Servicefahrten. Für jedes Depot d wird ein wie in Abschnitt 6.1 beschriebenes TSN basiertes Umlaufplanungsnetzwerk $G^d = (N^d, A^d)$ aufgestellt, das aus der Knotenmenge N^d und der Kantenmenge A^d besteht. Sei zusätzlich mit $A^d(t)$ die Menge der Kanten $(i, j) \in A^d$ definiert, die Servicefahrt $t \in \mathcal{T}$ im Umlaufplanungsnetzwerk von Depot $d \in \mathcal{D}$ repräsentieren. Falls t von Depot d aus ausgeführt werden kann, enthält $A^d(t)$ genau ein Element, ansonsten gilt $A^d(t) = \emptyset$.

Jeder Kante $(i, j) \in A^d$ werden Kosten c^d_{ij} zugeordnet, die Fahrzeugkosten entsprechen, die sich aus der durch die Kante repräsentierte Fahrzeugaktivität ableiten lassen. Die Zirkulationsflusskante eines Umlaufplanungsnetzwerks bzw. einer Netz-

werkschicht wird mit fixen Kosten für einen Fahrzeugeinsatz des entsprechenden Depots belegt. Die maximale Flusskapazität u_{ij}^d einer Kante $(i,j) \in A^d$ ist für alle Servicefahrten auf 1 beschränkt und für alle anderen Kanten durch die maximal pro Depot d verfügbaren Fahrzeuganzahl u^d.

Als Entscheidungsvariablen werden zwei Arten Variablen definiert: Flussvariablen und pfadbasierte Variablen. Über ganzzahlige Flussvariablen y_{ij}^d für jede Kante $(i,j) \in A^d$ wird angegeben, inwiefern Kante (i,j) mit einem Fluss belegt ist. Dabei repräsentiert jede Flusseinheit ein Fahrzeug, das die durch Kante (i,j) beschriebene Aktivität ausführt. Binäre Entscheidungsvariablen z_l^d mit $l \in L^d$ und zugehöriger potenzieller Verspätungspropagation w_l^d gemäß Formel 5.1 geben an, ob Umlauf l für Depot $d \in D$ ausgewählt wird oder nicht. Dabei sei mit L^d die Menge aller gültigen Umläufe definiert, die von Depot d aus durchgeführt werden können. Des weiteren sei $L^d(i,j) \subset L^d$ die Menge aller Umläufe, die die durch Kante $(i,j) \in A^d$ repräsentierte Fahrzeugaktivität bedienen. *Das robuste Umlaufplanungsproblem mit mehreren Depots und Berücksichtigung von Verspätungspropagation* (engl. *robust multiple-depot vehicle scheduling problem, R-MDVSP*) kann wie folgt formuliert werden.

R-MDVSP:

$$\sum_{d\in\mathcal{D}} \sum_{(i,j)\in A^d} y_{ij}^d c_{ij}^d + \alpha_{vs} \sum_{d\in\mathcal{D}} \sum_{l\in L^d} z_l^d w_l^d \to \min \tag{7.20}$$

$$\text{s.t.} \sum_{d\in\mathcal{D}} \sum_{(i,j)\in A^d(t)} y_{ij}^d = 1 \qquad \forall t \in \mathcal{T} \tag{7.21}$$

$$\sum_{\{j:(j,i)\in A^d\}} y_{ji}^d - \sum_{\{j:(i,j)\in A^d\}} y_{ij}^d = 0 \qquad \forall d \in \mathcal{D}, \forall i \in N^d \tag{7.22}$$

$$\sum_{l\in L^d(i,j)} z_l^d - y_{ij}^d = 0 \qquad \forall d \in \mathcal{D}, \forall (i,j) \in A^d \tag{7.23}$$

$$0 \le y_{ij}^d \le u_{ij}^d,\ y_{ij}^d \in \mathbb{N} \qquad \forall d \in \mathcal{D}, \forall (i,j) \in A^d \tag{7.24}$$

$$z_l^d \in \{0,1\} \quad \forall d \in \mathcal{D}, \forall l \in L^d \tag{7.25}$$

Mit der Zielfunktion (7.20) werden zum einen die Fahrzeugkosten minimiert. Zusätzlich wird in Abhängigkeit des gewählten Bestrafungsfaktors α_{vs} die Propagation von Verspätungen innerhalb der Umläufe minimiert. Über die Restriktionen (7.21)-(7.22) ist das Umlaufplanungsproblem als Mehrgüter-Fluss-Formulierung abgebildet. Die Restriktionen garantieren einen gültigen Umlaufplan. Dabei stellen Nebenbedingungen (7.21) sicher, dass jede Fahrt genau einmal von einem Fahrzeug bedient wird, während Nebenbedingungen (7.22) durch Flusserhaltung an jedem Knoten einen zulässigen Fluss gewährleisten. Die Kopplung der flussbasierten mit der pfadbasierten Umlaufformulierung erfolgt durch die Nebenbedingungen (7.23). Bei jeder durch

einen Fahrzeugfluss abgedeckten Kante, muss das durch sie repräsentierte Umlaufelement in ebenso vielen Umläufen enthalten sein. Nebenbedingungen (7.24) stellen sicher, dass der maximal mögliche Fluss auf den Kanten eingehalten wird. Die Nebenbedingungen (7.25) beschreiben den binären Charakter der Entscheidungsvariablen, die pfadbasierte Umläufe repräsentieren.

Alternativ lässt sich der Zielkonflikt zwischen Kostenminimierung und Robustheitserhöhung mit Hilfe einer Elastic-Constraint-Formulierung (siehe Abschnitt 3.1.2) im Modell abbilden. Ein Vorteil einer solchen Formulierung besteht in der Möglichkeit, einen genauen Rahmen vorgeben zu können, zu welchen Kosten die Robustheit erhöht werden darf. Dabei bezeichne ϵ das vorgegebene Kostenlimit (in Abhängigkeit einer kostenoptimalen Lösung). Schlupfvariable s^l und Surplusvariable s^u stellen sicher, dass die (elastische) Kostenrestriktion immer erfüllt sein kann. Wird das Kostenlimit überschritten, wird diese Abweichung durch den der Variable s^u zugeordneten Faktor p^u in der Zielfunktion bestraft. Das *robuste Umlaufplanungsproblem mit mehreren Depots und Berücksichtigung von Verspätungspropagation mit elastischer Kostenbedingung* (engl. *robust multiple-depot vehicle scheduling problem with elastic cost costraint, R-MDVSP-E*) kann dann wie folgt formuliert werden.

R-MDVSP-E:

$$\sum_{d\in\mathcal{D}}\sum_{l\in L^d} z_l^d w_l^d + p^u s^u \to \min \tag{7.26}$$

$$\text{s.t.} \sum_{d\in\mathcal{D}}\sum_{(i,j)\in A^d} y_{ij}^d c_{ij}^d + s^l - s^u = \epsilon \tag{7.27}$$

$$\sum_{d\in\mathcal{D}}\sum_{(i,j)\in A^d(t)} y_{ij}^d = 1 \qquad \forall t \in \mathcal{T} \tag{7.28}$$

$$\sum_{\{j:(j,i)\in A^d\}} y_{ji}^d - \sum_{\{j:(i,j)\in A^d\}} y_{ij}^d = 0 \qquad \forall d \in \mathcal{D}, \forall i \in N^d \tag{7.29}$$

$$\sum_{l\in L^d(i,j)} z_l^d - y_{ij}^d = 0 \qquad \forall d \in \mathcal{D}, \forall (i,j) \in A^d \tag{7.30}$$

$$0 \le y_{ij}^d \le u_{ij}^d,\ y_{ij}^d \in \mathbb{N} \qquad \forall d \in \mathcal{D}, \forall (i,j) \in A^d \tag{7.31}$$

$$z_l^d \in \{0,1\} \quad \forall d \in \mathcal{D}, \forall l \in L^d \tag{7.32}$$

$$s^l, s^u \ge 0 \tag{7.33}$$

Die Zielfunktion (7.26) minimiert die Verspätungspropagation innerhalb der Umläufe. Die elastische Kostenrestriktion (7.27) beschränkt dabei die Kosten in Bezug auf eine z.B. bekannte kostenoptimale Lösung. Die übrigen Restriktionen (7.28) - (7.32) entsprechen der Formulierung (7.21) - (7.25). Zusätzlich sind in Bedingung

(7.33) Schlupf- und Surplusvariable noch nach unten beschränkt. Da jedoch die Formulierung R-MDVSP rechentechnisch vorteilhafter ist, wird im weiteren Verlauf der Arbeit, wenn nicht anders angegeben, Modell R-MDVSP verwendet.

7.4.2 Lösungsverfahren

Das mit der Formulierung (7.20) - (7.25) beschriebene Umlaufplanungsproblem unter Berücksichtigung von Verspätungspropagation (R-MDVSP) wird durch einen Column Generation Ansatz in Kombination mit Lagrange-Relaxation gelöst. Die Column Generation Pricing-Probleme zur Generierung neuer Spalten bzw. Umläufe werden dabei ähnlich zu den im Abschnitt 7.3.2 beschriebenen Pricing-Problemen als ressourcenbeschränktes Kürzeste-Wege-Problem formuliert, in denen insbesondere die Propagation von Verspätungen eines Umlaufes abgebildet werden. Zulässige Lösungen werden mit Hilfe einer Lagrange-Heuristik ermittelt. Die generellen Vorgehensweisen bei der Initialisierung des Verfahrens, beim Lösen des Master-Problems und der Pricing-Probleme sowie beim Erstellen der zulässigen Lösungen werden nachfolgend genauer vorgestellt.

Initialisierung

Als Initiallösung dient ein aus der Flusslösung dekomponierter (kostenoptimaler) Umlaufplan. Dieser kann auch der robusteste unter den kostenoptimalen Plänen sein, welcher mit dem in Abschnitt 7.3 beschriebenen Verfahren der globalen Flussdekomposition erzeugt wurde. Die durch den Plan beschriebenen Umläufe bilden dann die initialen Spalten des (beschränkten) Master-Problems für das Column Generation Verfahren.

Master-Problem: Lagrange-Relaxation

Um die Betrachtung des (beschränkten) Master-Problems zu vereinfachen, werden die Kopplungsbedingungen (7.23) Lagrange-relaxiert und mit einem Lagrange-Multiplikator assoziiert in die Zielfunktion überführt. Dadurch entstehen zwei separate Teilprobleme: Ein klassisches MDVSP mit Zielsetzung der Kostenminimierung und ein als einfaches Auswahlproblem formuliertes Umlaufplanungsproblem zur Minimierung von Verspätungspropagation. Da das MDVSP ein NP-schweres Problem ist, können auch die Fahrtabdeckungsbedinungen (7.21) Lagrange-relaxiert werden, um das Problem weiter zu vereinfachen (vgl. [Steinzen et al., 2010]). Die Zielfunktion des rela-

xierten Modells kann dann wie folgt formuliert werden:

$$\begin{aligned} \min \quad & \sum_{d\in\mathcal{D}} \sum_{(i,j)\in A^d} y_{ij}^d c_{ij}^d + \alpha_{vs} \sum_{d\in\mathcal{D}} \sum_{l\in L^d} z_l^d w_l^d \\ + \sum_{d\in\mathcal{D}} \sum_{(i,j)\in A^d} \nu_{ij}^d \left(\sum_{d\in\mathcal{D}} y_{ij}^d - \sum_{l\in L^d(i,j)} z_l^d \right) & + \sum_{t\in\mathcal{T}} \pi_t \left(1 - \sum_{d\in\mathcal{D}} \sum_{(i,j)\in A^d(t)} y_{ij}^d \right) \end{aligned} \tag{7.34}$$

Das Lagrange-Unterproblem kann daher geschrieben werden als:

$$\Phi(\pi,\nu) = \Phi_y(\pi,\nu) + \Phi_z(\nu) + \sum_{t\in T} \pi_t \tag{7.35}$$

mit dem Umlaufplanungs-Unterproblem zur Kostenminimierung

$$\Phi_y(\pi,\nu) \;=\; \Bigg\{ \min \sum_{d\in D} \sum_{(i,j)\in A^d} y_{ij}^d \bar{c}_{ij}^d \;| \tag{7.36}$$

$$\sum_{\{j:(j,i)\in A^d\}} y_{ji}^d - \sum_{\{j:(i,j)\in A^d\}} y_{ij}^d = 0, \quad \forall d\in\mathcal{D}, \forall i\in N^d, \tag{7.37}$$

$$0 \le y_{ij}^d \le u_{ij}^d, \qquad \forall d\in\mathcal{D}, \forall (i,j)\in A^d \Bigg\} \tag{7.38}$$

und dem Umlauf-Auswahlproblem zur Minimierung der Verspätungspropagation

$$\Phi_z(\nu) \;=\; \Bigg\{ \min \sum_{d\in D} \sum_{l\in L^d} z_l^d \bar{w}_l^d \;| \tag{7.39}$$

$$z_l^d \in \{0,1\}, \quad \forall l\in L^d \Bigg\} \tag{7.40}$$

Dabei sind die reduzierten Kosten $\bar{c}_{ij}^d$ auf Kante $(i,j)\in A^d$ definiert als:

$$\bar{c}_{ij}^d = \begin{cases} c_{ij}^d + \nu_{ij}^d - \pi_t & \text{für } : (i,j)\in \widetilde{A}^d \text{ und } \exists t\in\mathcal{T} : (i,j)\in A^d(t) \\ c_{ij}^d + \nu_{ij}^d & \text{für } : (i,j)\in \widetilde{A}^d \text{ und } \neg\exists t\in\mathcal{T} : (i,j)\in A^d(t) \\ c_{ij}^d & \text{für } : (i,j)\notin \widetilde{A}^d \end{cases} \tag{7.41}$$

Für die reduzierten Kosten $\bar{w}_l^d$ für Umlauf $l\in L^d$ gilt:

$$\bar{w}_l^d = \alpha_{vs} w_l^d - \sum_{(i,j)\in A^d(l)} \nu_{ij}^d \tag{7.42}$$

Zu gegebenen Lagrange-Multiplikatoren ν und π kann das Umlaufplanungs-Unterproblem $\Phi_y(\pi, \nu)$ in einzeln zu lösende Minimalkosten-Fluss-Probleme aufgeteilt werden. Das Umlaufplanungs-Unterproblem $\Phi_z(\nu)$ stellt zu gegebenen Multiplikatoren ν ein einfaches Auswahlproblem dar, in welchem jede Spalte $l \in L^d$ mit negativen reduzierten Kosten ausgewählt wird. Damit für die aktuell im Modell enthaltene Spaltenmenge eine möglichst gute untere Schranke zur nicht relaxierten Formulierung gegeben werden kann, werden die Lagrange-Multiplikatoren ν und π genutzt, die den Wert von $\Phi(\pi, \nu)$ maximieren. Dazu wird das Lagrange-Duale-Problem mit einem modifizierten Subgradientenverfahren gelöst. Für Details über das verwendete Subgradientenverfahren sei an dieser Stelle auf [Gintner, 2008] bzw. [Steinzen et al., 2010] verwiesen.

Pricing-Problem: Teilumlauf-basierte ressourcenbeschränkte kürzeste Wege

Mit Hilfe der Dualwerte bzw. der ermittelten Lagrange-Multiplikatoren aus der Lösung des (beschränkten) Master-Problems, können im Pricing-Problem neue Spalten mit negativen reduzierten Kosten erzeugt werden, die den aktuellen Zielfunktionswert verbessern können. Analog zum Vorgehen bei der globalen Flussdekomposition wird aufgrund der extrem hohen Anzahl möglicher Spalten wie in Abschnitt 7.3.2 ein zweiphasiges Pricing zur Erstellung von Umläufen verwendet.

In der ersten Phase werden unter Berücksichtigung der Propagation von Verspätungen Pfade erstellt, die Teilumläufen bzw. Umlaufblöcken entsprechen. In der zweiten Phase werden aus den Teilumläufen – unter Berücksichtigung der Verspätungspropagation – zulässige Fahrzeugumläufe gebildet. In beiden Phasen wird dabei je Column Generation Iteration nicht nur eine Spalte, die Spalte mit den besten negativen reduzierten Kosten, hinzugefügt, sondern wie in Ansätzen zur Ressourceneinsatzplanung etabliert mehrere (Multiple-Pricing), um die Konvergenz des Verfahrens zu verbessern.

Pricing Phase I: Teilumlauferstellung unter Minimierung der Verspätungspropagation

Analog zum Pricing-Problem der ersten Phase der globalen Flussdekomposition in Abschnitt 7.3.2 können für jede Depot-Fahrzeugtypkombination $d \in D$ Teilumlaufgenerierungsnetzwerke $\widetilde{G}^d_{\widetilde{V}}$ basierend auf den Netzwerken $\widetilde{G}^d = (N^d, \widetilde{A}^d)$ erstellt werden. Im Gegensatz zu den Netzwerken $\widetilde{G}^d_{B_{\widetilde{V}}}$ aus dem ersten Pricing-Problem der globalen Flusslösung beinhalten die Teilumlaufgenerierungsnetzwerke $\widetilde{G}^d_{\widetilde{V}}$ allerdings nicht nur die mit einer Flusslösung belegten Kanten und Knoten außerhalb des De-

pots, sondern mit der der Kantenmenge $\widetilde{A}^d$ alle Kanten außerhalb des Depots sowie mit der Knotenmenge N^d alle Knoten des ursprünglichen Umlaufplanungsnetzwerkes. Die Knoten- und Kantenmengen werden jeweils ergänzt um einen Quellknoten s und einen Senkeknoten t mit Verbindungskanten zu Time-Space-Knoten, von denen mindestens eine Ausrückfahrtkante ausgeht bzw. in die mindestens eine Einrückfahrtkante eingeht. Auf allen Kanten wird ein Ressourcenverbrauch abgebildet, wobei als Ressourcen die Anzahl der durch die Kante repräsentierten Servicefahrten und die primäre Verspätungspropagation entsprechen assoziiert werden.

Ähnlich zur Definition (7.7) sind Teilumläufe mit negativen reduzierten Kosten gesucht, die sich aus kanten- und pfadbasierten (reduzierten) Kosten zusammensetzen:

$$\bar{c}_l = w_l - \sum_{\{(i,j)\in\widetilde{A}^d(l)\}} \nu_{ij}^d \tag{7.43}$$

wobei $\widetilde{A}^d(l)$ die Menge aller Kanten bezeichnet, die im durch Teilumlauf $l \in \widetilde{L}^d$ repräsentierten Pfad abgedeckt werden, und w_l die pfadabhängige Summe der propagierten Verspätungen auf den Servicefahrten des Teilumlaufs.

Die so entstehenden ressourcenbeschränkten Kürzeste-Wege-Probleme (R-VSP-CG-P I) werden mit dem zuvor in Abschnitt 7.3.2 beschriebenen Label-Setting Ansatz 7.3.2 nach [Desrosiers et al., 1995] gelöst. Die einzelnen Label an den Knoten des Teilumlaufgenerierungsnetzwerkes $\widetilde{G}^d_{\widetilde{V}}$ werden dabei gebildet, wie in Phase I der globalen Flussdekomposition dargestellt. Wie bei der globalen Flussdekomposition bilden die erstellten Teilumläufe aus der ersten Phase die Eingabe für die Generierung von Fahrzeugumläufen in einer zweiten Phase.

Pricing Phase II: Umlauferstellung unter Minimierung der Verspätungspropagation

Nachdem zulässige Teilumläufe gebildet wurden, werden diese genutzt, um in einem zweiten Schritt robuste Umläufe zu bilden. Dafür wird ein ähnliches Netzwerk erstellt, welches Teilumläufe und ihre möglichen Verknüpfungen miteinander repräsentiert, wie in der zweiten Phase der globalen Flussdekomposition in Abschnitt 7.3.2 beschrieben. Es werden diejenigen zuvor generierten Teilumläufe miteinander verknüpft, die hintereinander ausführbar sind. Auf allen Kanten wird ein Ressourcenverbrauch abgebildet, der dem aus Abschnitt 7.3.2 bekannten Verbrauch entspricht. Die entstehenden ressourcenbeschränkten Kürzeste-Wege-Probleme (R-VSP-CG-P II) werden mit dem angepassten Label Setting Ansatz auf Basis dynamischer Programmierung 7.3.2 zur korrekten Abbildung von Verspätungspropagation in den Umläufen erstellt.

Die einzelnen Label an den Knoten des Umlaufgenerierungsnetzwerkes G_V^d werden dabei so gebildet wie in Phase II der globalen Flussdekomposition beschrieben.

Die in der zweiten Phase erstellten Umläufe werden als Spalten in das (beschränkte) Master-Problem aufgenommen. Werden keine Umläufe mehr mit negativen reduzierten Kosten erstellt oder ist die Verbesserung des Zielfunktionswertes durch Hinzufügen neuer Spalten über mehrere Iterationen unter einem bestimmten Schwellenwert, wird mit Hilfe einer Lagrange-Heuristik eine zulässige Lösung des ursprünglichen Umlaufplanungsproblems bestimmt.

Ganzzahlige Lösung

Eine zulässige Flusslösung wird ermittelt, indem zunächst bei Formulierung R-MDVSP nur noch Restriktion (7.23) Lagrange-relaxiert wird. Das entsprechende Lagrange-Duale-Problem wird dann wie zuvor mit einem Subgradientenverfahren gelöst, wobei die Lagrange-Multiplikatoren mit den zuletzt in den Column Generation Schritten ermittelten Werten initialisiert werden, um die Konvergenz des Verfahrens zu verbessern. In Abhängigkeit der erstellten (robusten) Umläufe ergibt sich so ein zulässiger (robuster) Fluss für das ursprüngliche Umlaufplanungsproblem. Als zulässige Flusslösung kann dann die beste Lösung (bezüglich der Ziele Kosteneffizenz und Robustheit) aus den letzten Iterationen des Subgradientenverfahrens übernommen werden. In dem in dieser Arbeit entwickelten teilintegrierten Verfahren zur (robusten) Umlauf- und Dienstplanung kann z.B. die Flusslösung ausgewählt werden, durch die in Interaktion mit der nachfolgenden Dienstplanung die bestmöglichen Ressourceneinsatzpläne bezüglich der Zielkriterien erreicht werden können.

7.5 Dienstebene

Innerhalb des teilintegrierten Verfahrens basiert die Dienstplanung wie bei einer sequenziellen Vorgehensweise auf Vorgaben aus der Umlaufplanung. Dadurch, dass das teilintegrierte Dienstplanungsproblems allerdings nicht basierend auf einem festen Umlaufplan formuliert wird, sondern basierend auf einer Flusslösung des Umlaufplanungsproblems, also einem Bündel möglicher Umlaufpläne, existieren mehr Freiheitsgrade für die Erstellung kosteneffizienter und robuster Dienste. Abschnitt 7.5.1 stellt das Modell zur teilintegrierten Dienstplanung vor. In Abschnitt 7.5.2 wird das entsprechende Lösungsverfahren genauer erläutert.

7.5.1 Modellierung

Mathematisch kann das (teilintegrierte) Dienstplanungsproblem als Set-Partitioning-Problem formuliert werden. Dabei ist die Menge der durch Dienste abzudeckenden Aufgaben je Depot $d \in \mathcal{D}$ durch die Kanten $\widetilde{A}_B^d$ des Umlaufplanungsnetzwerkes G^d (siehe Abschnitt 7.3.1) vorgegeben, die Fahrzeugkaktivitäten außerhalb des Depots beschreiben und im Lösungsfluss des Umlaufplanungsproblems enthalten sind. Wie oft eine Aufgabe abzudecken ist, sei je Kante $(i,j) \in \widetilde{A}_B^d$ durch b_{ij}^d beschrieben. Binärvariablen x_k^d für jeden Dienst $k \in K_B^d$ geben an, ob der Dienst für Depot d in der Lösung ausgewählt ist oder nicht. Dabei sei die Menge K_B^d die Menge aller möglichen Dienste für den gegebenen Lösungsfluss einer Netzwerkschicht. Die Menge aller Dienste, die Kante $(i,j) \in \widetilde{A}_B^d$ enthalten, sei durch $K_B^d(i,j)$ definiert. Mit jedem Dienst in der Zielfunktion assoziiert sind entsprechende Dienstkosten f_k^d sowie die Summe der propagierten Verspätungen v_k^d der vom Dienst auszuführenden Servicefahrten gemäß Formel 5.1, gewichtet mit einem Faktor α_{cs}.

Das *robuste Dienstplanungsproblem mit Berücksichtigung von Verspätungspropagation* (engl. *robust crew scheduling problem, R-CSP*) kann dann wie folgt formuliert werden.

R-CSP:

$$\sum_{k \in K_B^d} x_k^d (f_k^d + \alpha_{cs} v_k^d) \to \min \tag{7.44}$$

$$\text{s.t.} \sum_{k \in K_B^d(i,j)} x_k^d = b_{ij}^d \qquad \forall (i,j) \in \widetilde{A}_B^d \tag{7.45}$$

$$x_k^d \in \{0,1\} \qquad \forall k \in K_B^d \tag{7.46}$$

Die Zielfunktion (7.44) minimiert die Kosten des Diensteinsatzes sowie in Abhängigkeit des Penalty-Parameters α_{cs} die Propagation von Verspätungen innerhalb der Dienste. Die Aufgabenabdeckungbedingungen (7.45) stellen sicher, dass jede durch Kante $(i,j) \in \widetilde{A}_B^d$ repräsentiert Fahrzeugaktivität, die sich aus der Flusslösung des Umlaufplanungsproblems ergibt und einen Fahrer erfordert, so oft wie benötigt von einem Dienst abgedeckt wird. Die Nebenbedingungen (7.46) charakterisieren die binäre Eigenschaft der Entscheidungsvariablen für Dienste.

Alternativ lässt sich wie bereits für die Umlaufebene in Abschnitt 7.4.1 beschrieben der Zielkonflikt zwischen Kosteneffizienz und Robustheit auch für die Dienstebene mit Hilfe einer elastischen Kostenrestriktion modellieren. Mit ϵ sei der Kostenrahmen bezeichnet. Des weiteren wird mit Hilfe einer Schlupfvariablen s^l und einer Surplusvariablen s^u sichergestellt, dass der Kostenrahmen immer erfüllt sein kann. Eine Über-

schreitung des Kostenrahmens wird gewichtet mit dem Faktor p^u in der Zielfunktion bestraft. Das *robuste Dienstplanungsproblem mit Berücksichtigung von Verspätungspropagation und elastischer Kostenrestriktion* (engl. *robust crew scheduling problem with elastic cost constraint, R-CSP-E*) kann dann folgendermaßen formuliert werden.

R-CSP-E:

$$\sum_{k \in K_B^d} x_k^d v_k^d + p^u s^u \to \min \tag{7.47}$$

$$\text{s.t.} \sum_{k \in K_B^d} x_k^d f_k^d + s^l - s^u = \epsilon \tag{7.48}$$

$$\sum_{k \in K_B^d(i,j)} x_k^d = b_{ij}^d \qquad \forall (i,j) \in \widetilde{A}_B^d \tag{7.49}$$

$$x_k^d \in \{0,1\} \qquad \forall k \in K_B^d \tag{7.50}$$

$$s^l, s^u \geq 0 \tag{7.51}$$

Mit der Zielfunktion (7.47) wird die Propagation von Verspätungen minimiert. Durch die elastische Kostenrestriktion (7.48) wird der zur Verspätungminimierung einsetzbare Kostenrahmen eingehalten. Die Aufgabenabdeckungsbedingungen (7.49) entsprechen den Restriktionen (7.45) und die Nebenbedingungen (7.50) entsprechen den Restriktionen (7.46). Darüber hinaus sind über Bedingung (7.51) Schlupf- und Surplusvariable nach unten beschränkt. Da sich allerdings bei der Modellevaluation gezeigt hat, dass Modell R-CSP für umfangreiche Testrechnungen vorteilhafter ist, wird im weiteren Verlauf der Arbeit Formulierung (7.44)-(7.46) genutzt, um ein robustes Dienstplanungsproblem zu formulieren.

7.5.2 Lösungsverfahren

Das in Formulierung (7.44) - (7.46) beschriebene Dienstplanungsproblem mit Berücksichtigung von Verspätungspropagation kann mit einem Column Generation Ansatz gelöst werden, in welchem das (beschränkte) Master-Problem als LP-Relaxation mit einem LP-Solver gelöst wird. Die Column Generation Pricing-Probleme zur Erzeugung neuer Spalten bzw. Dienst werden dabei als ressourcenbeschränkte Kürzeste-Wege-Probleme formuliert, die insbesondere die Propagation möglicher Verspätungen abbilden. Ganzzahlige Lösungen können mit Hilfe eines Branch-and-Bound-Verfahrens abgeleitet werden. Die Initialisierung des Column Generation Verfahrens, das Lösen des Master-Problems und der Pricing-Probleme und das Bestimmen einer ganzzahligen Lösung werden im Folgenden genauer dargestellt.

Initialisierung

Eine initiale Dienstmenge kann heuristisch bestimmt werden. Dazu werden die sich aus der Flusslösung ergebenden Aufgaben bzw. Fahrten so auf einzelne Diensten verteilt, dass alle Aufgaben abgedeckt sind und Kurzdienste entstehen, die nur aus einem Dienststück bestehen und keine Pausenzeitregelungen erfüllen müssen (siehe Tabelle 11.1 zur Diensttypenübersicht in Abschnitt 11.1.1 für eine genaue Ausprägung eines solchen Diensttyps). Die Menge der so erstellten (Kurz-)Dienste wird mit entsprechenden Zielfunktionswerten, d.h. der Summe aus den geplanten Kosten und den sich bei Berücksichtigung von Verspätungspropagation in Abhängigkeit der auszuführenden Aufgaben ergebenden Propagationskosten, als initiale Spaltenmenge dem Master-Problem hinzugefügt. Bei einer Nichteinhaltung von Dienstregeln, z.B. einer Unterschreitung der Mindestdauer oder einer Überschreitung der Maximaldauer, wird der betreffende Dienst zusätzlich mit Strafkosten bewertet.

Master-Problem: LP-Relaxation

Das (beschränkte) Master-Problem kann als LP-Relaxation mit einem LP-Solver gelöst werden. Die Dualwerte μ zu Aufgabenabdeckungsbedingungen (7.45) werden genutzt, um in den Pricing-Problemen Spalten bzw. Dienste zu finden, die den aktuellen Zielfunktionswert verbessern können.

Pricing-Problem: Dienstelement-basierte ressourcenbeschränkte kürzeste Wege

Die Pricing-Probleme zur Erstellung von kosteneffizienten und robusten Diensten (R-CSP-CG-P) sind als ressourcenbeschränkte Kürzeste-Wege-Probleme auf Dienstelement-basierten Dienstgenerierungsnetzwerken formuliert – auf den von [Kliewer et al., 2012] und [Amberg, 2014] vorgeschlagenen Netzwerkstrukturen zur Dienstgenerierung. Die Dienstelemente ergeben sich im teilintegrierten Lösungsansatz aus den in der (robusten) Flusslösung des Umlaufplanungsproblems enthaltenen Fahrzeugaktivitäten, die einen Fahrer erfordern. Während der Verknüpfung von Dienstelementen zu Diensten wird dann die Propagation von Verspätungen berücksichtigt.

Zu jeder Kombination aus Diensttyp t und Depot d sei ein Dienstgenerierungsnetzwerk $G_C^{d,t} = (N_C^{d,t}, A_C^{d,t})$ definiert mit der Knotenmenge $N_C^{d,t}$ und der Kantenmenge $A_C^{d,t}$. Die Knotenmenge enthält neben Time-Space-Knoten, die Anfang oder bzw. und Ende eines Dienstelementes beschreiben, einen Quellknoten und einen Senkeknoten. Die Kantenmenge $A_C^{d,t} \supset \widetilde{A}_C^d$ enthält vier verschiedene Kantentypen: *Aufrüst-, Abrüst-, Dienstelement- und Pausenkanten.* Aufrüstkanten verbinden Quell-

knoten mit Dienstelementen und repräsentieren den möglichen Start eines Dienstes. Abrüstkanten verknüpfen analog Dienstelemente mit dem Senkeknoten und stellen das mögliche Ende eines Dienstes dar. Dienstelement-Kanten repräsentieren von einem Fahrer auszuführende Fahrzeugaktivitäten, die aus der Durchführung von Servicefahrten, Leerfahrten, und dem Warten außerhalb des Depots abgeleitet werden. Sie verbinden Time-Space-Knoten miteinander die Start und Ende der entsprechenden Aktivität abbilden. Pausenkanten beschreiben Pausen zwischen zwei Dienststücken. Im Dienstelement-basierten Netzwerk verbinden sie die Knoten aller Dienstelemente miteinander, die potenzielle Enden und Starts von Dienststücken repräsentieren. Welche Knoten verbunden werden ist abhängig von der im zugrunde liegenden Diensttyp definierten minimalen und maximalen Pausenlänge.

Jeder zulässige Pfad von Quell- zu Senkeknoten repräsentiert dann einen zulässigen Dienst. Dabei sind die reduzierten Kosten $\bar{f}_k^{d,t}$ eines Dienstes $k \in K^d$ zu der vom Dienst abgedeckten Kantenmenge $\widetilde{A}_B^d(k)$ beschrieben durch

$$\bar{f}_k^{d,t} = f_k^{d,t} + \alpha_{CS} v_k^{d,t} - \sum_{(i,j) \in \widetilde{A}_B^d(k)} \mu_{ij}^d \tag{7.52}$$

mit den (kantenabhängigen) Dienstkosten $f_k^{d,t}$ und den Dualwerten μ_d sowie den pfadabhängigen Kosten für die Summe der im Dienst propagagierten Verspätungen $v_k^{d,t}$ gemäß Formel 5.1.

Die Zulässigkeitskriterien zur Diensterstellung sind neben den strukturell im Netzwerk abbildbaren Restriktionen, wie z.B. Start- und Endzeitpunkte eines bestimmten Diensttyps, über Ressourcenverbräuche auf den Kanten des Dienstgenerierungsnetzwerkes modellierbar. Tabelle 7.1 fasst die generellen Ressourcenverbräuche für einen Diensttyp t auf den Kanten der Netzwerke zusammen. Zusätzlich zu den beschriebenen Verbräuchen werden primäre Verspätungen auf den Kanten abgebildet.

Mit F^t sind die Fixkosten des Diensttyps t definiert, mit a^t die variablen Arbeitszeitkosten und mit b^t die variablen Pausenzeitkosten pro Minute, wobei l_{ij} die Dauer in Minuten der durch Kante $(i,j) \in A_C^t$ repräsentierten Aktivität bezeichnet. Ferner sind mit $lot(i,j)$ (evtl. pausenfähige) Aufenthaltszeiten am Ende einer Service- oder Leerfahrt beschrieben sowie mit $t(i,j)$ mögliche Transferzeiten zwischen den durch Knoten $i,j \in N_C^t$ repräsentierten Orten. Bei den mit den Kanten assoziierten Zeiten wird unterschieden zwischen der Arbeitszeit AZ, der Lenkzeit LZ, der Dienstzeit DZ und (evtl. pausenfähigen) Wartezeiten WZ. Dienstelement-Kanten sind mit DE abkürzend bezeichnet. Um je Diensttyp eine minimale und maximale Anzahl an Dienststücken sicherzustellen, wird die Ressource *#Dienststücke* verwendet. Damit in den Dienststücken mindestens eine Servicefahrt bedient wird, wird

Typ der Kante $(i,j) \in A^t$	**Kosten**	**AZ**	**LZ**	**DZ**	**WZ**	**#Service**	**#Dh**	**#Dienststücke**
Aufrüsten	$F^t + l_{ij}a^t$	l_{ij}	0	l_{ij}	0	0	0	0
Abrüsten	$l_{ij}a^t$	l_{ij}	0	l_{ij}	0	0	0	1
DE Service	$l_{ij}a^t$	l_{ij}	l_{ij}	l_{ij}	$lot(i,j)$	1	0	0
DE Ausrücken	$l_{ij}a^t$	l_{ij}	l_{ij}	l_{ij}	$lot(i,j)$	0	0	0
DE Einrücken	$l_{ij}a^t$	l_{ij}	l_{ij}	l_{ij}	0	0	0	0
DE Verb.fahrt	$l_{ij}a^t$	l_{ij}	l_{ij}	l_{ij}	$lot(i,j)$	0	1	0
DE Warten	$l_{ij}a^t$	l_{ij}	l_{ij}	l_{ij}	l_{ij}	0	0	0
Pause	$l_{ij}b^t$	0	0	l_{ij}	l_{ij}	0	0	1
Pause+Transfer	$l_{ij}b^t$ $+t(i,j)a^t$	$t(i,j)$	0	l_{ij}	$l_{ij} - t(i,j)$	0	0	1

Tabelle 7.1: Generelle Ressourcenverbräuche auf den Kanten im Dienstelement-Abschnitt basierten Time-Space-Dienstgenerierungsnetzwerk nach [Amberg, 2014]

zusätzlich die Ressource *#Service*(fahrten) abgebildet. Um zu gewährleisten, dass bei Durchführung einer Leerfahrt in den Diensten stets die direkte Verbindung zwischen zwei Orten genutzt wird, wird außerdem die Ressource *#Dh* genutzt, mit der bei der Diensterstellung die Anzahl aktuell durchgeführter Leerfahrten gezählt werden kann.

Die so beschriebenen ressourcenbeschränkten Kürzeste-Wege-Probleme zur Dienstgenerierung werden mit einem Label-Setting Algorithmus ähnlich zu dem mit Algorithmus 7.3.2 vorgestellten gelöst, bei dem explizit eine mögliche Propagation von Verspätungen in die Diensterstellung einbezogen wird. Zur Beschleunigung des Verfahrens kann dabei auf die von [Kliewer et al., 2012] beschriebenen Techniken zur Reduzierung der zu betrachtenden Anzahl Label bzw. (Teil-)Pfade zurückgegriffen werden.

Pufferzeiten die während des Label-Algorithmus eine mögliche Propagation von Verspätungen im erstellten Dienst reduzieren können, sind üblicherweise mit Kanten assoziiert, die Warten, das Durchführen einer Leerfahrt oder einer Pause repräsentieren. Dabei sei darauf hingewiesen, dass auch im Falle einer vermeintlich langen Pause Verspätungen nach dieser weiter propagiert werden können, da die Pufferzeit auf Pausenkanten als Differenz zwischen Länge der Pause und minimaler für den aktuell betrachteten Diensttyp benötigten Pausenlänge bestimmt wird.

Ganzzahlige Lösungen

Werden in den Pricing-Problemen keine Dienste mehr mit negativen reduzierten Kosten gefunden oder die Verbesserung des Zielfunktionswertes ist bezogen auf eine bestimmte Anzahl Iterationen unter einem definierten Schwellenwert, wird die aktuel-

le Spaltenmenge des (beschränkten) Master-Problems genutzt, um eine ganzzahlige Lösung zu finden. Dazu wird Modell R-CSP (Formulierung (7.44)-(7.46)) mit einem MIP-Solver gelöst. Dabei liefert die Lösung des MIPs mittels Branch-and-Bound-Verfahren basierend auf der aktuellen Spaltenmenge ausreichend gute Lösungen, wie die im Rahmen der Arbeit durchgeführten Rechenläufe haben erkennen lassen. Der Gap zwischen dem Lösungswert der LP-Relaxation und des MIPs liegt in der Regel deutlich unter 1%. Daher wird auf eine rechenaufwändige Bestimmung der ganzzahligen Lösung mittels Branch-and-Price-Verfahren verzichtet.

7.6 Verknüpfung der Umlauf- und Dienstebene

Nachdem ein (robuster) Dienstplan mit dem im vorherigen Abschnitt 7.5 beschriebenen Ansatz erstellt wurde, müssen die einzelnen Dienstelemente so zu Umläufen zusammengefasst werden, dass ein zur berechneten Flusslösung gültiger Umlaufplan gebildet wird. Im Gegensatz zur in Abschnitt 7.3 vorgestellten Dekomposition der Flusslösung allein auf Umlaufebene müssen nun bei der Dekomposition Vorgaben aus der Dienstplanung berücksichtigt werden. Insbesondere durch die Dienststücke sind dabei Sequenzen von Aufgaben bzw. Fahrten vorgegeben, die von einem Dienst auf demselben Fahrzeug ausgeführt werden sollen.

Robustheitsaspekte bezüglich der Verknüpfung von Umlauf- und Dienstebene können dann dadurch einbezogen werden, dass aus den Fahrtensequenzen die Umläufe gebildet werden, die eine Propagation von Verspätung minimieren. Die Dekomposition des Lösungsflusses zielt also darauf ab, kompatible Dienststücke aus dem Dienstplan so zu verknüpfen, dass die bestmögliche Robustheit erreicht wird. Da dabei die auszuführenden Aufgaben auf Umlaufebene und die konkreten Dienste bereits vorgegeben sind, werden die Kosten des Ressourceneinsatzes in diesem Schritt nicht mehr verändert.

In den folgenden Abschnitten 7.6.1 und 7.6.2 werden Modellierung und Lösungsverfahren genauer vorgestellt. Dabei ist die Vorgehensweise bei der globalen Flussdekomposition basierend auf der Dienstebene bzw. basierend auf Dienststücken mit der Vorgehensweise der in Abschnitt 7.3.2 beschriebenen zweiten Phase der globalen Flussdekomposition basierend auf Teilumläufen vergleichbar.

7.6.1 Modellierung

Statt wie in Abschnitt 7.3.2 erläutert, Teilumläufe unter Berücksichtigung von Verspätungspropagation zu (Tages-)Umläufen zu verknüpfen, erfolgt die Verknüpfung

nun mit Hilfe der berechneten Dienststücke. Mathematisch kann das entstehende Optimierungsproblem dann wie folgt beschrieben werden.

Sei $S_B^d = \{1, ..., n\}$ die Menge der zur Depot-Fahrzeugtypkombination $d \in \mathcal{D}$ gehörigen Dienststücke. Sei L^d die Menge aller Tagesumläufe und L_{max}^d die aufgrund der Flusslösung maximal erlaubte Anzahl von Fahrzeugumläufen je Depot d. Binäre Entscheidungsvariablen x_l^d mit zugehörigem Zielfunktionskoeffizienten w_l^d zeigen an, ob Umlauf $l \in L^d$ in der Lösung ausgewählt ist oder nicht. Dabei repräsentiert Koeffizient w_l^d die gesamte Verspätungspropagation eines Umlaufes über (mehrere) Dienste, in Abhängigkeit der im Umlauf ausgeführten Dienststücke. Sei überdies $L^d(s)$ die Menge aller Umläufe, die das Dienststück $s \in S_B^d$ enthalten. Dann kann pro Depot bzw. Netzwerkschicht $d \in \mathcal{D}$ analog zur Formulierung GFD-R-VSP II das Problem *zur globalen Flussdekomposition unter Berücksichtigung von Verspätungspropagation basierend auf Dienststücken* (*GFD-R-VSP II Pow*) formuliert werden.

GFD-R-VSP II Pow:

$$\sum_{l \in L^d} x_l^d w_l^d \to \min \tag{7.53}$$

$$\text{s.t.} \sum_{l \in L^d(s)} x_l^d = 1 \qquad \forall s \in S_B^d \tag{7.54}$$

$$\sum_{l \in L^d} x_l^d \leq L_{max}^d \tag{7.55}$$

$$x_l^d \in \{0, 1\} \qquad \forall l \in L^d \tag{7.56}$$

Die Zielfunktion (7.53) minimiert die gesamte Verspätungspropagation aller Umläufe (über alle Dienste). Dabei stellen die Nebenbedingungen (7.54) sicher, dass jedes im Dienstplan vorgesehene Dienststück genau einmal in einem Umlauf enthalten ist. Nebenbedingung (7.55) begrenzt die Anzahl maximal erlaubter Umläufe pro Depot auf die durch die zugrunde liegende Flusslösung vorgegebene Anzahl. Die Nebenbedingungen (7.56) beschreiben den binären Charakter der Entscheidungsvariablen für Umläufe.

7.6.2 Lösungsverfahren

Eine explizite Enumeration aller Verbindungsmöglichkeiten von Dienststücken zu Umläufen führt bereits bei Problemstellungen mit geringer Anzahl zu verplanender Servicefahrten schnell zu Modellen mit einer extrem hohen Anzahl an Variablen. Deshalb wird die Formulierung (7.53) - (7.56) (R-GFD-VSP II Pow) mit einem Column Generation Ansatz gelöst, der in seiner Ausgestaltung nahezu der Phase II der glo-

balen Flussdekomposition in Abschnitt 7.3.2 entspricht. Nachfolgend werden Initialisierung des Verfahrens, das Lösen des Master-Problems und der Pricing-Probleme sowie die Bestimmung einer ganzzahligen Lösung genauer erläutert.

Initialisierung

Als Initiallösung werden die aus dem Dienstplan stammenden Dienststücke mit Hilfe einer Formulierung als Netzwerkflussproblem zu einem zulässigen Umlaufplan zusammengesetzt. In den entsprechenden Netzwerken $G^d = (N^d, A^d)$ je Depot $d \in \mathcal{D}$ besteht die Knotenmenge N^d aus den auszuführenden Dienststücken sowie einem Quell- und einem Senkeknoten. Es werden dann diejenigen Dienststücke durch Kanten miteinander verbunden, die zeitlich und örtlich zueinander kompatibel sind. Dabei müssen folgende Regeln beachtet werden:

- Dienststücke, die mit einer Einrückfahrt enden, werden mit Dienststücken verbunden, die mit einer Ausrückfahrt beginnen.
- Dienststücke, die mit einer Ausrückfahrt enden, werden mit allen Dienststücken verbunden, außer mit denen, die mit einer Ausrückfahrt beginnen.
- Dienststücke können auch außerhalb eines Depots beginnen und enden – an jeder Station bzw. Haltestelle, an der ein Fahrer- bzw. Fahrzeugtausch möglich ist. Deshalb werden alle Dienststücke, die außerhalb eines Depots enden, mit allen Dienststücken verbunden, die nicht mit einer Ausrückfahrt beginnen.

Außerdem verbinden Kanten den Quellknoten mit allen Knoten, die ein Dienststück repräsentieren, das mit einer Ausrückfahrt startet. Ebenso werden alle Knoten mit dem Senkeknoten verbunden, die ein Dienststück darstellen, dass mit einer Einrückfahrt endet. Des weiteren verknüpft eine Zirkulationsflusskante Senke- und Quellknoten miteinander.

Ganzzahlige Entscheidungsvariablen y_{ij}^d für jede Kante $(i, j) \in A^d$ zeigen die Menge der Flusseinheiten an, die in der Lösung über die entsprechende Kante fließen. Um bereits Robustheitsaspekte beim Lösen zu berücksichtigen und so eine schnellere Konvergenz des Verfahrens zu ermöglichen, werden den Kanten dabei Kosten $\widehat{w}_{ij}^d$ zugewiesen, die den jeweiligen propagierten Verspätungen eines Dienststückes bei nachfolgender Ausführung des verbundenen Dienststückes entsprechen. Im Gegensatz zu einer vollständigen Berücksichtigung von Verspätungspropagation werden bei der Initialisierung also zunächst Propagationseffekte unterschätzt, da nur die mögliche Propagation zwischen zwei aufeinanderfolgenden Dienststücken betrachtet wird und keine Übertragung der Verspätung auf weitere Verbindungen erfolgt. Eine korrekte

Berechnung des Propagationswertes wird dann vor Hinzufügen der entstandenen initialen Umläufe ins Master-Problem durchgeführt. Die maximale Flusskapazität u_{ij}^d auf einer Kante ist für die Zirkulationsflusskante mit der Fahrzeuganzahl L_{max}^d definiert und für alle anderen Kanten mit 1. Das Netzwerkflussproblem mit einfacher Berücksichtigung von Verspätungspropagation zur Erstellung eines initialen Umlaufplans kann dann wie folgt formuliert werden.

VSP Pow NWF:

$$\sum_{(i,j)\in A^d} y_{ij}^d \widehat{w}_{ij}^d \to \min \tag{7.57}$$

$$\text{s.t.} \sum_{\{j:(j,i)\in A^d\}} y_{ji}^d - \sum_{\{j:(i,j)\in A^d\}} y_{ij}^d = 0 \qquad \forall i \in N^d \tag{7.58}$$

$$\sum_{\{j:(i,j)\in A^d\}} y_{ij}^d = 1 \qquad \forall i \in N^d \tag{7.59}$$

$$0 \leq y_{ij}^d \leq u_{ij}^d,\ y_{ij}^d \in \mathbb{N} \qquad \forall (i,j) \in A^d \tag{7.60}$$

Mit der Zielfunktion (7.57) wird die Verspätungspropagation zwischen direkt aufeinander folgenden Dienststücken in der Netzwerkflusslösung minimiert. Dabei stellen die Bedingungen (7.59) sicher, dass jedes Dienststück genau einmal in der Flusslösung besucht wird. Die Flusserhaltungsbedingungen (7.58) und die Beschränkung der Kapazitäten (7.60) stellen einen gültigen Netzwerkfluss sicher. Da sich durch die explizite Formulierung der möglichen Verbindungen zwischen den Dienststücken eindeutige Pfade von Quelle zu Senke ergeben, können die entstandenen Umläufe direkt aus der Netzwerkflusslösung abgeleitet werden. Die gefundenen Umläufe werden dann mit ihrem tatsächlichen Verspätungspropagationswert gemäß Formel 5.1 basierend auf den in den Umläufen ausgeführten Fahrten dem Master-Problem als initiale Spaltenmenge hinzugefügt.

Master-Problem: LP-Relaxation

Das (beschränkte) Master-Problem wird als LP-Relaxation mit einem LP-Solver gelöst. Mit Hilfe der Dualwerte μ und ν der Dienststückabdeckungsbedingungen (7.54) und der Fahrzeuganzahlrestriktion (7.55) können in den Pricing-Problemen Spalten bzw. Umläufe gesucht werden, die den aktuellen Zielfunktionswert verbessern können.

Pricing-Problem: Dienststück-basierte ressourcenbeschränkte kürzeste Wege

Für die Pricing-Probleme werden ähnlich der in Abschnitt 7.3.2 beschriebenen Phase II zur globalen Flussdekomposition Umlaufgenerierungsnetzwerke aufgestellt. Statt wie bei der globalen Flussdekomposition auf Basis von Teilumläufen werden die Netzwerke nun auf Basis von Dienststücken modelliert. Dabei werden Dienststücke als Kanten dargestellt, die zusammengehörige Start- und Endknoten miteinander verbinden. Die Knoten repräsentieren jeweils Start- und Endzeitpunkt eines Dienststückes. Zusätzliche Kanten werden eingefügt, die potenzielle kompatible Verbindungen zu anderen Dienststücken repräsentieren. Zwei Dienststücke werden dann miteinander verbunden, wenn der Endzeitpunkt des ersten Dienststückes früher bzw. gleich dem Startzeitpunkt des zweiten Dienststückes ist und sie örtlich zueinander kompatibel sind. Um gültige Umläufe zu erstellen, werden daher die kompatiblen Dienststücke unter den Regeln verbunden werden, die oben bei der Beschreibung der Initialisierung des zugrunde liegenden Column Generation Verfahrens formuliert sind. Analog zum Vorgehen in Phase II werden die vorgestellten Umlaufgenerierungsnetzwerke mit einem Quell- und einem Senkeknoten erweitert sowie mit entsprechenden Verbindungskanten vom Quellknoten zu allen möglichen Knoten, die den Start eines Umlaufs aus dem Depot repräsentieren, bzw. mit Verbindungskanten von allen Knoten, die das Ende eines Umlaufs im Depot darstellen, zum Senkeknoten. So können die Pricing-Probleme dann als ressourcenbeschränkte Kürzeste-Wege-Probleme formuliert und mit dem (angepassten) Label Setting Algorithmus 7.3.2 gelöst werden. Dabei werden dann analog je Label bzw. (Teil-)Pfad die kumulierten reduzierten Kosten, die kumulierte Summe der Verspätungspropagation auf den Servicefahrten und die aktuelle Verspätungspropagation berücksichtigt.

Ganzzahlige Lösungen

Wird das Column Generation Verfahren beendet, nachdem keine Umläufe mehr mit negativen reduzierten Kosten gefunden wurden bzw. die Verbesserung des Zielfunktionswertes in einer bestimmten Anzahl letzter Iterationen unter einem definierten Schwellenwert liegt, können je Netzwerkschicht mit der ermittelten Spaltenmenge ganzzahlige Lösungen durch einem MIP-Solver bestimmt werden. Wie bereits für die Dienstebene beschrieben (vgl. Abschnitt 7.5.2), wird diese heuristische Vorgehensweise genutzt und auf ein rechenaufwändiges Branch-and-Price-Verfahren verzichtet, da der Gap zwischen LP-Relaxation und der anschließend per MIP-Solver bestimmten Lösung gering ist. Es ergeben sich dann Ressourceneinsatzpläne mit der geringsten Verspätungspropagation von Umläufen bei Berücksichtigung von Diensten, bei

denen der Umlaufplan kostengleich zur ursprünglich berechneten (kostenoptimalen) Flusslösung nach Abschnitt 7.4 ist.

7.7 Zusammenfassung und Implikationen

In diesem Kapitel wurde ein Lösungsansatz entwickelt, der in der Lage ist, Robustheitsaspekte in die Planung auf Umlauf- und Dienstebene einzubeziehen. Dabei kann die Robustheit der Pläne zusätzlich zur Kosteneffizienz sowohl in den einzelnen Planungsebenen beeinflusst werden, als auch durch wechselseitige Beziehungen zwischen Umlauf- und Dienstplanung.

Eine Besonderheit des Lösungsansatzes besteht darin, dass eine mögliche Propagation von Verspätungen des Ressourceneinsatzes direkt bei der Berechnung von (kosteneffizienten und) robusten Umläufen und Diensten berücksichtigt werden kann. Dazu wurden mehrkriterielle Entscheidungsprobleme modelliert, in welchen zusätzlich zu Kosten auch die Propagation von Verspätungen minimiert werden soll. Gelöst werden diese Entscheidungsprobleme im Rahmen von Column Generation Verfahren. Beim Lösen der als ressourcenbeschränkte Kürzeste-Wege-Probleme formulierten Pricing-Probleme zur Generierung von kosteneffizienten und robusten Umläufen und bzw. oder Diensten wird dann die Verspätungspropagation der Ressourceneinsätze ermittelt. Dabei werden Umläufe zur Reduktion der zu lösenden Problemgröße in einem (heuristischen) zwei-phasigen Pricing-Schema berechnet, bei dem zunächst Umläufblöcke bzw. Teilumläufe generiert werden und dann komplette Fahrzeugumläufe. Dienste werden basierend auf den auszuführenden Dienstelementen bestimmt.

Um bei der Planung von Umläufen und Diensten weiterhin den Vorteil einer TSN-Modellierung des zugrunde liegenden Umlaufplanungsproblems ausnutzen zu können, wurde ein Konzept zur Erstellung robuster Flusslösungen entwickelt. So können basierend auf Bündeln von kosteneffizienten und robusten Umlaufplänen, kosteneffiziente und robuste Dienste ermittelt werden. Dazu wurde eine netzwerkflussbasierte Formulierung des Umlaufplanungsproblems mit einer pfadbasierten Formulierung kombiniert.

Aufgrund der Konzeption, bei der mehrere kosten- und aufgabengleiche Ressourceneinsatzpläne betrachtet werden können, eignet sich der Ansatz um der Forschungsfrage F2 nachzugehen und herauszufinden, wie viel Robustheit in kostenoptimalen Plänen realisiert werden kann.

Die Möglichkeiten für eine (robuste) Dienstplanung sind im teilintegrierten Lösungsansatz allerdings auch abhängig von der (Fluss-)Lösung des Umlaufplanungsproblems – wenn auch im Vergleich zu einer sequenziellen Planung deutlich mehr Freiheitsgrade

für die Dienstplanung zur Verfügung stehen. Um Bedürfnisse der Dienstplanung noch stärker zu berücksichtigen, d.h. um eine durch Dienste verursachte Verspätungspropagation noch stärker in die Ressourceneinsatzplanung einbeziehen zu können, wird im nächsten Kapitel 8 ein vollständig integrierter Ansatz vorgeschlagen, bei welchem Umlauf- und Dienstplan simultan erstellt werden.

Kapitel 8

Robuste Effizienz durch integrierte Planung

In diesem Kapitel wird ein Lösungsansatz vorgeschlagen, der bei der Planung simultan die Umlauf- und Dienstebene betrachtet und dabei sowohl geplante Kosten als auch Robustheitsaspekte in Form von potenzieller Verspätungspropagation bei Fahrzeugen und Fahrern einbezieht. Dieser *integrierte* Ansatz basiert auf dem von [Steinzen et al., 2010] vorgeschlagenen Ansatz für reine kosteneffiziente Umlauf- und Dienstplanung und dem im vorangegangenen Kapitel 7 entwickelten Ansatz zur robusten und kosteneffizienten teilintegrierten Planung. Im Vergleich zu einem sequenziellen, teilintegrierten oder iterativen Verfahren kann durch die gleichrangige Behandlung von Umlauf- und Dienstplanungsproblem der Ressourceneinsatz bezüglich der Ziele Kosteneffizienz und Robustheit besser aufeinander abgestimmt werden. Denn es bieten sich nicht nur mehr Freiheitsgrade für die Dienstplanung, sondern es wird auch eine Berücksichtigung von gemeinsamen ressourcenübergreifenden Robustheitsaspekten ermöglicht. Abbildung 8.1 zeigt die gemeinsame, gleichrangige Kopplung der verschiedenen Planungsprobleme schematisch als Bestandteile der in diesem Kapitel entwickelten integrierten Modellformulierung.

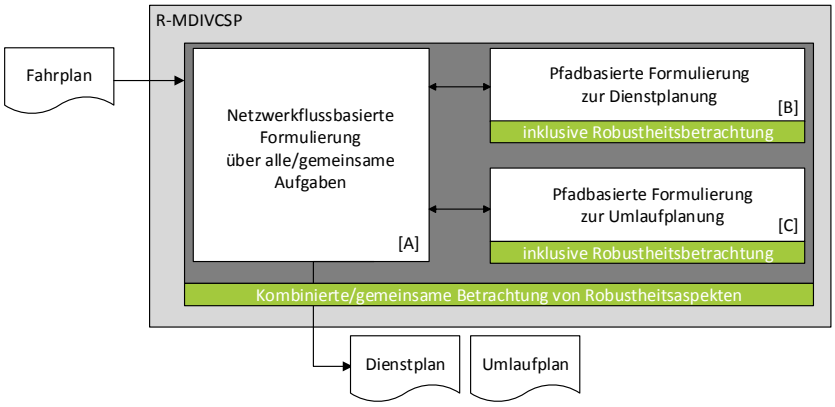

Abbildung 8.1: Verknüpfung der Planungsprobleme zur integrierten Umlauf- und Dienstplanung mit Einbezug von Robustheitsaspekten in die Planung

Grundidee

Das integrierte Modell koppelt drei wesentliche Formulierungen miteinander: Eine netzwerkflussbasierte Formulierung (A) basierend auf dem in Abschnitt 6.1 vorgestelltem Netzwerkmodell, welche die gleichzeitige Betrachtung von (kostengleichen) Bündeln von Ressourceneinsatzplänen ermöglicht und so die im vorangegangenen Kapitel beschriebenen Vorteile beim Lösen bietet, eine pfadbasierte Formulierung zur Dienstplanung (B) und eine pfadbasierte Formulierung zur Umlaufplanung (C). Formulierung (A) sichert die Verknüpfung der pfadbasierten Formulierungen. Die beiden pfadbasierten Formulierungen ermöglichen es wiederum, aufgabensequenzenbasierte Effekte wie Verspätungspropagation auf Dienst- bzw. Umlaufebene korrekt abzubilden.

Robustheitsheitsaspekte werden dann während der Planung gleichzeitig und gleichrangig auf den beiden Planungsebenen betrachet. Insbesondere sollen so eine potenzielle Verspätungspropagation durch beide Planungsebenen und mögliche kombinierte Effekte bei der Planerstellung berücksichtigt werden. Die durch das integrierte Modell abgebildete gemeinsame Berechnung von Umläufen und Diensten kann dann vor allem auf beiden Ressourcenebenen gleichlaufende Verspätungspropagationseffekte schaffen, die höchstenfalls innerhalb einzelner gemeinsamer Aufgabensequenzen eines Umlaufes und Dienstes entstehen und nicht zum Aufschaukeln von Verspätungspropagation zwischen den Ressourcenebenen führen. Umläufe und Dienste sind also idealerweise so ausgeprägt, dass keinerlei Verspätungspropagation von einer auf die andere Ressourcenebene weitergegeben wird, die entstehenden Ressourceneinsatzpläne aber weiterhin kosteneffizient geplant sind.

Aufbau des Kapitels

Nach der Definition des zugrunde liegenden Planungsproblems in Abschnitt 8.1wird in Abschnitt 8.2 die mathematische Modellierung zur robusten integrierten Umlauf- und Dienstplanung erläutert. Anschließend stellt Abschnitt 8.3 das Lösungsverfahren vor, wobei insbesondere auf das Einbeziehen potenzieller Verspätungspropagation in die Planung eingegangen wird. Eine Zusammenfassung und Implikationen werden am Schluss des Kapitels in Abschnitt 8.4 gegeben.

8.1 Problemdefinition R-MDIVCSP

Das *integrierte (Mehrdepot-)Umlauf- und Dienstplanungsproblem mit der Zielsetzung nach robuster Effizienz* (engl. *robust integrated multiple-depot vehicle and crew sche-*

duling problem, R-MDIVCSP) kann wie folgt definiert werden: Gesucht wird ein kosteneffizienter und in Bezug auf (betrachtete) Verspätungsszenarien möglichst robuster Umlauf- und Dienstplan für eine gegebene Menge an Servicefahrten, Ablösepunkten, Leerfahrten und Möglichkeiten zum Transfer zwischen (End-)Haltestellen sowie verschiedenen Depots. Umlaufplan und Dienstplan sollen dabei zulässig und zueinander kompatibel sein, d.h. jede Servicefahrt soll von genau einem Umlauf und einem Dienst ausgeführt werden und jede zusätzlich auszuführende Fahrzeugaktivität von genau einem Dienst, während die Dienste keine Fahrzeugaktivitäten enthalten, die nicht Bestandteil der Umläufe sind.

Wie beim teilintegrierten Ansatz soll dabei die Verspätungspropagation nach der in Abschnitt 5.2.1 beschriebenen Form als Nicht-Robustheitsmaß zur Bewertung der Robustheit der Umlauf- und Dienstpläne genutzt werden.

8.2 Modellierung des R-MDIVCSP

Die mathematische Formulierung zur Minimierung von geplanten Kosten und Verspätungspropagation bei der gleichzeitgen Erstellung von Umläufen und Diensten beruht hauptsächlich auf der in Abschnitt 6.1 beschriebenen TSN-Struktur zur Abbildung des Umlaufplanungsproblems. Die Formulierung kombiniert eine Mehrgüter-Fluss-Formulierung für eine (kosteneffiziente) Umlaufplanung mit einer zusätzlichen Set-Partitioning Formulierung für die Umlaufplanung zur Minimierung von Verspätungspropagation und einer Set-Partitioning Formulierung zur kosteneffizienten und robusten Dienstplanung.

Sei $\mathcal{D} = \{1, 2, ..., m\}$ die Menge von Depot-Fahrzeugtypkombinationen und sei $\mathcal{T} = \{1, 2, ..., n\}$ die Menge aller Servicefahrten. Für jedes Depot $d \in \mathcal{D}$ wird ein Umlaufplanungsnetzwerk $G^d = (N^d, A^d)$ definiert, wie in Abschnitt 7.3.1 beschrieben. Darüber hinaus beschreibe $\widetilde{A}^d \subset A^d$ die Menge derjenigen Kanten, die eine Fahrzeugaktivität repräsentieren, für deren Ausführung ein Dienst notwendig ist. Zusätzlich definiert $A^d(t) : \mathcal{T} \to A^d$ eine Funktion, die die zu Fahrt $t \in \mathcal{T}$ und Depot $d \in \mathcal{D}$ zugehörige Servicefahrtkante liefert. Falls die Fahrt t nicht von Depot d ausgeführt werden kann, sei $A^d(t) = \emptyset$.

Jeder Kante $(i, j) \in A^d$ werden fixe oder variable Fahrzeugkosten c^d_{ij} zugewiesen. Fixe Fahrzeugkosten werden mit den Zirkulationsflusskanten assoziiert, um den Gebrauch eines Fahrzeugs aus Depot d darzustellen. Variable Fahrzeugkosten werden allen anderen Kanten in Abhängigkeit der Fahrt- und Wartezeit der zuhörigen Fahrzeugaktivität zugewiesen.

Im Allgemeinen wird für jede Servicefahrt, wie auch für Ausrück- und Einrückfahrtkanten die maximale Flusskapazität u_{ij}^d der entsprechenden Kante auf 1 gesetzt. Falls im TSN-Modell allerdings mehr als eine Servicefahrt an einem Time-Space-Knoten startet bzw. endet, wird die Flusskapazität der zugehörigen Ausrückfahrtkante bzw. Einrückfahrtkante auf die entsprechende Anzahl der beginnenden bzw. endenden Sercivefahrten gesetzt. Für alle anderen Kanten beträgt die maximale Flusskapazität die maximale Anzahl verfügbarer Fahrzeuge u^d des entsprechenden Depots d.

Es werden zwei Arten von Entscheidungsvariablen genutzt: Flussvariablen und pfadbasierte Variablen. Ganzzahlige Flussvariablen y_{ij}^d werden mit jeder Kante $(i,j) \in A^d$ assoziiert und geben an, in welchem Ausmaß die durch Kante (i,j) repräsentierte Fahrzeugaktivität in der Lösung verwendet wird.

Binäre Entscheidungsvariablen x_k^d geben an, ob Dienst $k \in K^d$ in der Lösung aus der Menge K^d der Dienste ausgewählt ist, welche von Depot $d \in \mathcal{D}$ aus durchgeführt werden können, oder nicht. Jeder Variable x_k^d sind dabei geplante Dienstkosten f_k^d sowie mit Verspätungspropagation assoziierte Kosten v_k^d basierend auf den ausgeführten Aufgaben zugewiesen. Die Verspätungsspropagation wird berechnet wie in Formel 5.1 beschrieben. Ferner beschreibe $K^d(i,j) \subset K^d$ die Menge von Diensten, welche die durch Kante $(i,j) \in \widetilde{A}^d$ repräsentierte Aufgabe bedienen.

Binäre Entscheidungsvariablen z_l^d mit $l \in L^d$ und zugehörigen Kosten w_l^d geben an, ob Umlauf l für Depot d ausgewählt wird oder nicht. Dabei stellt w_l^d in Abhängigkeit der von Umlauf l ausgeführten Fahrten die mit Verspätungspropagation assoziierten Kosten des Umlaufs dar. Die Verspätungsspropagation wird gemäß Formel 5.1 berechnet. Analog zum Dienstplanungsteil beschreibe $L^d(i,j) \subset L^d$ die Menge von Umläufen, welche die durch Kante $(i,j) \in A^d$ repräsentierte Fahrt abdecken.

Da die zugrunde liegende Flussformulierung für den Umlaufplanungsteil des Modells es nicht erlaubt, eine korrekte Propagation von Verspätungen zu berücksichtigen, dient die Menge der Variablen z als Art Hilfsstruktur. Dadurch wird es ermöglicht, die exakte Berechnung von Verspätungspropagation innerhalb von Umläufen mit dem beibehaltenen Vorteil der zugrunde liegenden Flussformulierung zu kombinieren. Das *robuste integrierte Umlauf- und Dienstplanungsproblem mit mehreren Depots und Berücksichtigung möglicher Verspätungspropagation* (engl. *robust integrated multiple-depot vehicle and crew scheduling problem, R-MDIVCSP*) kann dann wie folgt formuliert werden.

R-MDIVCSP:

$$\begin{aligned}
&\sum_{d\in\mathcal{D}}\sum_{(i,j)\in A^d} y_{ij}^d c_{ij}^d + \sum_{d\in\mathcal{D}}\sum_{k\in K^d} x_k^d f_k^d \\
&+\alpha_{vs}\sum_{d\in\mathcal{D}}\sum_{l\in L^d} z_l^d w_l^d + \alpha_{cs}\sum_{d\in\mathcal{D}}\sum_{k\in K^d} x_k^d v_k^d \to \min
\end{aligned} \tag{8.1}$$

$$\text{s.t.} \sum_{d\in\mathcal{D}}\sum_{(i,j)\in A^d(t)} y_{ij}^d = 1 \qquad \forall t\in\mathcal{T} \tag{8.2}$$

$$\sum_{\{j:(j,i)\in A^d\}} y_{ji}^d - \sum_{\{j:(i,j)\in A^d\}} y_{ij}^d = 0 \qquad \forall d\in\mathcal{D}, \forall i\in N^d \tag{8.3}$$

$$\sum_{k\in K^d(i,j)} x_k^d - y_{ij}^d = 0 \qquad \forall d\in\mathcal{D}, \forall (i,j)\in \widetilde{A}^d \tag{8.4}$$

$$\sum_{l\in L^d(i,j)} z_l^d - y_{ij}^d = 0 \qquad \forall d\in\mathcal{D}, \forall (i,j)\in A^d \tag{8.5}$$

$$0\le y_{ij}^d \le u_{ij}^d,\ y_{ij}^d\in\mathbb{N} \qquad \forall d\in\mathcal{D}, \forall (i,j)\in A^d \tag{8.6}$$

$$x_k^d\in\{0,1\} \quad \forall d\in\mathcal{D}, \forall k\in K^d \tag{8.7}$$

$$z_l^d\in\{0,1\} \quad \forall d\in\mathcal{D}, \forall l\in L^d \tag{8.8}$$

Mit der Zielfunktion (8.1) werden Fahrzeugkosten und Dienstkosten minimiert. Zusätzlich werden in Abhängigkeit des gewählten Penalty-Parameters α_{vs} die Propagation von Verspätungen innerhalb der Umläufe minimiert, sowie in Abhängigkeit des Penalty-Parameters α_{cs} die Propagation von Verspätungen innerhalb von Diensten. So erlaubt die einfache Formulierung als gewichtete Summe in der Zielfunktion die Skalierung zwischen der Minimierung von geplanten Kosten und Kosten assoziiert mit Verspätungspropagation. Die Restriktionen (8.2)-(8.3) entsprechen einer Mehrgüter-Fluss-Formulierung und stellen einen gültigen Umlaufplan sicher. Die Fahrtenabdeckungsbedingungen (8.2) gewährleisten dabei, dass jede Fahrt genau einmal von Fahrzeug aus einem Depot bedient wird. Nebenbedingungen (8.3) garantieren Flusserhaltung in jedem Netzwerkknoten. Die Verknüpfung der flussbasierten Umlaufplanung mit der pfadbasierten Dienstplanung erfolgt durch die Kopplungsbedingungen (8.4), indem für jede Kante (außerhalb eines Depots), auf der ein Fahrzeugfluss fließt, das entsprechende Dienstelement so oft Bestandteil von Diensten ist, wie durch den Fahrzeugfluss vorgegeben. Gleiches gilt für die Nebenbedingungen (8.5), die die flussbasierte mit der pfadbasierten Umlaufformulierung koppeln. Nebenbedingungen (8.6) gewähren, dass der maximale Fluss auf den Kanten nicht überschritten wird. Die Nebenbedingungen (8.7) und (8.8) beschreiben den binären Charakter der Entscheidungsvariablen für Dienste und Umläufe.

8.3 Lösungsverfahren

In diesem Abschnitt wird ein Verfahren zur Lösung der im vorangegangen Abschnitt vorgestellten Formulierung (8.1) - (8.8) zur robusten integrierten Umlauf- und Dienstplanung entwickelt. Aufgrund der Problemkomplexität der Formulierung, kombiniert das Lösungsverfahren Column Generation mit Lagrange Relaxation wie es für eine integrierte Ressourceneinsatzplanung von [Freling, 1997], [Huisman et al., 2005] und [Steinzen et al., 2010] vorgeschlagen wird. Der Lösungsansatz von [Steinzen et al., 2010] wird dazu um die Berücksichtigung von Robustheitsaspekten erweitert, d.h. um die Einbeziehung von Verspätungspropagation in die Planung. Während des Verfahrens werden auf Umlauf- und Dienstebene dynamisch mögliche Verspätungspropagationen berechnet, so dass für Umläufe und Dienste sowohl geplante Kosten als auch mit Verspätungspropagation assoziierte Kosten berücksichtigt werden.

Für eine algorithmische Zusammenfassung des Vorgehens in Anlehnung an [Steinzen et al., 2010] sei auf Algorithmus 8.3.1 verwiesen. Abbildung 8.2 gibt einen detaillierten schematischen Überblick über das Verfahren.

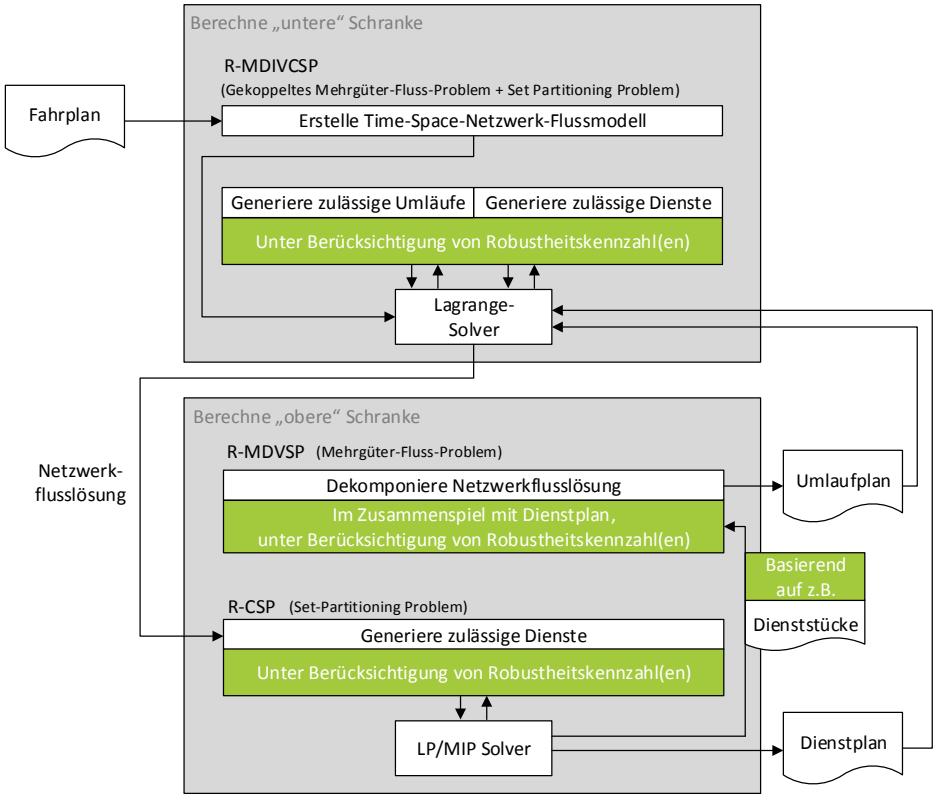

Abbildung 8.2: Integrierte Umlauf- und Dienstplanung mit Berücksichtigung der Zielsetzungen Kosteneffizienz und Robustheit orientiert an [Amberg, 2014]

Im Wesentlichen umfasst das integrierte Lösungsverfahren zwei Phasen. In der ersten Phase (Schritte 2 bis 5) wird die Formulierung (8.1) - (8.8) zur Reduktion der Problemkomplexität Lagrange-relaxiert und eine möglichst gute Approximation des Zielfunktionswertes berechnet. In die Berechnung geht dabei insbesondere die Erstellung guter, d.h. kosteneffizienter und robuster, (zueinander) passender Dienste und Umläufe ein. Beim Lösen der entsprechenden Pricing-Probleme werden ähnlich zu [Dunbar et al., 2012] im Airline-Bereich die Informationen über mögliche Verspätungspropagation für sowohl die Berechnung der tatsächlichen Kosten der Spalten als auch zur Auswahl der Spalten während der Pricing-Schritte eingesetzt. Darüber hinaus wird während des Verfahres die Umlaufplanungsbündel repräsentierende Flusslösung des als Netzwerkflussproblem formulierten Umlaufplanungsteils mit Hilfe der errechneten Dienste und Umläufe so angepasst, dass die Zielsetzung nach Kosteneffizienz und bzw. oder Robustheit auf allen Planungsebenen erreicht werden kann. In der zweiten Phase des Verfahrens (Schritt 6) werden dann ausgehend von solchen angepassten Flusslösungen zulässige ganzzahlige Lösungen, also zulässige und wechselseitig kompatible Umlauf- und Dienstpläne ermittelt. Dazu wird, wie von [Huisman et al., 2005] bzw. [Steinzen et al., 2010] für integrierte Umlauf- und Dienstplanungsprobleme vorgeschlagen, eine Lagrange-Heuristik genutzt.

Die Initialisierung des Verfahrens wird in Abschnitt 8.3.1 kurz erläutert. Abschnitt 8.3.2 beschreibt die Lagrange-Relaxation des Master-Problems und die zur Bestimmung dualer Informationen über die aktuelle Spaltenmenge eingesetzte Lösungsmethode. Wie mit den Dualwerten in den Pricing-Problemen Spalten erzeugt werden, die negative reduzierte Kosten besitzen und so im Rahmen des Column Generation Ansatzes zum Master-Problem hinzugefügt werden können, fasst Abschnitt 8.3.3 zusammen. Auf die Bestimmung zulässiger Umlauf- und Dienstpläne wird in Abschnitt 8.3.4 eingegangen.

8.3.1 Initialisierung

Eine zulässige Initiallösung für das Verfahren wird durch eine sequenzielle oder eine teilintegrierte Planung bestimmt, wie in Abschnitt 6.3 bzw. in Kapitel 7 über teilintegrierte Planung erläutert. Die für den Umlaufplanungsteil (R-)VSP bzw. für den Dienstplanungsteil (R-)CSP berechneten Spalten dienen als initiale Spaltenmenge des vollständig integrierten Verfahrens.

Algorithmus 8.3.1 : Lösungsverfahren für R-MDIVCSP

(Schritt 1) **Initialisierung**
Löse MD-(R-)VSP und anschließend
(R-)CSP für jede Depot-Fahrzeugtypkombination
mit Hilfe des SEQ oder PINT Lösungsschemas.
Nimm die (R-)VSP- und die (R-)CSP-Lösung als initiale
Spaltenmenge $L' \cup K'$.
Setze $t = 0$.

(Schritt 2) **Löse Beschränktes Master Problem**
Löse ein Lagrange-Dual-Problem mit der aktuellen Spaltenmenge
$L' \cup K'$.
Speichere die untere Schranke für die aktuelle Spaltenmenge sowie
die dualen Informationen.

(Schritt 3) **Löse Pricing-Probleme**
Generiere neue Spalten bzw. Fahrzeugumläufe L'' mit negativen
reduzierten Kosten mit Hilfe von Modell R-VSP-CG-P I und II.
Generiere neue Spalten bzw. Dienste K'' mit negativen reduzierten
Kosten mit Hilfe von Modell R-CSP-CG-P.
Falls $|L'' = 0|$ und $|K'' = 0|$, gehe zu Schritt 5.

(Schritt 4) **Spaltenmanagement**
Füge neue Spalten zum beschränkten Master Problem hinzu:
$L' := L' \cup L''$ und $K' := K' \cup K''$.
Lösche Spalten mit hohen positiven reduzierten Kosten von L' (K')
falls $|L'|$ ($|K'|$) groß und über einem definierten Schwellenwert ist.

(Schritt 5) **Prüfe Abbruchbedingungen**
Falls eine der folgenden Abbruchbedingungen gilt:
- max. Anzahl Iterationen erreicht: $t \geq t_{max}$
- keine signifikante Verbesserung der unteren Schranke während
einer bestimmten Anzahl letzter Iterationen,
gehe zu Schritt 6;
ansonsten setze $t = t + 1$ und gehe zu Schritt 2.

(Schritt 6) **Bestimme eine zulässige Lösung**
Nutze eine Lagrange-Heuristik, um einen zulässigen, kompatiblen
Umlauf- und Dienstplan zu erstellen

8.3.2 Master-Problem: Lagrange-Relaxation

Die Formulierung (8.1) - (8.8) zur integrierten Umlauf- und Dienstplanung mit den Zielsetzungen von sowohl Minimierung der geplanten Kosten als auch von Minimierung der potenziellen Verspätungspropagation von Umläufen und Diensten stellt ein schwer lösbares Optimierungsproblem dar. Als Lösungsverfahren wird daher eine Kombination aus Column Generation und Lagrange-Relaxation eingesetzt, bei der das Problem zunächst vereinfacht werden kann. Das dabei entstehende Lagrange-Duale-Problem wird approximativ mit einem Subgradientenverfahren gelöst. So können jeweils die (optimalen) dualen Lösungen – mit Hilfe berechneter Lagrange-Multiplikatoren – approximiert und untere Schranken für das (beschränkte) Master Problem formuliert werden.

Aufgrund der Kopplungsbedingungen (8.4) und (8.5) sowie der Fahrtenabdeckungsbedingungen (8.2) ist Modell R-MDIVCSP schwer lösbar. Die entsprechenden Restriktionen werden daher Lagrange-relaxiert: Durch die Relaxierung der Kopplungsbedingungen entfällt zunächst die restriktive simultane Betrachtung von Umlauf- und Dienstplanung. Die Relaxierung der Fahrtenabdeckungsbedinungen löst die ursprünglich $\mathcal{NP}$-schwere Formulierung zur Mehrdepot-Umlaufplanung auf.

Bei der Lösung des Modells sind dann drei Arten Unterprobleme zu berücksichtigen: (1) Eindepot-Umlaufplanungsprobleme zur Minimierung von geplanten Kosten. (2) Je Depot ein Umlaufplanungsproblem zur Minimierung von mit Verspätungspropagation assoziierten Kosten. (3) Je Depot ein Dienstplanungsproblem zur Minimierung von geplanten Kosten und von mit Verspätungspropagation assoziierten Kosten. Diese Unterprobleme sind im Lösungsprozess miteinander verbunden, indem zueinander unpassende Umlauf- und Dienstplanlösungen in der Zielfunktion bestraft werden. Durch die Reduzierung der Komplexität ergibt sich ein Laufzeitvorteil: Die Eindepot-Umlaufplanungsprobleme sind in polynomieller Laufzeit lösbar. Für die Umlaufplanungs-Unterprobleme zur Minimierung von Verspätungspropagation sowie für die Dienstplanungs-Unterprobleme können effizient Lösungen berechnet werden, indem die Variablen z und x Variablen mit negativen reduzierten Kosten ausgewählt werden.

Die Zielfunktion lautet dann:

$$\begin{aligned}\min \quad & \sum_{d\in\mathcal{D}}\sum_{(i,j)\in A^d} y_{ij}^d c_{ij}^d + \sum_{d\in\mathcal{D}}\sum_{k\in K} x_k^d f_k^d + \alpha_{vs}\sum_{d\in\mathcal{D}}\sum_{l\in L^d} z_l^d w_l^d + \alpha_{cs}\sum_{d\in\mathcal{D}}\sum_{k\in K^d} x_k^d v_k^d \\ & + \sum_{d\in\mathcal{D}}\sum_{(i,j)\in\widetilde{A}^d} \mu_{ij}^d \left(\sum_{d\in\mathcal{D}} y_{ij}^d - \sum_{k\in K^d(i,j)} x_k^d\right) + \sum_{d\in\mathcal{D}}\sum_{(i,j)\in A^d} \nu_{ij}^d \left(\sum_{d\in\mathcal{D}} y_{ij}^d - \sum_{l\in L^d(i,j)} z_l^d\right) \\ & + \sum_{t\in\mathcal{T}} \pi_t \left(1 - \sum_{d\in\mathcal{D}}\sum_{(i,j)\in A^d(t)} y_{ij}^d\right)\end{aligned} \tag{8.9}$$

Das Lagrange-Unterproblem kann daher geschrieben werden als:

$$\Phi(\mu,\pi,\nu) = \Phi_y(\mu,\pi,\nu) + \Phi_z(\nu) + \Phi_x(\mu) + \sum_{t\in T}\pi_t \tag{8.10}$$

mit dem Umlaufplanungs-Unterproblem zur Minimierung geplanter Kosten

$$\Phi_y(\mu,\pi,\nu) = \Bigg\{\min \sum_{d\in D}\sum_{(i,j)\in A^d} y_{ij}^d \bar{c}_{ij}^d \mid \tag{8.11}$$

$$\sum_{\{j:(j,i)\in A^d\}} y_{ji}^d - \sum_{\{j:(i,j)\in A^d\}} y_{ij}^d = 0, \quad \forall d\in\mathcal{D}, \forall i\in N^d, \tag{8.12}$$

$$0 \le y_{ij}^d \le u_{ij}^d, \qquad \forall d\in\mathcal{D}, \forall (i,j)\in A^d\Bigg\} \tag{8.13}$$

und dem Umlaufplanungs-Unterproblem zur Minimierung potenzieller Verspätungspropagation

$$\Phi_z(\nu) = \Bigg\{\min \sum_{d\in D}\sum_{l\in L^d} z_l^d \bar{w}_l^d \mid \tag{8.14}$$

$$z_l^d \in \{0,1\}, \quad \forall l\in L^d\Bigg\} \tag{8.15}$$

und dem Dienstplanungs-Unterproblem

$$\Phi_x(\mu) = \Bigg\{\min \sum_{d\in D}\sum_{k\in K^d} x_k^d \bar{f}_k^d \mid \tag{8.16}$$

$$x_k^d \in \{0,1\}, \quad \forall k\in K^d\Bigg\} \tag{8.17}$$

Als Lagrange-Multiplikatoren sind π_t, ν_{ij}^d und μ_{ij}^d den entsprechenden Restriktionen (8.2), (8.4) und (8.5) zugeordnet.

Die reduzierten Kosten $\bar{c}_{ij}^d$ auf Kante $(i,j) \in A^d$ eines Eindepot Umlaufplanungs-Unterproblems für Depot $d \in D$ sind folglich definiert als

$$\bar{c}_{ij}^d = \begin{cases} c_{ij}^d + \nu_{ij}^d + \mu_{ij}^d - \pi_t & \text{für } : (i,j) \in \widetilde{A}^d \text{ und } \exists t \in \mathcal{T} : (i,j) \in A^d(t) \\ c_{ij}^d + \nu_{ij}^d + \mu_{ij}^d & \text{für } : (i,j) \in \widetilde{A}^d \text{ und } \neg\exists t \in \mathcal{T} : (i,j) \in A^d(t) \\ c_{ij}^d + \nu_{ij}^d & \text{für } : (i,j) \notin \widetilde{A}^d \end{cases} \tag{8.18}$$

Für die reduzierten Kosten $\bar{w}_l^d$ eines Umlaufs $l \in L^d$ gilt:

$$\bar{w}_l^d = \alpha_{vs} w_l^d - \sum_{(i,j) \in A^d(l)} \nu_{ij}^d \tag{8.19}$$

wobei $A^d(l) \subseteq A^d$ die Menge der Kanten beschreibt, die durch Umlauf $l \in L^d$ abgedeckt werden. In ähnlicher Form gilt für die reduzierten Kosten $\bar{f}_k^d$ von Dienst $k \in K^d$:

$$\bar{f}_k^d = f_k^d + \alpha_{cs} v_k^d - \sum_{(i,j) \in \widetilde{A}^d(k)} \mu_{ij}^d \tag{8.20}$$

wobei mit $\widetilde{A}^d(k) \subseteq \widetilde{A}^d$ die Menge an Kanten definiert ist, die durch Dienst $k \in K^d$ abgedeckt werden. Die Parameter w und v bezeichnen jeweils die exakte Verspätungspropagation der berechneten Umläufe und Dienste.

Um eine gute untere Schranke zu bekommen, wird die Zielfunktion des Lagrange-Dualen Problems maximiert. Es werden also die Werte für die Lagrange-Multiplikatoren π, μ und ν gesucht, mit denen $\Phi(\mu, \pi, \nu)$ maximiert wird. Für eine detaillierte Beschreibung des dazu geeigneten Subgradientenverfahrens sei auf [Steinzen, 2007] bzw. [Steinzen et al., 2010] verwiesen.

8.3.3 Pricing-Probleme zur Generierung von Umläufen und Diensten

Die durch das Lösen des (beschränkten) Master-Problems erhaltenen Dualwerte bzw. Lagrange-Multiplikatoren werden verwendet, um Spalten mit negativen reduzierten Kosten zu erzeugen und diese dann dem (beschränkten) Master-Problem als neue Variablen hinzuzufügen, so dass durch Einbeziehen der neuen Spalten in der Lösung der Zielfunktionswert verbessert werden kann. Bei der Berechnung werden zwei Ar-

ten von Spalten unterschieden: Spalten, die Umläufe repräsentieren, und Spalten, die Fahrerdienste darstellen. Für die Generierung von Umläufen wird eine zweiphasige (heuristische) Pricing-Methode vorgeschlagen, in der zunächst Teilumläufe und dann Fahrzeugumläufe erstellt werden. Die Generierung von Diensten erfolgt basierend auf dem Pricing-Schema, welches von [Kliewer et al., 2012] und [Amberg, 2014] vorgeschlagen wird. Zusätzlich zur einfachen Möglichkeit vielfältige Diensttypen abzubilden, erlaubt die dort vorgeschlagene Dienstelement(abschnitt)-basierte Betrachtung das direkte Einbeziehen der Berechnung von Verspätungspropagation zwischen Dienstelementen sowie innerhalb von Sequenzen von Dienstelementen in einzelnen Diensten.

Beide Pricing-Verfahren umfassen das Lösen von ressourcenbeschränkten Kürzeste-Wege-Problemen durch dynamische Programmierung und die exakte Berechnung von Verspätungspropagation. Die Vorgehensweisen werden in den nachfolgenden Abschnitten kurz zusammengefasst.

Generierung von Teilumläufen

Die Ermittlung von Teilumläufen als erster Schritt zur Bestimmung von Umläufen mit negativen reduzierten Kosten erfolgt wie bei der in Abschnitt 7.4.2 beschriebenen Pricing Phase I. Die ressourcenbeschränkten Kürzeste-Wege-Probleme (R-VSP-CG-P I) sind also auf Teilumlaufgenerierungsnetzwerken abgebildet, die alle möglichen Fahrzeugaktivitäten außerhalb eines Depots enthalten. Gelöst werden die Probleme mit dem in Algorithmus 7.3.2 vorgestellten Verfahren, bei welchem die mögliche Verspätungspropagation bei der Teilumlaufgenerierung einbezogen wird. Basierend auf den erstellten Teilumläufen werden dann in einem zweiten Schritt die Fahrzeug- bzw. Tagesumläufe bestimmt.

Generierung von Umläufen

Fahrzeugumläufe mit negativen reduzierten Kosten werden auf Basis der zuvor ermittelten Teilumläufe wie in den in Abschnitten 7.3.2 und 7.4.2 vorgestellten Pricing Phasen II generiert. Mit Hilfe der Teilumläufe wird ein Umlaufgenerierungsnetzwerk aufgestellt, in welchem die kompatiblen, d.h. nacheinander ausführbaren, Teilumläufe miteinander verbunden sind, während auf den Kanten, wie auch bei den Teilumlaufgenerierungsnetzwerken, u.a. (primäre) Verspätungen abgebildet sind. Die entstehenden ressourcenbeschränkte Kürzeste-Wege-Probleme (R-VSP-CG-P II) werden mit Algorithmus 7.3.2 zur korrekten Einbeziehung von Verspätungspropagation bei der Umlaufgenerierung gelöst.

Generierung von Diensten

Um zulässige Dienste mit negativen reduzierten Kosten zu generieren, werden ressourcenbeschränkte Kürzeste-Wege-Probleme (R-CSP-CG-P) basierend auf Dienstelement-basierten Time-Space Dienstgenerierungsnetzwerken gelöst. Der Aufbau der Netzwerke erfolgt dabei analog zum in Abschnitt 7.5.2 beschriebenen Netzwerkaufbau für das Pricing von Diensten im teilintegrierten Lösungsschema. Ein Unterschied des integrierten Falls zum teilintegrierten besteht darin, dass im integrierten Fall statt einer Auswahl an Fahrzeugaktivitäten, die einen Fahrer erfordern, alle Aktivitäten im Dienstgenerierungsnetzwerk abgebildet sind, die sich aus allen möglichen Fahrzeugaktivitäten zur Bedienung der zu auszuführenden Servicefahrten ergeben können.

In den ressourcenbeschränkten Kürzeste-Wege-Problemen werden Dienste bzw. Pfade mit negativen reduzierten Kosten über einen Label-Setting Algorithmus ähnlich zu dem in Abschnitt 7.3.2 vorgestellten Algorithmus 7.3.2 bestimmt. Dabei wird eine mögliche Propagation von Verspätungen explizit in die Diensterstellung einbezogen durch die Betrachtung der aktuellen kumulierten Verspätungspropagation und der gesamten kumulierten Verspätungspropagation eines (Teil-)Pfades an einem Knoten im Dienstgenerierungsnetzwerk.

Beschleunigungstechniken

Beim Lösen der ressourcenbeschränkten Kürzeste-Wege-Probleme werden einfache Techniken genutzt, mit denen die zur Lösung notwendige Rechenzeit verringert werden kann. Insbesondere mit Hilfe von Dominanztests zwischen den Labeln an einem Knoten wird die Menge der im weiteren Lösungsverlauf zu betrachtenden Label reduziert. Dazu werden die Informationen der durch die Label repräsentierten (Teil-)Pfade miteinander verglichen. Ein Label dominiert ein anderes, wenn alle betrachteten Ressourcenverbräuche des (Teil-)Pfades mindestens genauso gut sind und mindestens ein Ressourcenverbrauch besser, d.h. je nach Ressource kleiner oder größer, als die auf dem anderen (Teil-)Pfad. Als Ressourcen in den Dominanztest einbezogen werden auf Umlauf- und Dienstebene, neben der Anzahl bedienter Fahrten und Aufgaben, die zeitliche Dauer und die auf dem (Teil-)Pfad kumulierte Verspätungspropagation. Zusätzlich kann die Anzahl der jeweils an einem Knoten zu betrachtenden (Teil-)Pfade zur Laufzeitreduktion eingegrenzt werden (vgl. [Kliewer et al., 2012]).

8.3.4 Ganzzahlige Lösungen

Die am Ende der ersten Phase errechnete Lösung kann unzulässig sein, da bei der Berechnung Restriktionen (8.2), (8.4) und (8.5) relaxiert wurden. Deshalb wird in der

zweiten Phase eine Lagrange-Heuristik zur Erstellung von zulässigen und zueinander kompatiblen Umlauf- und Dienstplänen genutzt (vgl. [Huisman et al., 2005]). Um dabei die Verspätungspropagation zwischen Fahrzeugen und Fahrerdiensten korrekt zu berücksichtigen, wird ein Lösungsschema ähnlich der teilintegrierten (robusten) Planung von Umläufen und Diensten verwendet (vgl. Abschnitt 7.2): Basierend auf einer zulässigen Netzwerkflusslösung für den Mehrdepot-Umlaufplanungsteil des Planungsproblems wird ein Dienstplan berechnet. Anhand des Dienstplans wird dann wiederum ein geeigneter Umlaufplan aus der Umlaufplanbündel repräsentierenden Flusslösung bestimmt, der eine Propagation von Verspätungen minimiert.

In einem ersten Schritt, wird zunächst eine zulässige Flusslösung für das Mehrdepot-Umlaufplanungsproblem berechnet (vgl. [Huisman et al., 2005] und [Steinzen, 2007]). Dafür werden nur noch die Kopplungsbedingungen (8.4) zwischen Umlaufplanungsteil und Dienstplanungsteil Lagrange-relaxiert, ebenso wie die Kopplungsbedingungen (8.5) zwischen der netzwerkflussbasierten und der pfadbasierten Formulierung des Umlaufplanungsteils. Ähnlich wie in der ersten Phase des Lösungsverfahrens, werden die zugehörigen Lagrange-Dualen-Probleme mit einem Subgradientenalgorithmus gelöst. Um den Rechenaufwand zu verringern, wird das Verfahren mit den Lagrange-Multiplikatoren aus der letzten Iteration der ersten Phase gestartet. Während des Verfahrens werden die zuletzt berechneten Flusslösungen der Umlaufplanungs-Unterprobleme gespeichert. Ergebnis des ersten Schrittes sind also mehrere Bündel von zulässigen Umlaufplänen.

In einem zweiten Schritt werden die gespeicherten Flusslösungen benutzt, um zulässige Umlauf- und Dienstpläne zu generieren. Für jede gespeicherte Flusslösung werden ein Dienstplan und ein dazu kompatibler Umlaufplan konstruiert. Basierend auf der gerade betrachteten Flusslösung wird das zugehörige R-CSP als SPP formuliert und mit einem Column Generation Ansatz gelöst. Master-Problem und Pricing-Probleme werden dabei wie in Abschnitt 7.5.2 beschrieben modelliert und gelöst. Die Bestimmung einer ganzzahligen Lösung, d.h. die Bestimmung eines zulässigen Dienstplans, erfolgt dann ebenfalls wie in Abschnitt 7.5.2 vorgestellt. Der so erstellte, (robuste) Dienstplan dient als Eingabe für das Erstellen eines kompatiblen, robusten Umlaufplanes. Die Vorgehensweise entspricht dabei der in Abschnitt 7.6 dargestellten: Basierend auf den Dienststücken aus dem Dienstplan werden mit Hilfe eines Column Generation Ansatzes (robuste) Umläufe gebildet.

In einem dritten Schritt wird die Kombination aus ermitteltem Umlauf- und Dienstplan als Lösung ausgewählt, die hinsichtlich der Zielsetzungen nach Robustheit und bzw. oder Kosteneffizienz im Vergleich zu den anderen Lösungen bestmöglich bewertet ist.

8.4 Zusammenfassung und Implikationen

In diesem Kapitel wurde eine integrierte Modellformulierung zur Umlauf- und Dienstplanung vorgeschlagen, die steuerbar sowohl geplante Kosten als auch eine mögliche Verspätungspropagation bei der kombinierten Ressourceneinsatzplanung minimiert. Im Vergleich zu Ansätzen aus der Literatur anderer Problemdomänen, wie z.B. bei der Ressourceneinsatzplanung im Flugverkehr, besteht die Besonderheit darin, dass darüber hinaus zum Lösen dieser integrierten Formulierung auch tatsächlich ein integriertes Verfahren genutzt wird. Als Lösungsverfahren wird ein Column Generation Ansatz kombiniert mit Lagrange-Relaxation verwendet. Die korrekte Abbildung möglicher Verspätungspropagation erfolgt dabei in den als ressourcenbeschränkte Kürzeste-Wege-Probleme formulierten Pricing-Problemen zur Erstellung zulässiger Umläufe und Dienste.

Um eine Propagation von Verspätungen auf Umlaufebene berücksichtigen zu können, die Vorteile einer TSN-basierten Formulierung des zugrunde liegenden Umlaufplanungsproblems aber beizubehalten, wird zusätzlich zur Netzwerkflussformulierung des Umlaufplanungsproblems ein Hilfskonstrukt in das integrierte Planungsmodell eingefügt: Analog zur Kopplung der Dienstebene, wird die netzwerkflussbasierte Formulierung des Umlaufplanungsproblems mit einer pfadbasierten Formulierung verknüpft.

Aufgrund der Freiheitsgrade bei der integrierten Umlauf- und Dienstplanung ermöglicht der vorgeschlagene Lösungsansatz, den Ressourceneinsatz zu den gewählten Zielsetzungen (Kosteneffizienz und bzw. oder Robustheit) bestmöglich in Bezug auf das durch einen Fahrplan vorgegebene Servicefahrtenangebot berechnen zu können. Mit dem entwickelten Ansatz können also Referenzwerte bestimmt werden, wenn zur Beantwortung der Forschungsfrage F1 (vgl. Kapitel 4) der Einfluss unterschiedlicher Planungsmethoden auf die mögliche Robustheit und Effizienz der erstellten Pläne untersucht wird. Dabei gibt der Fahrplan den Rahmen für die erreichbare Kosteneffizienz und Robustheit vor (vgl. Abschnitt 2.1). Um die Freiheitsgrade in der Planung noch weiter zu erhöhen, wird im nächsten Kapitel ein Lösungsansatz entwickelt, der Teilentscheidungen der Fahrplanung in die Ressourceneinsatzplanung einbezieht.

Kapitel 9

Robuste Effizienz durch integrierte Planung mit Zeitfenstern

In diesem Kapitel wird ein Lösungsansatz vorgestellt, der zusätzlich zur simultanen Betrachtung des Umlauf- und Dienstplanungsproblems zu einem gewissen Grad die vorgelagerte Planungsstufe der Fahrplanung einbezieht. So werden Abfahrtszeitpunkte für Servicefahrten nicht mehr als fest vorgeben angesehen, sondern die Servicefahrten können in einem bestimmten Zeitfenster nach vorne oder hinten verschoben werden, um weitere Freiheitsgrade für die Ressourceneinsatzplanung zu bieten. Dieser vorgestellte *integrierte* Ansatz basiert auf dem Ansatz für kosteneffiziente Umlauf- und Dienstplanung mit Zeitfenstern der von [Kliewer et al., 2012] vorgeschlagen wird und dem im vorangeganenen Kapitel 8 entwickelten Ansatz zur robusten und kosteneffizienten integrierten Planung. Abbildung 9.1 skizziert die Möglichkeiten, die das Verschieben von Fahrten in kleinen Zeitfenstern für die Ressourceneinsatzplanung bietet.

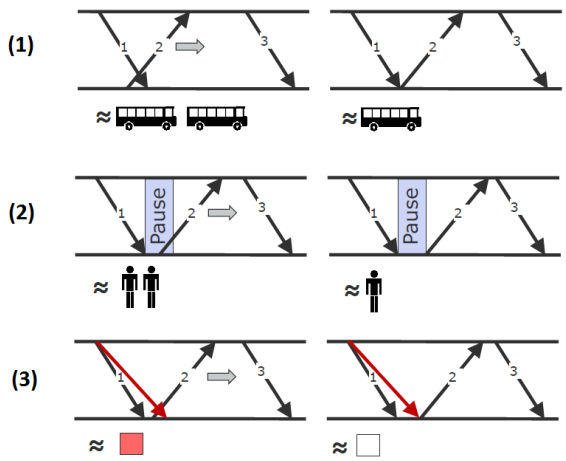

Abbildung 9.1: Potenziale durch Verschiebeintervalle in Bezug auf Kosteneffizienz und Robustheit des Ressourceneinsatzes

Grundidee

Auf der einen Seite können durch das zeitliche Verschieben von Fahrten die Anzahl benötigter Fahrzeuge und Fahrerdienste reduziert und somit die Kosteneffizienz erhöht werden. Es können zusätzliche Verknüpfungen zwischen Fahrten geschaffen werden, die vorher nicht gemeinsam ausgeführt werden konnten (Bild 1 in Abbildung 9.1). Darüber hinaus können zusätzliche Pausenmöglichkeiten für die Diensterstellung geschaffen werden, die die Möglichkeiten zum Erstellen potenziell zulässiger Dienste erhöhen (Bild 2). Auf der anderen Seite ermöglicht das Verschieben von Fahrten die Propagation von Verspätungen zu reduzieren und somit die Robustheit der Pläne zu erhöhen, indem zwischen der Ausführung von Fahrten zusätzlicher zeitlicher Puffer geschaffen wird (Bild 3).

Bezogen auf Robuste Effizienz werden mit dem in diesem Kapitel entwickelten Ansatz folglich zwei eigentlich gegenläufige Effekte gleichzeitig erzeugt: Zum einen das Schaffen von Fahrtverknüpfungen, d.h. eine tendenzielle Verkürzung der Pufferzeit zwischen der Ausführung einzelner Fahrten für mehr Kosteneffizienz. Zum anderen das Auseinanderziehen von Fahrten, d.h. eine tendenzielle Ausweitung der Pufferzeit zwischen der Ausführung einzelner Fahrten für mehr Robustheit.

Im Vergleich zu den in den vorangegangenen Kapiteln 7 und 8 vorgestellten Lösungsansätzen zur Ressourceneinsatzplanung ist die Planung mit Zeitfenstern statt auf der operativen Ebene eher auf taktischer Ebene bzw. auf Ebene der langfristigen Planug einzuordnen (siehe Abbildung 2.1). Ein Eingriff in den aktuellen Fahrplan durch Anpassungen der Abfahrtszeiten findet nicht wöchentlich, sondern eher zu den turnusmäßigen (halb- bzw. jährlichen) Fahrplanwechseln statt. Die Erkenntnisse einer solchen Planung dienen also vor allem auch zum Aufzeigen von noch möglichen Potenzialen, d.h. zum Aufdecken maximal erreichbarer Grenzwerte des aktuell genutzten Ressourceneinsatzes.

Aufbau des Kapitels

Nach der Definition des betrachteten Planungsproblems in Abschnitt 9.1, wird in Abschnitt 9.2 ein Retiming-Ansatz skizziert, dessen Anwendung für die durchgeführten Experimente zur integrierten Betrachtung von Zeitfenstern als Vergleichsgrundlage dient. Anschließend werden die Modellierung der integrierten robusten Umlauf- und Dienstplanung mit Zeitfenstern in Abschnitt 9.3, sowie das Lösungsverfahren inklusive möglicher Beschleunigungstechniken in Abschnitt 9.4 erläutert. Abschnitt 9.5 fasst dieses Kapitel zusammen.

9.1 Problemdefinition R-MDIVCSP-TW

Das *integrierte (Mehrdepot-)Umlauf- und Dienstplanungsproblem mit der Berücksichtigung von Zeitfenstern für Servicefahrten und der Zielsetzung robuster Effizienz* (engl. *robust integrated multiple-depot vehicle and crew scheduling problem with time windows, R-MDIVCSP-TW*) kann wie folgt definiert werden:

Gesucht wird ein kosteneffizienter und in Bezug auf (betrachtete) Verspätungsszenarien möglichst robuster Umlauf- und Dienstplan für eine gegebene Menge an Servicefahrten mit jeweils zugehörigen Verschiebezeitfenstern, sowie für eine zugehörige Menge von Ablösemöglichkeiten, Leerfahrtmöglichkeiten und Transfermöglichkeiten zwischen (End)Haltestellen, und verschiedenen Depot-Fahrzeugtypkombinationen. Die Abfahrts- und Ankunftszeiten der durch den Fahrplan vorgebebenen Servicefahrten dürfen dabei in zuvor festgelegten Zeitfenstern zeitlich nach vorne oder nach hinten verschoben werden, wobei die ursprüngliche Dauer der Servicefahrt immer beibehalten wird. Für die einzelnen Intervalle im Verschiebezeitfenster werden diskrete Werte angenommen, da die An- und Abfahrtszeiten im Fahrplan in der Regel minutengenau ausgewiesen werden. Der resultierende Fahrplan sowie Umlaufplan und Dienstplan sollen zulässig und zueinander kompatibel sein, so dass jede Servicefahrt von einem Umlauf und Fahrerdienst und jede im Umlaufplan definierte zusätzliche Fahrzeugaktivität (z.B. eine Leerfahrt) durch genau einen Dienst ausgeführt wird.

Die Bewertung der Robustheit des Umlauf- und Dienstplans soll dabei über die potenzielle Verspätungspropagation nach der in Abschnitt 5.2.1 beschriebenen Form als Nicht-Robustheitsmaß erfolgen.

9.2 Retiming-Ansatz als Vergleichsgrundlage

Beim nachträglichen Anpassen eines Fahrplanes (engl. *Retiming*) werden basierend auf zuvor berechneten Ressourceneinsatzplänen die Abfahrts- und Ankunftszeiten einzelner Fahrten verschoben, um Ausführung und Belegung der Fahrten mit Ressourcen besser aufeinander abzustimmen. Ist dies in anderen Problemdomänen ein gebräuchliches Mittel, um die Robustheit des geplanten Ressourceneinsatzes nachträglich zu erhöhen (siehe Abschnitt 3.4), so findet sich in der Forschung zur Ressourceneinsatzplanung im ÖPNV bisher keine Untersuchung dazu (siehe Abschnitt 3.3).

Im Folgenden werden unterschiedliche Einsatzmöglichkeiten von Retiming zur Erhöhung der Robustheit herausgearbeitet. Die Unterschiede ergeben sich aus dem Umfang erlaubter Änderungsmöglichktien im Umgang mit den bestehenden Pläne. Dabei sei darauf hingewiesen, dass Retiming im Rahmen dieser Arbeit als rein proaktive

Planungsaufgabe zu verstehen ist. Ein Justieren von Fahrten als reaktive, dispositive Maßnahme findet nicht statt (bei der Simulation zur Evaluation des Ressourceneinsatzes), um eine Verwässerung der Robustheitsbewertung des geplanten Ressourceneinsatzes zu vermeiden.

Kategorisierung von Retiming-Ansätzen

Wird die Zuordnung von Fahrtenfolgen zu Ressourcen und Diensten zu Umläufen beibehalten, bleiben zwei grundlegende Möglichkeiten durch Retiming den Ressourceneinsatz robuster zu gestalten: (A) die in den Plänen vorhandene Pufferzeit zwischen der Ausführung von Fahrten umverteilen, (B) durch das Verschieben von Fahrten zusätzliche Pufferzeit einfügen. Diese zwei Einsatzmöglichkeiten sind in Abbildung 9.2 beispielhaft dargestellt. Grundsätzlich wird Pufferzeit in beiden Fällen so eingesetzt, dass eine durch primäre Verspätungen (gelb) induzierte Verspätungspropagation (rot) verhindert werden kann.

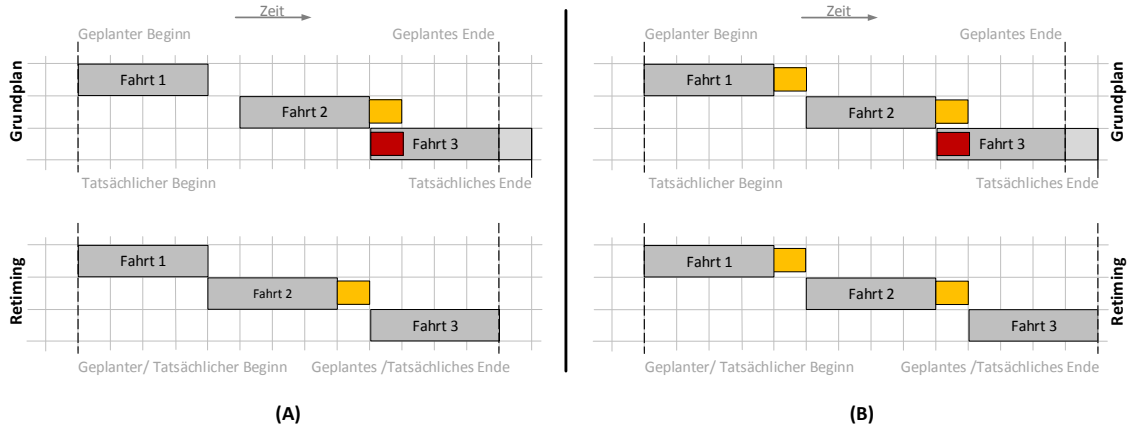

Abbildung 9.2: Einsatz von Retiming zur Reduktion von Verspätungspropagation: (A) vorhandene Pufferzeit kostenneutral umverteilen, (B) zusätzliche Pufferzeit einfügen

Während im ersten Fall vorhandene Pufferzeit kostenneutral an geeigneten Stellen eingefügt wird, wird im zweiten Fall eine Weitergabe von Verspätungen durch eine Verlängerung der Dauer des Ressourceneinsatzes erkauft. Zwar ist die resultierende Verspätungspropagation geringer, Verspätungen werden aber lediglich durch Wartezeit ersetzt bei letztendlich in der Ausführung gleicher Gesamtdauer des Ressourceneinsatzes. So werden Verspätungen gegen zusätzliche geplante Kosten herausgerechnet. Eine vergleichende Bewertung des ursprünglichen Ressourceneinsatzes im Hinblick auf Robustheit wird daher erschwert und der Effekt der Fahrplananpassung überschätzt.

Aus diesem Grund soll der Einsatz von Retiming in dieser Arbeit kostenneutral erfolgen und durch folgende Restriktionen eingeschränkt werden: die ursprüngliche Gesamtdauer einzelner Dienststücke, Dienste und Umläufe soll immer bestehen bleiben. Dafür dürfen Fahrten zu Beginn bzw. Ende eines Dienststückes nicht einzeln, unabhängig von den anderen Fahrten verschoben werden.

Genutzter Retiming-Ansatz

Als Lösungsansatz und somit als Vergleichsgrundlage für die robuste integrierte Planung mit Verschiebeintervallen wird das von [Ahmadbeygi et al., 2010] formulierte Modell zum Retiming im Flugverkehr adaptiert (siehe auch Abschnitt 3.4.1). Es ermöglicht, die Weitergabe einer Verspätung über mehrere Ressourcenebenen abzubilden. Dafür wird für jede Kombination aus einer Fahrt und einer primären Verspätung für diese Fahrt im Vorfeld der Optimierung ein Propagationsbaum aufgebaut. Dieser bildet eine mögliche Propagation der Verspätung über den Ressourceneinsatz für den Fall ab, dass nacheinander ausgeführte Fahrten im Rahmen der Vorgaben so verschoben werden, dass nur noch die geringstmögliche Pufferzeit zwischen diesen Fahrten vorhanden ist. Es wird also der größtmögliche Einfluss einer Verspätungspropagation abgebildet. Mit der Zielfunktion werden die Fahrtenanpassungen so ausgewählt, dass über alle Propagationsbäume gesehen die Verspätungspropagation minimiert wird. Dabei können die entstehenden Propagationswerte gemäß der Eintrittswahrscheinlichkeit der jeweiligen primären Verspätung gewichtet werden. Überdies ist die Formulierung total unimodular, so dass sie mit einem LP-Solver gelöst werden kann.

Die Formulierung ermöglicht es, individuelle Verschiebeintervalle für einzelne Fahrten vorzugeben. Dies wird genutzt, um im Rahmen dieser Arbeit für Fahrten zu Beginn und am Ende eines Dienststückes keine Verschiebungen zugelassen. So kann zum einen vermieden werden, dass die Dauern der Ressourceneinsätze verändert werden und durch z.B. die Überschreitung von Arbeitszeitrestriktionen zu unzulässigen Diensten führen. Zum anderen werden dadurch vor allem nur kostenneutrale Verschiebungen zugelassen (vgl. Abbildung 9.2).

9.3 Modellierung des R-MDIVCSP-TW

Dieser Abschnitt beschreibt die Erweiterung der in Abschnitt 8.2 vorgestellten (robusten) integrierten Umlauf- und Dienstplanungsformulierung (8.1) - (8.8) zu einer integrierten Umlauf- und Dienstplanung mit Berücksichtigung von Verspätungspropagation und der Möglichkeit Fahrplananpassungen durch Zeitfenster bzw. Verschiebeintervalle durchzuführen.

Erweiterung der zugrunde liegenden Time-Space-Netzwerkstruktur

Dafür werden dem zugrunde liegenden Netzwerk zusätzlich zu der bestehenden Kantenmenge (siehe Abschnitt 6) mit *Zeitfenster-Kanten* ein neuer Kantentyp hinzufügt. Diese stellen die Multiplikation von Servicefahrtkanten dar und ermöglichen so, das zeitliche Verschieben der orginalen Servicefahrten zu modellieren. Wie in [Kliewer et al., 2012] beschrieben, stellt jede Zeitfenster-Kante eine Gesamtverschiebung der Servicefahrt um einen bestimmten Zeitraum dar. Da im Rahmen dieser Arbeit eine minutengenaue Planung des Ressourceneinsatzes erfolgt, entspricht somit jede Zeitfensterkante einer zeitlichen Verschiebung um eine Minute. Die mit der entsprechenden Original-Servicefahrtkante assoziierten primären Verspätungen in den bei der Optimierung einbezogenen Szenarien werden für die zugehörigen Zeitfensterkanten übernommen, da in der Regel nur kleine Zeitfenster von wenigen Minuten (im Bereich ± 4) betrachtet werden. Werden große Verschiebeintervalle in Betracht gezogen, sollten in Abhängigkeit der genutzten Verspätungsverteilung, mögliche primäre Verspätungen angepasst werden, um beispielsweise Verschiebungen von Servicefahrten in Rush-Hour-Zeiten hinein oder heraus für die Planung korrekt berücksichtigen zu können.

Mathematische Modellierung

Für das mathematische Modell werden pro neuer Zeitfenster-Kante zusätzliche Flussvariablen definiert. Damit jede Servicefahrt genau einmal abgedeckt wird, werden die Variablen der originalen Servicefahrt-Kanten gemeinsam mit den Variablen der ihnen zugehörigen Zeitfenster-Kanten in den Restriktionen zur Servicefahrtenabdeckung berücksichtigt. Die Menge dieser Restriktionen bleibt daher gleich. Die übrigen Restriktionengruppen werden jeweils um die Menge der Restriktionen ergänzt, welche die einzelnen Zeitfenster-Kanten einbeziehen. Es werden darüber hinaus keine neuen Restriktionengruppen erstellt, allein die Anzahl der aus Abschnitt 8.2 bekannten Restriktionen ändert sich in Abhängigkeit der betrachteten Zeitfenster, resultierend in größeren, noch schwerer lösbaren Optimierungsmodellen.

Das Modell basiert auf der in Abschnitt 6.1 beschriebenen TSN-Struktur und kann somit ähnlich Abschnitt 8.2 wie folgt formuliert werden. Mit $\mathcal{D} = \{1, 2, ..., m\}$ sei die Menge von Depot-Fahrzeugtypkombinationen und mit $\mathcal{T} = \{1, 2, ..., n\}$ die Menge aller Servicefahrten bezeichnet. Im Vergleich zu den vorher betrachteten Umlaufplanungsnetzwerken mit Knotenmenge N^d und Kantenmenge A^d bestehend aus Servicefahrt-, Leerfahrt-, Wartekanten, sowie einer Zirkulationsflusskante, wird für jedes Depot d ein erweitertes Umlaufplanungsnetzwerk $G^d = (extN^d, extA^d)$ definiert.

Um die Verschiebung von Servicefahrten zu modellieren, enthält die darin betrachtete Kantenmenge $extA^d$ zusätzlich die Menge der zur Fahrtenverschiebung benötigten Zeitfenster-Kanten. Durch neu entstehende Fahrtenverknüpfungen ändert sich auch die zugrunde liegende Knotenmenge in die Knotenmenge $extN^d$, um Start- bzw. Endzeitpunkte korrekt abbilden zu können. Die Menge der Kanten, die gleichzeitig Fahrzeug- und Fahrerdienstaktivitäten darstellen sei mit $ext\widetilde{A}^d \subset extA^d$ beschrieben. Diese Menge umfasst also alle Kanten, bis auf Wartekanten im Depot. Zusätzlich wird durch $extA^d(t) : \mathcal{T} \rightarrow extA^d$ eine Funktion definiert, die zu Servicefahrt $t \in \mathcal{T}$ und Depot $d \in \mathcal{D}$ die korrespondierende Servicefahrt-Kante sowie die entsprechenden Zeitfenster-Kanten $(i, j) \in extA^d$ liefert. Für Fahrten t, die nicht von Depot d aus durchgeführt werden können, sei $extA^d(t) = \emptyset$.

Die Flusskapazitäten der Kanten werden wie in Abschnitt 8.2 beschrieben belegt. Zeitfenster-Kanten werden dabei so wie die zugehörigen Servicefahrt-Kanten behandelt.

Die den Kanten $(i, j) \in extA^d$ zugewiesenen Fahrzeugkosten c^d_{ij} werden wie bei der Formulierung R-MDIVCSP in fixe und variable Kosten unterteilt. Dabei erhalten die Flusszirkulationskanten fixe Fahrzeugkosten, die den Einsatz eines Fahrzeuges aus einem Depot repräsentieren. Allen anderen Kanten werden variable Kosten basierend auf der abgebildeten Fahrt- und Wartezeit zugewiesen. Die variablen Kosten der Zeitfenster-Kanten werden (bei Bedarf) mit zusätzlichen Strafkosten für das Verschieben einer Fahrt versehen. Um zu gewährleisten, dass Verschiebungen hauptsächlich stattfinden, wenn dadurch Ressourcen eingespart oder potenziell weniger Verspätungen propagiert werden, werden die variablen Kosten von Zeitfenster-Kanten wie von [Kliewer et al., 2012] vorgeschlagen leicht erhöht. Die zu einer Servicefahrt zugehörigen Zeitfenster-Kanten erhalten Kosten proportional zur abgebildeten Verschiebung. Die geringsten Kosten hat somit die originale Servicefahrt-Kante, die höchsten Kosten besitzen die Zeitfenster-Kanten mit dem größten Verschiebeintervall. So ergeben sich die variablen Kosten c^d_a einer Zeitfenster-Kante als $c^d_a = c^d_0 + |t_a| * \epsilon$, wobei t_a das Verschiebeintervall wiedergibt, c^d_0 den Kosten der original Servicefahrt-Kante entspricht und ϵ den Bestrafungsfaktor für das zeitliche Verschieben darstellt. Durch Variation von ϵ in Bezug zu den anderen gewählten Kostenkomponente des Modells lässt sich ein Trade-Off zwischen Verschiebungen und Kosteneinsparungen bzw. Robustheit abbilden.

Auch für diese Modellformulierung werden mit Flussvariablen und pfadbasierten Variablen zwei Arten von Entscheidungsvariablen definiert. Die ganzzahligen Flussvariablen y^d_{ij} sind mit jeder Kante $(i, j) \in extA^d$ assoziiert. Sie geben an, in welchem Ausmaß die entsprechende Kante in der Lösung benutzt wird. Eine Flusseinheit

entspricht dabei einem Umlauf, der die durch Kante (i,j) repräsentierte Aktivität ausführt.

Sei zusätzlich durch K^d die Menge der von Depot $d \in \mathcal{D}$ ausführbaren Dienste definiert, sowie durch $K^d(i,j) \subset K^d$ die Menge von Diensten, welche die durch Kante $(i,j) \in ext\widetilde{A}^d$ repräsentierte Aufgabe bedienen. Mit den binären Entscheidungsvariablen x_k^d für $k \in K^d$ wird angegeben, ob ein Dienst k für die Lösung genutzt wird oder nicht. In Abhängigkeit der auszuführenden Aufgabe des durch die Variable x_k^d repräsentierten Dienstes, werden jeder solchen Variable geplante Dienstkosten f_k^d und mit Verspätungspropagation assoziierte Kosten v_k^d zugewiesen. Die Verspätungspropagation wird dabei gemäß Formel 5.1 berechnet.

Analog zur Definition der Fahrerdienste, erfolgt die Definition von Umläufen, die eine konkrete Abbildung von Verspätungspropagation im Modell ermöglichen. So sei L^d die Menge der Umläufe, welche von Depot $d \in \mathcal{D}$ aus durchgeführt werden können, sowie $L^d(i,j) \subset L^d$ die Menge von Umläufen, welche die durch Kante $(i,j) \in extA^d$ repräsentierte Fahrt abdecken. Die binären Entscheidungsvariablen z_l^d mit $l \in L^d$ geben an, ob Umlauf l für die Lösung genutzt wird oder nicht. In Abhängigkeit der auszuführenden Fahrten des durch die Variable z_l^d repräsentierten Umlaufes, werden jeder solchen Variable mit Verspätungspropagation assoziierte Kosten w_l^d zugewiesen. Die Propagation von Verspätungen wird ebenfalls anhand Formel 5.1 berechnet.

Das *robuste, integrierte Umlauf- und Dienstplanungsproblem mit mehreren Depots und dem Einbeziehen von sowohl potenzieller Verspätungspropagation der Ressourcen als auch dem Berücksichtigen von Zeitfenstern zum Verschieben von Servicefahrten* (engl. *robust integrated multiple-depot vehicle and crew scheduling problem with time window, R-MDIVCSP-TW*), kann wie folgt formuliert werden.

R-MDIVCSP-TW:

$$\begin{aligned}
&\sum_{d\in\mathcal{D}} \sum_{(i,j)\in extA^d} y_{ij}^d c_{ij}^d + \sum_{d\in\mathcal{D}} \sum_{k\in K^d} x_k^d f_k^d \\
&+\alpha_{vs} \sum_{d\in\mathcal{D}} \sum_{l\in L^d} z_l^d w_l^d + \alpha_{cs} \sum_{d\in\mathcal{D}} \sum_{k\in K^d} x_k^d v_k^d \to \min
\end{aligned} \tag{9.1}$$

$$\text{s.t.} \sum_{d\in\mathcal{D}} \sum_{(i,j)\in extA^d(t)} y_{ij}^d = 1 \qquad \forall t \in \mathcal{T} \tag{9.2}$$

$$\sum_{\{j:(j,i)\in extA^d\}} y_{ji}^d - \sum_{\{j:(i,j)\in extA^d\}} y_{ij}^d = 0 \qquad \forall d \in \mathcal{D}, \forall i \in extN^d \tag{9.3}$$

$$\sum_{k\in K^d(i,j)} x_k^d - y_{ij}^d = 0 \qquad \forall d \in \mathcal{D}, \forall (i,j) \in ext\widetilde{A}^d \tag{9.4}$$

$$\sum_{l\in L^d(i,j)} z_l^d - y_{ij}^d = 0 \qquad \forall d \in \mathcal{D}, \forall (i,j) \in extA^d \tag{9.5}$$

$$0 \leq y_{ij}^d \leq u_{ij}^d,\ y_{ij}^d \in \mathbb{N} \qquad \forall d \in \mathcal{D}, \forall (i,j) \in extA^d \tag{9.6}$$

$$x_k^d \in \{0,1\} \quad \forall d \in \mathcal{D}, \forall k \in K^d \tag{9.7}$$

$$z_l^d \in \{0,1\} \quad \forall d \in \mathcal{D}, \forall l \in L^d \tag{9.8}$$

Die Zielfunktion (9.1) minimiert die geplanten Kosten von Umläufen und Diensten, sowie in Abhängigkeit der Gewichtung der Parameter α die mit Verspätungspropagation assoziierten Kosten auf Umlauf- bzw. Dienstebene. Der Trade-Off zwischen Robustheit und Kosteneffizienz ist hierbei über die einfache Formulierung als gewichtete Summe der verschiedenen Kosten in der Zielfunktion steuerbar. Die Restriktionen (9.2) und (9.3) bilden eine Mehrgüter-Fluss-Formulierung auf Umlaufebene ab. Sie stellen die Gültigkeit eines Bündels von Umlaufplänen sicher. Aufgrund der Nebenbedingungen (9.2) wird jede Servicefahrt genau einmal ausgeführt. Zusätzlich ist die Summe der zu einer Servicefahrt gehörigen Flusseinheiten auf der original Servicefahrt-Kante und den korrespondierenden Zeitfenster-Kanten immer genau eins. Dadurch wird für jede Fahre eine Auswahl zwischen Servicefahrt-Kanten oder entsprechenden Zeitfenster-Kanten getroffen. Die Nebenbedingungen (9.3) sichern Flusserhaltung in jedem Netzwerkknoten. Die flussbasierte Umlaufplanungsformulierung wird durch die Restriktionengruppen (9.4) und (9.5) mit einer pfadbasierten Dienstplanungsformulierung bzw. eine pfadbasierten Umlaufplanungsformulierung gekoppelt. Jede Flusseinheit auf einer durch eine Fahrzeugaktivität beschriebene Kante (außerhalb eines Depots), muss durch ebenso viele Dienste abgedeckt werden. Gleiches gilt für den Zusammenhang zwischen Flusseinheiten und die durch die pfadbasierte Umlaufplanungsformulierung addressierten Umläufe zur Minimierung von Verspätungspropagation. Die Restriktionen (9.6) gewährleisten, dass die maximale Flusskapazität auf den einzelen Kanten eingehalten wird. Wie auch in der Modellformulierung zur integrierten, robusten Umlauf- und Dienstplanung in Abschnitt 8.2, bestimmt die Anzahl der Kanten in der Flussformulierung zur Umlaufplanung und somit die Anzahl der Flussvariablen den Großteil der Anzahl der insgesamt benötigten Restriktionen. Mit zunehmendem Umfang der einbezogenen, potenziellen Verschiebeintervalle, steigt folglich auch die zu betrachtende Problemgröße drastisch. Die Nebenbedingungen (9.7) und (9.8) beschreiben den binären Charakter der pfadbasierten Entscheidungsvariablen, die Dienste und Umläufe repräsentieren.

Erweiterungen der Modellierung für den Praxiseinsatz

Die Formulierung (9.1) - (9.8) kann noch um weitere Restriktionen ergänzt werden, um zusätzliche Anforderungen erfüllen zu können, die sich aus der Anwendungspraxis von ÖPNV-Unternehmen ergeben. Diese beziehen sich insbesondere auf die Bewah-

rung bestimmter Fahrplaneigenschaften und das Einbeziehen der Passagiersicht in die Planerstellung (vgl. auch Abschnitt 2.1). So können Restriktionen hinzugefügt werden, die der Anschlusssicherung zwischen bestimmten Servicefahrten dienen, um die ursprünglich bei der Fahrplanerstellung forcierte Servicequalität beizubehalten und lange Wartezeiten bei Umstiegen zwischen einzelnen Fahrten zu vermeiden. Oder es können zusätzliche Restriktionen ergänzt werden, die das Beibehalten vorhandener, konstanter Taktfrequenzen auf einzelnen Linien sichern. Für eine zusätzliche Diskussion über den Einsatz von Zeitfenstern und das Abbilden von Restriktionen im Praxiseinsatz sei auf [Kliewer et al., 2012] verwiesen.

9.4 Lösungsverfahren

Dieser Abschnitt diskutiert Verfahren zum Lösen der im vorangegangenen Abschnitt vorgestellten Formulierung R-MDIVCSP-TW (9.1) - (9.8) für robuste integrierte Planung mit Zeitfenstern. Im Unterabschnitt 9.4.1 wird zunächst auf das grundlegende Lösungsverfahren eingegangen, während in den daran anschließenden Abschnitten Techniken zur Reduktion der Problemgröße vorgestellt werden. Unterabschnitt 9.4.2 beschäftigt sich mit Problemgrößenreduktion vor dem eigentlichen Lösen des Planungsproblems, wohingegen Unterabschnitt 9.4.3 mit Kantengenerierung ein Verfahren zur Reduktion der zu betrachtenden Problemgröße während der Laufzeit addressiert.

9.4.1 Generelles Verfahren

Das Verfahren zum Lösen der integrierten Formulierung mit Zeitfenstern entspricht in der Grundversion dem in Abschnitt 8.3 beschriebenen Lösungsverfahren. Die vorgestellte Formulierung (9.1) - (9.8) wird also mit einer Kombination aus Column Generation und Lagrange-Relaxation gelöst. Wobei in der ersten Phase durch das Lösen einer Lagrange-relaxierten Problemstellung eine möglichst gute Approximation des Zielfunktionswertes berechnet wird und in der zweiten Phase ganzzahlige Lösungen als zueinander kompatible Umlauf- und Dienstpläne basierend auf berechneten Flusslösungen erstellt werden (vgl. Abbildung 8.2 bzw. Algorithmus 8.3.1). Allein die einzubeziehende Menge der Kanten bzw. möglichen Aufgaben und Fahreraktivitäten im Modell ändert sich in Abhängigkeit der gewählten Zeitfenster pro Servicefahrt im Vergleich zum Vorgehen in Abschnitt 8.3.

Eine zulässige *Initiallösung* wird durch sequenzielle oder teilintegrierte Planung berechnet, wie in Abschnitt 6.3 beschrieben oder in Kapitel 7 für die teilintegrierte

Planung ausführlich erläutert. Um die Konvergenz des Verfahrens mit einer guten Ausgangslösung zu beschleunigen, werden dabei bereits Zeitfenster (TW) in die Planung einbezogen bzw. Robustheitsaspekte (R) betrachtet. So werden auf Umlaufebene bereits Fahrtenverschiebungen als Vorgabe für die weitere Planung zugelassen, die dann beim Bilden passender Dienste berücksichtigt werden. Die für den Umlaufplanungsteil (R-)VSP(-TW) und für den Dienstplanungsteil (R-)CSP(-TW) errechneten Spalten dienen als initiale Spaltenmenge des vollständig integrierten Verfahrens mit Zeitfenstern.

Zur Vereinfachung der Problemstellung werden zur Formulierung des *(beschränkten) Master-Problems* analog Abschnitt 8.3.2 die Kopplungsbedingungen (9.4) und (9.5), sowie die Fahrtenabdeckungsbedingungen (9.2) lagrange-relaxiert. Das entstehende Lagrange-Duale-Problem mit dem Umlaufplanungs-Unterproblem zur Minimierung geplanter Umlaufkosten, dem Umlaufplanungs-Unterproblem zur Minimierung von Verspätungspropagation sowie dem Dienstplanungs-Unterproblem zur Minimierung von geplanten Dienstkosten und Verspätungspropagation wird approximativ mit einem Subgradientenverfahren gelöst. Der Unterschied zur integrierten Planung ohne Zeitfenster besteht darin, dass in Abhängigkeit der Zeitfenstergröße und der dadurch zu berücksichtigenden Zeitfensterkanten die einzelnen Restriktionengruppen der ursprünglichen Formulierung R-MDIVCSP-TW deutlich mehr Restriktionen umfassen.

Ebenso sind im Rahmen der *Pricing Probleme* zur Generierung neuer Spalten, die Diensten bzw. Umläufen entsprechen, je nach Zeitfenstergrößen deutlich umfangreichere Netzwerke zu berücksichtigen. Das Pricing-Verfahren selbst erfolgt analog Abschnitt 8.3.3 für die Umlaufgenerierung in zwei Phasen (R-VSP-CG-P I und II) auf Teilumlaufgenerierungs- bzw. Umlaufgenerierungsnetzwerken je Depot-Fahrzeugtyp, sowie für die Dienstgenerierung (R-CSP-CG-P) auf Dienstelement-basierten Time-Space Dienstgenerierungsnetzwerken je Kombination aus Depot-Fahrzeugtyp und Diensttyp.

Die Berechung einer zulässigen *ganzzahligen Lösung* erfolgt wie in Abschnitt 8.3.4 beschrieben mit einer Lagrange-Heuristik. Die Ausgestaltung entspricht dem teilintegrierten Lösungsschema (vgl. Abschnitt 7.2).

Die zusätzlich im Lösungsverfahren zu berücksichtigenden Zeitfensterkanten und die damit einhergehenden drastisch steigenden Problemgrößen führen bei reiner Anwendung des soeben beschriebenen Lösungsverfahrens zu einem deutlichen Anstieg der Lösungszeiten im Vergleich zu integrierter Planung ohne Zeitfenster (für exemplarische Netzwerkgrößen sowie Lösungszeiten sei auf die später aufgeführten Tabellen 9.1 und 9.2 verwiesen). Im Folgenden werden deshalb Techniken vorgeschlagen, um

die zu betrachtende Problemgröße vor Beginn des integrierten Lösungsverfahrens zu verringern (Abschnitt 9.4.2) oder während des integrierten Lösungsverfahrens einzuschränken (Abschnitt 9.4.3).

9.4.2 Problemgrößenreduktion durch Preprocessing-Techniken

Zur Reduktion der zu betrachtenden Problemgröße und damit zur Reduktion der Laufzeit kann die Menge der möglichen Fahrtenverschiebungen bereits im Vorfeld (heuristisch) begrenzt werden, so dass weniger Flussvariablen und zugehörige Restriktionen, sowie potenzielle pfadbasierte Variablen im mathematischen Modell zu berücksichtigen sind. Dabei sollte die Einschränkung der Fahrtenverschiebungen so ermittelt werden, dass gleichzeitig ähnlich gute Ergebnisse wie bei der erlaubten Verschiebung aller Fahrten erreichbar sein können. Vor dem Hintergrund der Zielsetzungen nach Kosteneffizienz und Robustheit ist zu bestimmen, ob durch Fahrtenverschiebungen zusätzliche Freiheitsgrade durch neue Fahrtenverknüpfungen bzw. Pausenmöglichkeiten oder zusätzlich eingefügte Pufferzeiten ermöglicht werden sollen. Im Einzelnen ist abzuwägen:

- Welcher Zielsetzung dienen die Fahrtenverschiebungen – der Robustheitserhöhung und/oder der Kosteneffizienzsteigerung?
- Für welche konkreten Fahrten sollten Verschiebungen erlaubt werden?
- In welchem Umfang sollten mögliche Verschiebungen individuell pro Fahrt vorgegeben werden?

Die Problematik besteht darin, die passende Menge von verschiebbaren Fahrten sowie das passende Ausmaß an potenziellen Verschiebungen zu finden. Die besten Vorgaben dafür würde die Anwendung der integrierten robusten Planung mit großzügigen Vorgaben für die Zeitfenstern aller Servicefahrten selbst liefern. Im Folgenden werden mit der aus der Literatur bekannten *Schnitt-Heuristik* sowie einer speziell in Bezug auf Robustheitsaspekte vorgeschlagenen *Pufferzeit-Zubilligen-Heuristik* Möglichkeiten für ein Preprocessing zur Identifkation der passenden Menge verschiebbarer Fahrten diskutiert.

Schnitt-Heuristik

Mit diesem Vorgehen werden mögliche Verschiebungen allein anhand der Servicefahrten des Fahrplanes bestimmt. In Ansätzen für rein kostenminimierende Planung

wird dafür das Betrachten von Verschiebeintervalle in Auslastungsspitzen vorgeschlagen (vgl. [Bunte et al., 2006] für Ergebnisse zur Umlaufplanung bzw. [Kliewer et al., 2012] für Ergebnisse zur Umlauf- und Dienstplanung). In solchen Spitzen ist durch die maximale Anzahl parallel auszuführender Fahrten eine untere Schranke für den zu leistenden, gesamten Ressourceneinsatz vorgegeben. Da im Bereich der Auslastungsspitzen am meisten Ressourcen eingesetzt werden, ist dort zum einen das größte Potenzial zur Reduktion des Ressourceneinsatzes vorhanden, indem neue Fahrtenverknüpfungen geschaffen werden. Über den Zeitraum der Auslastungsspitzen gesehen ist ebenso die größte Anzahl von Ressourcen gleichzeitig im Einsatz, die von Verspätungen betroffen sein könnten. So gilt mit Blick auf die gemeinsame Erfüllung von Kosteneffizienz und Robustheit bei der Planerstellung, dass Fahrtenverschiebungen in diesem Zeitraum zum anderen auch ein großes Potenzial bieten, die Propagation von Verspätungen in Aufgabensequenzen vieler Ressourcen für den weiteren Tagesverlauf einzudämmen.

Nach dem von z.B. [Bunte et al., 2006] beschriebenen Vorgehen wird zum Bestimmen der Fahrten, die in den Auslastungsspitzen ausgeführt werden, ein horizontaler Schnitt an der gesamten Auslastungskurve bzw. Ganglinie angelegt. Die Ganglinie stellt dabei über den durch den Fahrplan abgebildeten Zeitraum die Anzahl aller parallel auszuführenden Servicefahrten zu bestimmten Zeitpunkten (hier minutengenau) dar. Die Größe der verschiebbaren Fahrtenmenge wird durch einen Abstands-Parameter γ festgelegt. Dieser gibt an, wie groß der absolute Abstand zwischen der Auslastungsspitze und dem Schnitt sein soll, gemessen in der Anzahl parallel auszuführender Servicefahrten. Die Menge verschiebbarer Fahrten ergibt sich dann aus allen Fahrten die in dem Zeitintervall liegen, das durch die Schnittpunkte der Ganglinie mit der horizontalen Geraden definiert wird. In Abhängigkeit der Ausprägung der Ganglinie können so auch mehrere Intervalle mit verschiebbaren Fahrten bestimmt werden. Die Größe der Zeitfenster für die ermittelten Fahrten wird global vorgegeben und nicht individuell unterschieden. Zur individuellen Bestimmung von Zeitfenstern kann die Schnitt-Heuristik mit der im folgenden beschriebenen Heuristik kombiniert werden.

Pufferzeit-Zubilligen-Heuristik

Es ist naheliegend zur Erhöhung der Robustheit Verschiebeintervalle nicht nur allein anhand des Fahrplans, sondern in Abhängigkeit der historischen Verspätungsinformationen bzw. genutzten Verspätungsszenarien zu definieren, d.h. für diejenigen Fahrten Zeitfenster zu erlauben, die potenziell hohe primäre Verspätungen aufweisen. So können bei der Verknüpfung dieser Fahrten mit anderen Fahrten bei Bedarf durch

Fahrtenverschiebungen zusätzliche Pufferzeiten in den Ressourceneinsatz eingepflegt werden. Bei diesem Vorgehen sind verschiedene Ausprägungen mit unterschiedlich resultierenden Freiheitsgraden möglich.

(A) Die Auswahl erfolgt basierend auf den primären Verspätungen und Zeitfenster werden allein entweder für einen festgelegten Anteil Fahrten mit den größten Verspätungen oder für Fahrten mit Verspätungen größer als einem definierten Schwellenwert θ_{PD} erlaubt. Das Verschieben solcher Fahrten wird nur zeitlich nach vorne erlaubt, um eine mögliche primäre Verspätung kompensieren zu können. Dabei kann das Verschiebeintervall zusätzlich durch eine maximale Verschiebung δ_{max} eingegrenzt werden, um zu große Verschiebungen zu vermeiden. Dieses einfache Vorgehen ähnelt einer Planung mit fixierten Pufferzeiten für einzelne Fahrten mit dem Unterschied, dass die tatsächliche Ausgestaltung der Pufferzeitengröße durch Zeitfenstervorgaben variabel im anschließenden Lösungsverfahren zur Ressourceneinsatzplanung bestimmbar ist. Nachteilig ist, dass eine mögliche Verspätungspropagation der Vorgängerfahrt ebenso wie bei einer Planung mit fixierten Pufferzeiten bei der Bestimmung potenzieller Zeitfenster unberücksichtigt bleibt.

Algorithmus 9.4.1 : Pufferzeit-Zubilligen-Heuristik (Variante B) für die Berechnung von Verschiebeintervallen

```
// Schritt 1: Initialisierung
Setze TW_i^- = 0 ∀i ∈ T
Setze TW_i^+ = 0 ∀i ∈ T
// Schritt 2: Bestimme Verschiebeintervalle
// Wähle Station s
foreach s ∈ Station do
    // Prüfe kompatible Fahrtenpaare (i,j) an Station s,
    // mit i,j ∈ T und i ≠ j
    foreach (i,j) ∈ Paare_s do
        // Prüfe mögliche Verspätungspropagation
        // zwischen Fahrt i und j
        if PAT_i + PD_i > PDT_j then
            // Ermögliche zeitliche Verschiebung δ unter
            // Berücksichtigung bereits geprüfter Kompatibilitäten
            // und erlaubter Maximalverschiebung δ_max
            δ = min{PAT_i + PD_i − PDT_j, δ_max}
            if δ > |TW_i^-| then
                TW_i^- = −δ
            if δ > TW_j^+ then
                TW_j^+ = δ
```

(B) Für die Auswahl werden kompatible Fahrtenpaare an einer Station bzw. Haltestelle geprüft, also Fahrten, die nacheinander ausführbar sind. Algorithmus 9.4.1 beschreibt die Vorgehensweise. Dabei sei $\mathcal{T}$ die Menge aller Servicefahrten, TW_i^- die größtmögliche berechnete zeitliche Verschiebung einer Fahrt $i \in \mathcal{T}$ nach vorne, sowie TW_i^+ die zeitliche Verschiebung nach hinten. Darüber hinaus bezeichne PAT_i die geplante Ankunftszeit, PDT_i die geplante Abfahrtszeit und PD_i die mögliche primäre Verspätung einer Fahrt i. Ist die Propagation einer primären Verspätung zwischen Fahrten denkbar, werden Verschiebungen für die betreffende Fahrt und ihre Nachfolgerfahrt zulassen. Dabei wird das Zeitfenster zur Verschiebung für Fahrt 1 nur nach vorne erlaubt und für Fahrt 2 nur zeitlich nach hinten. Für jede einzelne Fahrt können so aufgrund der potenziellen eigenen primären Verspätung bzw. aufgrund der potenziellen Verspätungspropagation einer möglichen Vorgängerfahrt Zeitfenster vorgegeben werden, die zeitliche Verschiebungen sowohl nach vorne als auch nach hinten zulassen. Um zu große Verschiebungen zu vermeiden, wird das Verschiebeintervall in Abhängigkeit des möglichen Propagationsausmaßes bestimmt und kann (wie bei Variante A) zusätzlich durch eine maximal erlaubte Verschiebung δ_{max} eingegrenzt werden. Es ist anzumerken, dass für die Ausgestaltung der Verschiebeintervalle und möglicher Pufferzeiten nur lokale Entscheidungen während des Verfahrens getroffen werden, da zu ihrer Bestimmung keine Propagationseffekte über mehrere Fahrten abgebildet werden. Das Verfahren dient also allein der Vorgabe zusätzlicher Freiheitsgrade für die Planung. Die Entscheidungen über Fahrtenverschiebungen inklusive der Verschiebungsrichtung können dann in Abhängigkeit der ausgeführten Fahrtenfolgen dem Lösungsverfahren zur Ressourceneinsatzplanung überlassen werden, da dort eine mögliche Propagation über mehrere Aufgabensequenzen abgebildet wird.

Bei der Pufferzeit-Zubilligen-Heuristik liegt der Fokus auf der Erhöhung von Robustheit, dennoch werden zusätzliche Fahrtenverknüpfungen nicht ausgeschlossen, so dass durch Verschiebungen auch eine Reduktion des Ressourceneinsatz möglich ist. Das Vorgehen ist in Kombination mit der Schnitt-Heuristik verwendbar, indem es auf die bereits durch Schnitt-Heuristik vorausgewählten Fahrten zusätzlich noch angewendet wird.

9.4.3 Problemgrößenreduktion durch Kantengenerierung zur Laufzeit

In diesem Abschnitt wird mit dem Einbeziehen von *Kantengenerierung* in das in Abschnitt 9.4.1 beschriebene generelle Lösungsverfahren eine Methodik vorgestellt, um den betrachtenden Ausschnitt des Lösungsraumes während des Verfahrens so

gering und passend wie möglich zu halten. Das erweiterte Lösungsverfahren soll dadurch schneller zu einer guten Lösung konvergieren. Grundidee ist es, in jeder Column Generation Iteration nur diejenige Kantenmenge des zugrunde liegenden Netzwerkmodells in den Lösungsprozess einzubeziehen, die ausreicht, um die aktuelle Lösung möglicherweise zu verbessern. Für Informationen über das Netzwerkmodell sei an dieser Stelle auf den Beginn von Abschnitt 9.3 bzw. Abschnitt 6.1 verwiesen. Als Folge der Auswahl zu betrachtender Kantenmengen brauchen nicht mehr alle Leerfahrtmöglichkeiten sowie alle möglichen Fahrtenverschiebungen gleichzeitig bei der Planerstellung berücksichtigt werden. Stattdessen werden sukzessive verschiedene Bereiche der im Pricing eingesetzten Netzwerke freigegeben und beim Lösen genutzt. Diese Reduzierung der Umlauf- bzw. Dienstgenerierungsnetzwerke folgt dabei der Idee eingeschränkter Netzwerke als Beschleunigungsstrategie in Column Generation Pricing Verfahren (siehe [Desaulniers et al., 2002]).

Der Einsatz von Kantengenerierung ist unabhängig von der in diesem Kapitel adressierten Erweiterung der Problemstellung um das Einbeziehen von Zeitfenstern für Servicefahrten (siehe Abschnitt 9.1) und daher auch für das Lösungsverfahren zur normalen integrierten Planung (siehe Kapitel 8) mit und ohne Einbeziehen von Robustheitsaspekten anwendbar. Bei der normalen integrierten Planung sind bereits große Netzwerke mit vielfältigen Kombinationsmöglichkeiten zur Umlauf- und Dienstbildung zu berücksichtigen, so dass eine dynamische Einschränkung der Netzwerke während des Lösungsverfahrens ebenfalls deutliche Laufzeitvorteile bieten kann. Insbesondere zu Beginn des Column Generation Verfahrens existiert eine hohe Anzahl Variablen mit negativen reduzierten Kosten (vgl. auch Beobachtung von [Steinzen, 2007] zur Dienstgenerierung bei reiner kostenminimierenden integrierten Planung), so dass hier eine geschickte und drastische Einschränkung der im Pricing zu nutzenden Netzwerkkanten das gesamte Lösungsverfahren beschleunigen kann.

Für die Umsetzung ist zum einen zu klären, welche Netzwerkkanten initial als Grundlage für den Lösungsprozess bestimmt werden. Zum anderen ist zu klären, wie möglicherweise die Lösung verbessernde Kanten basierend auf dem Zusammenspiel der unterschiedlichen Modellkomponenten während des Lösungsprozesses identifiziert werden können. Es ist also vor der Generierung von Diensten und (robusten) Umläufen ein zusätzlicher Pricing Schritt zur Generierung von Kanten durchzuführen. Im Unterschied zum Pricing von Diensten und Umläufen sind dabei alle möglichen Kanten vorher explizit bekannt und können als Teil der zugrunde liegenden Netzwerkflussformulierung ausgewählt werden.

Abbildung 9.3 zeigt die Erweiterung des Lösungsschemas für eine integrierte Planung (mit Zeitfenstern) unter Einbeziehen von Kantengenerierung. Zusätzlich fasst

Algorithmus 9.4.2 : Lösungsverfahren für R-MDIVCSP-TW mit Kanten- und Spaltengenerierung

(Schritt 1) **Initialisierung**
Löse MD-(R-)VSP(-TW) und anschließend (R-)CSP(-TW)
für jede Depot-Fahrzeugtypkombination
mit Hilfe des SEQ oder PINT Lösungsschemas.
Nimm die (R-)VSP(-TW)- und die (R-)CSP(-TW)-Lösung als
initiale Spaltenmenge $L' \cup K'$.
Nimm die Kanten der der initialen Lösung zugrunde liegenden
Flusslösung als initiale Kantenmenge $extA'$.
Setze $t = 0$.

(Schritt 2) **Löse Beschränktes Master Problem**
Löse ein Lagrange-Dual-Problem mit der aktuellen Spaltenmenge
$L' \cup K'$, sowie der (gesamten) Kantenmenge $extA$.
Speichere die untere Schranke für die aktuelle Spaltenmenge sowie
die dualen Informationen.

(Schritt 3) **Löse Pricing-Probleme zur Kantengenerierung**
Generiere neue Kanten $extA''$ mit $extA'' \subseteq extA$ und
$extA'' \cap extA' = \emptyset$ mit Hilfe von Strategie Rn, NR oder FS

(Schritt 4) **Kantenmanagement**
Erweitere aktuelle Kantenmenge $extA' := extA' \cup extA''$
Füge mit $extA''$ korrespondierende Kanten den einzelnen
Pricing-Netzwerken hinzu.

(Schritt 5) **Löse Pricing-Probleme zur Spaltengenerierung**
Generiere neue Spalten bzw. Fahrzeugumläufe L'' mit negativen
reduzierten Kosten mit Hilfe von Modell R-VSP-CG-P I und II.
Generiere neue Spalten bzw. Dienste K'' mit negativen reduzierten
Kosten mit Hilfe von Modell R-CSP-CG-P.
(A) Betrachte dazu Netzwerke basierend auf Kantenmenge $extA'$.
(B) Betrachte dazu Netzwerke basierend auf Kantenmenge $extA$.
Falls $|L'' = 0|$ und $|K'' = 0|$, gehe zu Schritt 7.

(Schritt 6) **Spaltenmanagement**
Füge neue Spalten zum beschränkten Master Problem hinzu:
$L' := L' \cup L''$ und $K' := K' \cup K''$.
Lösche Spalten mit hohen positiven reduzierten Kosten von L' (K')
falls $|L'|$ ($|K'|$) groß und über einem definierten Schwellenwert ist.

(Schritt 7) **Prüfe Abbruchbedingungen**
Falls eine der folgenden Abbruchbedingungen gilt:
- max. Anzahl Iterationen erreicht: $t \geq t_{max}$
- keine signifikante Verbesserung der unteren Schranke während
einer bestimmten Anzahl letzter Iterationen,
gehe zu Schritt 8;
ansonsten setze $t = t + 1$ und gehe zu Schritt 2.

(Schritt 8) **Bestimme eine zulässige Lösung**
Nutze eine Lagrange-Heuristik, um einen zulässigen, kompatiblen
Umlauf- und Dienstplan zu erstellen

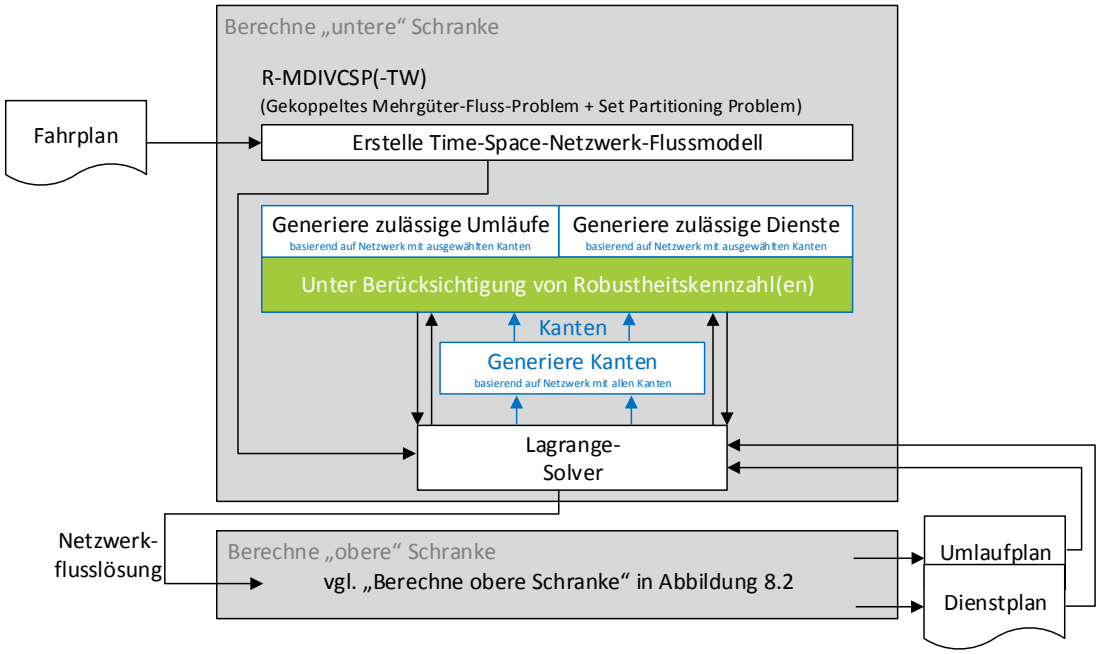

Abbildung 9.3: Einbeziehen von Kantengenerierung in das integrierte Lösungsverfahren zur Umlauf- und Dienstplanung (mit Zeitfenstern)

Algorithmus 9.4.2 das durch Einbeziehen von Kantengenerierung in den integrierten Lösungsprozess mit Zeitfenstern entstehende Lösungsverfahren zusammen. Für eine Planung ohne Zeitfenster ist statt der Kantenmenge $extA$ die Kantenmenge A zu verwenden (vgl. Abschnitte 8.2 bzw. 9.3). Im Folgenden werden insbesondere die Schritte näher erläutert, die sich durch Einbeziehen einer veränderlichen Kantenmenge im Vergleich zum Grundverfahren der integrierten Planung mit Zeitfenstern (vgl. Abschnitt 9.4.1) bzw. dem Verfahrensablauf bei integrierter Planung (vgl. Algorithmus 8.3.1 in Abschnitt 8.3) verändern. Im Einzelnen sind dies das Bestimmen einer *initialen Kantenmenge* (siehe Schritt 1), das *Pricing neuer Kanten* (siehe Schritt 3) und das *Pricing von Diensten und (robusten) Umläufen* (siehe Schritt 5), sowie das zugehörige zugehörige *Kanten-* und *Spaltenmanagement* (siehe Schritte 4 und 6).

Initiale Kantenmenge

Für die integrierte Planung (mit Zeitfenstern) wird zunächst eine Initiallösung berechnet wie in Abschnitt 9.4.1 beschrieben. Im Wesentlichen existieren dann zwei Varianten, eine initiale Kantenmenge zu bestimmen.

(A) Als Ausgangspunkt für den weiteren Lösungsprozess werden nur die in dieser Initiallösung genutzten Netzwerkkanten $extA'^d \subset extA^d$ je Depot-Fahrzeugtypkombi-

nation $d \in \mathcal{D}$ verwendet. So besteht das für den weiteren Verfahrensverlauf zugrunde liegende Umlaufplanungsnetzwerk G'^d aus den Kanten (sowie zugehörigen Knoten) der Flusslösung der initialen Lösung und die jeweiligen Dienstgenerierungs- bzw. Umlaufgenerierungsnetzwerke aus den dazu korrespondierenden Kanten (sowie zugehörigen Knoten). Zusätzlich werden alle Wartekanten im Depot ergänzt.

(B) Als Variante lässt sich für eine integrierte Planung mit Zeitfenstern das Lösungsverfahren mit Kantengenerierung in abgeschwächter Form unter Beibehaltung der kompletten originalen Problemstruktur durchführen. Dann bleiben alle originalen Netzwerkkanten bestehen und nur die durch Fahrtenverschiebungen hinzukommenden Kanten gehören nicht zur initialen Kantenmenge. Für die initiale Menge $extA'^d$ je Depot-Fahrzeugtypkombination $d \in \mathcal{D}$ gilt dann $A^d = extA'^d \subset extA^d$ und das für den weiteren Verfahrensverlauf zugrunde liegende Umlaufplanungsnetzwerk G'^d entspricht mit $G'^d = (N^d, A^d)$ dem für eine integrierte Planung definierten Umlaufplanungsnetzwerk (vgl. Abschnitt 8.2). So werden Freiheitsgrade zur Umlauf- und Dienstbildung vorgegeben, die zunächst mit reiner integrierter Planung vergleichbar sind und dann sukzessive erweitert werden. Die genutzten Netzwerke besitzen dann allerdings mindestens den Umfang der originalen Problemstellung, so dass die einzelnen Pricing-Schritte im Vergleich zum reinen integrierten Verfahren nicht schneller ausgeführt werden. Da durch das Einbeziehen von Kantengenerierung die Gesamtperformanz des integrierten Lösungsverfahrens merklich verbessert werden soll, wird im Rahmen dieser Arbeit – sofern nicht anders angegeben – eine größtmögliche initiale Einschränkung der Kantenmenge durch Variante A gewählt.

Tabelle 9.1 bietet einen Vergleich der sich ergebenden initialen Netzwerkgrößen bei integrierter Planung (INT) sowie bei Nutzung von integrierter Planung mit verschiedenenen Zeitfenstergrößen (von ± 1 bis ± 4 Minuten) ohne Einbeziehen von zusätzlichen Robustheitsaspekten (INT-TW). Die verwendeten Instanzen werden in Abschnitt 11.1.1 beschrieben. Die Netzwerkgrößen sind jeweils exemplarisch aufgeführt für eine Depot-Fahrzeugtypkombination $d \in \mathcal{D}$ inklusive zugehöriger Dienstgenerierungsnetzwerke. Es ist anzumerken, dass bei Einbeziehen von Robustheitsaspekten und damit einhergehend dem zusätzlichen Nutzen von Teilumlauf- bzw. Umlaufgenerierungsnetzwerken, die Netzwerkgrößen der Teilumlaufgenerierungsnetzwerke denen des korrespondierenden Umlaufplanungsnetzwerkes entsprechen zusätzlich der Verbindungen zum ergänzten Quell- bzw. Senkeknoten. Die Netzwerkgrößen der Umlaufgenerierungsnetzwerke sind in dem Fall allerdings abhängig von der Anzahl erstellter Teilumläufe. In der Tabelle angegeben sind die Netzwerkgrößen des zugrunde liegenden Umlaufplanungsnetzwerkes G'^d (V) sowie zugehöriger Dienstgenerierungsnetzwerke $G_C'^{d,t}$ ($D1$-$D5$) bei integrierter Planung *ohne Kantengenerierung*, sowie die

initialen Netzwerkgrößen bei integrierter Planung *mit Kantengenerierung*. Die durch die Dienstgenerierungsnetzwerke abgebildeten Diensttypen t entsprechen der Definition von Kurzdiensten (D1), geteilten Diensten (D2), sowie zusammenhängenden Dienste mit unterschiedlichen Mindestpausenzeiten (D3 bis D5). Die konkreten Ausprägungen dieser Diensttypen sind im Abschnitt 11.1.1 in Tabelle 11.1 wiedergegeben. Zur besseren Lesbarkeit von Tabelle 9.1 sind die Netzwerkgrößen für integrierte Planung mit Zeitfenstern als Faktoren in Bezug auf das jeweils zugrunde liegende normale integrierte Planungsnetzwerk aufgeführt (gerundet auf eine Nachkommastelle). Genaue Angaben der Netzwerkgrößen sind für die Planungsvarianten INT und INT-TW4 gegeben. Darüber hinaus ist für jedes Netzwerk beim Einsatz von Kantengenerierung eine Größenangabe relativ zur Größe des entsprechenden Netzwerkes ohne Einsatz von Kantengenerierung angegeben ($\leftarrow$ %).

Die Anzahl der Kanten im Umlaufplanungsnetzwerk entspricht jeweils der Anzahl der zu berücksichtigenden Kopplungsbedingungen (9.4) zwischen der flussbasierten Umlaufplanungsformulierung und dem pfadbasierten Dienstplanungsteil des Modells, sowie denen bei Einbeziehen von Verspätungspropagation zu berücksichtigen Kopplungsbedingungen (9.5) zwischen fluss- und pfadbasierter Umlaufplanung ohne Warten im Depot. Nicht eingerechnet ist jeweils die Zirkulationsflusskante. Dabei bedingt eine Kante des zugrunde liegenden Umlaufplanungsnetzwerk durchaus mehrere Kanten in den Dienstgenerierungsnetzwerken, da in Abhängigkeit der durch die Kante repräsentierten Aufgabe zusätzliche Pausen- sowie Aufrüst- und Abrüst-Kanten zu berücksichtigen sind (vgl. Netzwerkbeschreibung in Abschnitt 7.5.2).

Wie anhand der in der Tabelle aufgeführten Faktoren zu erkennen ist, führt das Einbeziehen von Zeitfenstern durch Multiplikation der ursprünglichen Servicefahrtkanten in verschiedene Ausprägungen zu explodierenden Netzwerkgrößen. Bei späterem Einsatz von Kantengenerierung und der Reduzierung der einzubeziehenden Kanten anhand der Initiallösung verdoppeln sich jedoch nur die Netzwerkgrößen maximal, auch wenn im größten Zeitfensterintervall von -4 bis $+4$ Minuten eigentlich das Neunfache an möglichen, zu bedienenden Servicefahrten zu berücksichtigen ist.

Pricing neuer Kanten

Bevor die durch das Lösen des (beschränkten) Master-Problems gewonnenen dualen Informationen zum Pricing von den Zielfunktionswert verbessernden Diensten bzw. Umläufen eingesetzt werden, werden die dualen Informationen genutzt, um neue Kanten für die einzelnen Pricing-Netzwerke freizuschalten. Zielsetzung dabei ist es, Kantenmengen zu identifizieren, deren Einbezug zu verbessernden Diensten bzw. Umläufen führen können ohne sämtliche Kanten des zugrunde liegenden Umlaufpla-

Tabelle 9.1: Initiale Netzwerkgrößen bei Einsatz von Kantengenerierung im Vergleich für rein kosteneffiziente Planung

Methode	—ohne Kantengenerierung— V(∗)	D1(∗)	D2(∗)	D3(∗)	D4(∗)	D5(∗)	—mit Kantengenerierung— V(∗)	← %	D1(∗)	← %	D2(∗)	← %	D3(∗)	← %	D4(∗)	← %	D5(∗)	← %
Testinstanz A																		
	2054	2856	34309	4506	3923	3492	832		1034		29260		1975		1657		1407	
INT	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	40,5	(∗1,0)	36,2	(∗1,0)	85,3	(∗1,0)	43,8	(∗1,0)	42,2	(∗1,0)	40,3
INT-TW1	(∗2,1)	(∗2,0)	(∗2,6)	(∗3,2)	(∗2,7)	(∗2,5)	(∗1,3)	24,7	(∗1,1)	19,9	(∗1,3)	43,3	(∗1,3)	17,7	(∗1,2)	18,5	(∗1,2)	18,7
INT-TW2	(∗3,2)	(∗2,9)	(∗4,3)	(∗5,6)	(∗4,7)	(∗4,1)	(∗1,7)	21,2	(∗1,2)	14,8	(∗1,9)	38,2	(∗1,7)	13,6	(∗1,5)	13,8	(∗1,4)	13,7
INT-TW3	(∗4,4)	(∗3,9)	(∗5,8)	(∗7,8)	(∗6,5)	(∗5,7)	(∗1,9)	17,8	(∗1,3)	11,8	(∗2,4)	35,8	(∗2,0)	11,4	(∗1,7)	11,4	(∗1,6)	11,2
INT-TW4	(∗5,7)	(∗4,9)	(∗7,0)	(∗9,9)	(∗8,1)	(∗7,2)	(∗2,1)	15,1	(∗1,3)	10,0	(∗2,8)	34,0	(∗2,3)	10,1	(∗1,9)	10,0	(∗1,8)	9,9
	11755	13932	240238	44460	31932	25135	1772		1395		81662		4491		3199		2477	
Testinstanz B																		
	3011	4250	80565	8821	7022	5958	1233		1555		68385		3814		2952		2409	
INT	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	40,9	(∗1,0)	36,6	(∗1,0)	84,9	(∗1,0)	43,2	(∗1,0)	42,0	(∗1,0)	40,4
INT-TW1	(∗2,2)	(∗2,0)	(∗2,4)	(∗3,0)	(∗2,7)	(∗2,5)	(∗1,4)	25,4	(∗1,2)	20,9	(∗1,5)	53,1	(∗1,4)	20,4	(∗1,3)	20,5	(∗1,3)	20,4
INT-TW2	(∗3,6)	(∗3,1)	(∗3,3)	(∗4,5)	(∗4,2)	(∗3,9)	(∗1,6)	18,2	(∗1,3)	15,0	(∗2,0)	50,0	(∗1,7)	16,5	(∗1,6)	16,4	(∗1,5)	16,0
INT-TW3	(∗5,2)	(∗4,2)	(∗3,8)	(∗5,6)	(∗5,3)	(∗5,0)	(∗1,7)	13,1	(∗1,3)	11,3	(∗2,2)	48,3	(∗1,9)	14,4	(∗1,8)	14,0	(∗1,7)	11,7
INT-TW4	(∗6,9)	(∗5,4)	(∗4,1)	(∗6,5)	(∗6,2)	(∗6,0)	(∗1,7)	10,2	(∗1,3)	9,0	(∗2,3)	47,0	(∗2,0)	13,3	(∗1,9)	12,7	(∗1,7)	11,7
	20676	23075	331169	57330	43633	35866	2109		2068		155690		7648		5524		4179	
Testinstanz C																		
	10594	12834	211830	22564	18589	16333	2253		2820		173165		6224		4907		3944	
INT	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	(∗1,0)	21,3	(∗1,0)	22,0	(∗1,0)	81,7	(∗1,0)	27,6	(∗1,0)	26,4	(∗1,0)	24,1
INT-TW1	(∗1,7)	(∗1,6)	(∗2,0)	(∗2,5)	(∗2,3)	(∗2,1)	(∗1,3)	15,9	(∗1,1)	15,2	(∗1,4)	44,1	(∗1,3)	14,5	(∗1,3)	14,9	(∗1,2)	14,4
INT-TW2	(∗2,5)	(∗2,4)	(∗2,4)	(∗3,4)	(∗3,1)	(∗2,9)	(∗1,6)	13,4	(∗1,4)	12,6	(∗1,6)	55,5	(∗1,6)	13,0	(∗1,5)	13,0	(∗1,5)	12,6
INT-TW3	(∗3,5)	(∗3,2)	(∗2,5)	(∗4,1)	(∗3,9)	(∗3,7)	(∗1,9)	11,6	(∗1,6)	10,9	(∗1,7)	55,4	(∗1,8)	12,3	(∗1,7)	11,9	(∗1,8)	11,5
INT-TW4	(∗4,7)	(∗4,2)	(∗2,6)	(∗4,7)	(∗4,6)	(∗4,5)	(∗2,2)	9,9	(∗1,8)	9,3	(∗1,8)	56,3	(∗2,0)	11,7	(∗1,9)	11,1	(∗1,9)	10,5
	50157	54010	556168	107153	85501	73393	4978		5049		313076		12486		9468		7675	

nungsnetzwerkes G^d je Depot-Fahrzeugtypkombination $d \in \mathcal{D}$ sowie der korrespondierenden Teilumlauf-, Umlauf- und Dienstgenerierungsnetzwerke berücksichtigen zu müssen.

Dabei können auf vielfältige Weise Entscheidungen über die Kantenauswahl getroffen werden. Aufgrund der dreigeteilten Modellierung mit zugrunde liegender Flussformulierung als gemeinsame Verbindungsstelle zwischen den verschiedenen Problembestandteilen (vgl. Abschnitt 9.3 dieses Kapitels bzw. Abbildung 8.1 im Kapitel zur integrierten Planung ohne Zeitfenster) bieten sich unterschiedliche Strategien an. Diese können beispielsweise die Kopplung der verschiedenen Modellteile adressieren oder einzelne Modellteile selbst. Im Folgenden werden die entwickelten Strategien vorgestellt. Die Performanz der einzelnen Strategien wird im Anschluss in Tabelle 9.2 mit dem generellen Lösungsverfahren ohne Kantengenerierung verglichen.

Zufallsbasiertes Pricing (Rn): Als Vergleichsmaßstab für die Funktionsweise von Kantengenerierung und die Performanz anderer Strategien werden bei diesem Vorgehen zufällig ausgewählte Kanten den einzelnen Netzwerken hinzugefügt. Dazu werden in jeder Iteration zufällig eine maximale Anzahl n (mit $n \in \mathbb{N}$) von Kanten aus der Menge $extA^d \setminus extA'^d$ je Depot-Fahrzeugtypkombination $d \in \mathcal{D}$ bestimmt, die dann in den korrespondierenden Netzwerken ergänzt werden. In den durchgeführten Experimenten hat sich gezeigt, dass so für kleinere Instanzen bei ausreichend vielen Iterationen gute Lösungen erreicht werden können, da die einzelnen Netzwerke immer größer werden und sich den originalen Netzwerken annähern. Die Laufzeiten bleiben dabei häufig unter den Laufzeiten des normal durchgeführten Pricings ohne Kantengenerierung. Bei Instanzen mit größeren Netzwerken werden in der Tendenz schlechte Lösungen erzielt, da die Kantenauswahl nicht zielgerichtet stattfindet und daher zu schwacher Konvergenz im Lösungsverfahren führt. Dies bewirkt teilweise einen frühzeitigen Abbruch des Lösungsverfahrens.

Pricing basierend auf Dualwerten der Kopplungsbedingungen bzw. Pricing basierend auf negativen reduzierten Kosten (NR): Eine naheliegende Variante der Kantenauswahl besteht darin, diejenigen Kanten den Pricing-Netzwerken hinzuzufügen, die basierend auf den Dualwerten bzw. zugehöriger Lagrange-Multiplikatoren negative reduzierte Kosten besitzen und deren Einbeziehen in die Lösung daher den Zielfunktionswert verbessern würde. Dazu werden für jede Kante $(i, j) \in extA^d \setminus extA'^d$ jeder Depot-Fahrzeugtypkombination $d \in \mathcal{D}$ die mit den Kopplungsbedingungen (9.4) auf Dienstebene assoziierten Lagrange-Multiplikatoren μ_{ij}^d bzw. bei Generierung (robuster) Umläufe auch die mit den Kopplungsbedingungen (9.5) auf Umlaufebene assoziierten Lagrange-Multiplikatoren ν_{ij}^d und zugehörige Kantenkosten berücksichtigt. Besitzt die Kante (in beiden Fällen) negative reduzierte Kosten,

wird sie der Kantenmenge $extA'^d$ hinzugefügt und zum Pricing neuer Dienste bzw. Umläufe in den Lösungsprozess einbezogen. Die Anzahl neu einzubeziehender Kanten wird pro Iteration begrenzt – ähnlich wie beim Zufallsbasierten Pricing. Dabei ist die Wahl der Begrenzung entscheidend für die Konvergenz des Verfahrens. Die Anzahl ist so zu wählen, dass auf der einen Seite die einzelnen Netzwerke nicht zu groß werden. Auf der anderen Seite sollte gleichzeitig eine zu rigorose Einschränkung der Erweiterung der Netzwerke und damit des betrachteten Lösungsraumes vermieden werden, da ansonsten übermäßig viele Column Generation Iterationen für einen Lösungsfortschritt benötigt werden. In den durchgeführten Experimenten hat sich gezeigt, dass für die Anzahl ein Viertel der gesamten Servicefahrtenanzahl einen guten Richtwert bietet. In den Experimenten ohne eine Begrenzung der hinzuzufügenden Kantenanzahl hat sich gezeigt, dass insbesondere in den ersten Column Generation Iterationen viele Verbesserungsmöglichkeiten existieren, so dass zunächst sehr viele Kanten den Netzwerken hinzugefügt werden. Dadurch entstehen bereits zu Beginn sehr große Netzwerke, einhergehend mit einer insgesamt großen Anzahl generierter Spalten und schlechter Konvergenz des Lösungsverfahrens.

Lagrangian Pricing bzw. Pricing basierend auf einer Flusslösung (FS): Bei der zuvor beschriebenen Strategie (NR) werden Kanten allein basierend auf ihren zugehörigen reduzierten Kosten ausgewählt und so lokale Entscheidungen über die Aufnahme von Kanten in den Lösungsprozess getroffen. Stattdessen werden bei der hier vorgestellten Pricing-Strategie (FS) Kanten im Zusammenspiel mit allen anderen Kanten des gesamten Netzwerkes beurteilt, um zur Auswahl geeignete Kanten zu identifizieren. So wird berücksichtigt, dass die durch die Kanten repräsentierten einzelnen Fahrten bzw. Aufgaben in einer Lösung miteinander innerhalb von Umläufen oder Diensten verbunden sind. Bisher nicht ausgewählte Kanten mit positiven reduzierten Kosten können also auch in Kombination mit bisher nicht ausgewählten Kanten mit negativen reduzierten Kosten zum Lösungsfortschritt bzw. zu einer Verbesserung des Zielfunktionswertes beitragen. Ein solches Vorgehen wird von [Löbel, 1997] im Rahmen von Column Generation bei Umlaufplanungsproblemen vorgeschlagen. Gemäß der Idee des *Lagrangean Pricing* von [Löbel, 1997] bzw. [Löbel, 1998] für die Umlaufplanung, die von [Bunte, 2009] für die Umlaufplanung auf einem Time-Space-Netzwerk übernommen wird, ist zur Identifikation solcher Kantenmengen ein vereinfachtes Problem unter Einbeziehen aller Kanten zu lösen basierend auf den Dualwerten der aktuellen Column Generation Iteration. Dafür sind die einzelnen Kanten anstelle der normalen Kantenkosten (vgl. Abschnitt 9.3) mit den jeweiligen reduzierten Kosten zu bewerten und als Zielsetzung wird eine Lösung mit den geringsten reduzierten Kosten bestimmt. Die in dieser Lösung verwendete Kantenmenge ist po-

tenziell dazu geeignet, den Zielfunktionswert des Gesamtproblems zu verbessern. Zur Vereinfachung der zu lösenden Problemstellung werden dabei einzelne Restriktionen des ursprünglichen Problems Lagrange-relaxiert. Im Kontext dieser Arbeit stellt sich die Frage, welches vereinfachte Planungsproblem geeignet ist, zum einen wirkungsvoll Kanten zu bestimmen ohne die Gesamtlaufzeit des Lösungsprozesses signifikant zu erhöhen, zum anderen gleichzeitig den aktuellen Lösungsstatus bzw. die verschiedenenen Problembestandteile in der Lösung berücksichtigen kann. Aufgrund der bereits Lagrange-relaxierten Abdeckungsbedingungen (9.2) und Kopplungsbedingungen (9.4) sowie (9.5) stellt das Unterproblem zur Minimierung der geplanten Umlaufkosten eine solche Problemstellung dar. Es wird als Verbindung zwischen den einzelnen Problemteilen unter Einbeziehen aller Multiplikatoren μ, π, ν gelöst. Das Pricing von Kanten basierend auf einer Flusslösung entspricht somit pro Column Generation Iteration dem nochmaligen Lösen eines (kostenminimierenden) Umlaufplanungs-Unterproblems mit den aktuellen Multiplikatoren aus der letzten Iteration des Subgradientenverfahrens. Das dafür zu lösende Umlaufplanungs-Unterproblem für jedes Depot $d \in \mathcal{D}$ auf um mögliche Verschiebungen von Servicefahrten erweiterten Netzwerken kann formuliert werden als

$$\Phi_y(\mu, \pi, \nu) = \Bigg\{ \min \sum_{d \in D} \sum_{(i,j) \in extA^d} y_{ij}^d \bar{c}_{ij}^d \mid \tag{9.9}$$

$$\sum_{\{j:(j,i) \in A^d\}} y_{ji}^d - \sum_{\{j:(i,j) \in A^d\}} y_{ij}^d = 0, \quad \forall d \in \mathcal{D}, \forall i \in extN^d, \tag{9.10}$$

$$0 \leq y_{ij}^d \leq u_{ij}^d, \qquad \forall d \in \mathcal{D}, \forall (i,j) \in extA^d \Bigg\} \tag{9.11}$$

Dabei sind die reduzierten Kosten $\bar{c}_{ij}^d$ auf Kante $(i,j) \in extA_{ij}^d$ eines Eindepot Umlaufplanungs-Unterproblems für Depot $d \in \mathcal{D}$ definiert als

$$\bar{c}_{ij}^d = \begin{cases} c_{ij}^d + \nu_{ij}^d + \mu_{ij}^d - \pi_t & \text{für } : (i,j) \in ext\widetilde{A}^d \text{ und } \exists t \in \mathcal{T} : (i,j) \in extA^d(t) \\ c_{ij}^d + \nu_{ij}^d + \mu_{ij}^d & \text{für } : (i,j) \in ext\widetilde{A}^d \text{ und } \neg\exists t \in \mathcal{T} : (i,j) \in extA^d(t) \\ c_{ij}^d + \nu_{ij}^d & \text{für } : (i,j) \notin ext\widetilde{A}^d \end{cases} \tag{9.12}$$

Das beschriebene Problem ist pro Depot als einfaches Flussproblem in polynomieller Zeit lösbar. Die Kantenmenge der Lösung entspricht pro Depot einem gültigen Bündel an Umlaufplänen. Sie wird der Kantenmenge $extA'^d$ hinzugefügt und damit im weiteren Lösungsprozess zum Pricing neuer Dienste bzw. (robuster) Umläufe verwendet.

Um bei der Kantenauswahl eine zu starke Fokussierung auf eine (kostenminimierende) Umlauflösung zu vermeiden, kann das Pricing neuer Dienste bzw. (robuster) Umläufe angepasst werden, wie im in Anschluss an die Performanzdiskussion folgenden Abschnitt *Pricing neuer Dienste und Umläufe* beschrieben wird.

Einfluss von Rn, NR und FS auf die Lösungsperformanz: Tabelle 9.2 fasst den Einfluss der verschiedenen Pricing-Strategien auf die Lösungsqualität und Laufzeit im Vergleich zum Lösen ohne Kantengenerierung zusammen. Dabei ist zu beachten, dass bereits ohne den Einsatz von Kantengenerierung die zuvor in Abschnitt 8.3.3 ausgewiesenen Beschleunigungsstrategien im Lösungsverfahren verwendet werden. So erfolgt das Pricing pro Column Generation Iteration parallelisiert, es werden Dominanztests verwendet und die Anzahl der in den ressourcenbeschränkten-Kürzeste-Wege Problemen jeweils an einem Knoten zu betrachtenden Pfade (States) wird eingegrenzt. Es werden exemplarische Ergebnisse zu integrierter Planung (INT) und integrierter Planung mit Zeitfenstern von ± 4 Minuten ($INT\text{-}TW4$) aufgeführt. Die verwendeten Testinstanzen sind in Abschnitt 11.1.1 beschrieben. Neben der Planungsmethode werden jeweils angegeben die Pricing-Strategie zur *Kantenauswahl*, die Anzahl benötigter Column Generation Iterationen ($\#It$) sowie erzeugter Spalten ($\#Cols$), die für das Lösen des Master-Problems aufgewendete Zeit (Cpu_Ma) sowie die für die Pricing-Probleme (d.h. für Kantenauswahl und Diensterstellung) aufgewendete Zeit (Cpu_Pr) in Stunden, Minuten und Sekunden (hh:mm:ss). Eine Verringerung oder Erhöhung der Zeiten wird jeweils ergänzend in Relation zum Verfahren ohne Kantengenerierung angegeben ($\Delta\%$). Zusätzlich werden die berechnete Lower Bound (LB) und die Anzahl Fahrzeuge bzw. Dienste in der besten zulässigen Lösung aufgeführt. Die zur Berechnung verwendeten Kostensätze entsprechen dabei den im späteren Abschnitt 11.1.3 beschriebenen Einstellungen.

Zunächst einmal verdeutlicht die Tabelle die Unterschiede zwischen integrierter Planung und integrierter Planung mit Zeitfenstern im Hinblick auf sowohl Laufzeit als auch Lösungsqualität. Während die Laufzeiten bei INT-TW massiv ansteigen, kann der Ressourceneinsatz deutlich gegenüber INT reduziert werden. Beim Lösungsverfahren der integrierten Planung mit Zeitfenstern ohne Kantengenerierung werden in der Regel mehr Column Generation Iterationen sowie erzeugte Spalten benötigt als bei integrierter Planung. Dabei ist anzumerken, dass im Pricing zur Diensterstellung beim Lösen der ressourcenbeschränkten-Kürzeste-Wege-Probleme der größten Instanz C weniger States verwendet werden als beim Lösen der beiden Instanzen A und B. Dies erklärt die im instanzenübergreifenden Vergleich geringeren Laufzeiten für das Pricing.

Tabelle 9.2: Performanz des integrierten Lösungsverfahrens (mit Zeitfenstern) mit und ohne zusätzliche Kantengenerierung

Methode	Kanten-auswahl	#It	#Cols	Cpu_Ma	Δ%	Cpu_Pr	Δ%	LB	#v	#d
				Testinstanz A						
INT	-	31	182.330	00:01:40	x	00:02:38	x	52.802.5	17	25
	R100	38	233.001	00:02:27	+47,0	00:01:27	-44,9	52.599.3	17	25
	NR	31	107.849	00:00:47	-53,0	00:00:42	-73,4	52.028,4	17	25
	FS	33	39.463	00:00:32	-68,0	00:00:47	-70,3	52.468,8	17	25
INT-TW4	-	69	574.172	00:52:00	x	00:38:46	x	50.062,2	15	25
	R200	39	264.081	00:08:38	-83,4	00:03:33	-90,8	48.894,0	15	25
	NR	63	339.863	00:11:47	-77,3	00:10:10	-73,8	49.747,5	15	25
	FS	49	57.265	00:05:33	-89,3	00:03:27	-91,1	49.852,6	15	25
				Testinstanz B						
INT	-	44	310.128	00:06:46	x	00:10:47	x	63.628,7	21	29
	R100	38	252.313	00:05:15	-22,4	00:04:15	-60,6	63.659,9	21	30
	NR	41	187.597	00:03:09	-53,4	00:03:50	-64,5	63.517,8	21	29
	FS	36	111.973	00:02:06	-69,0	00:02:39	-75,4	63.623,8	21	29
INT-TW4	-	73	597.176	02:31:06	x	02:21:26	x	57.671,2	18	26
	R200	59	460.542	00:42:30	-71,9	00:16:57	-88,0	58.007,1	18	28
	NR	35	244.132	00:26:20	-82,6	00:21:38	-84,7	57.605,9	18	26
	FS	39	73.716	00:21:51	-85,5	00:05:22	-96,2	57.567,5	18	26
				Testinstanz C						
INT	-	90	317.109	00:14:25	x	00:12:13	x	126.039,4	33	65
	R200	106	756.592	00:33:31	+132,5	00:10:04	-17,6	128.101,4	33	68
	NR	95	242.977	00:13:39	-5,3	00:05:04	-58,5	125.290,0	33	65
	FS	36	120.995	00:04:00	-72,3	00:01:47	-85,4	125.438,7	33	65
INT-TW4	-	101	777.445	06:15:58	x	01:33:10	x	122.679,7	32	62
	R400	40	332.633	00:56:52	-84,9	00:04:42	-95,0	126.394,7	32	67
	NR	65	391.330	01:40:38	-73,2	00:22:36	-75,7	123.622,8	32	62
	FS	43	118.855	00:48:44	-87,0	00:06:51	-92,6	123.110,6	32	62

Innerhalb der Planungsvarianten INT und INT-TW zeigt die Tabelle eine drastische Reduzierung der Laufzeiten beim Einsatz von Kantengenerierung, bei nahezu gleichbleibender Lösungsqualität im Vergleich zum Lösungsverfahren ohne Kantengenerierung. So können trotz (heuristischer) Einschränkung der Netzwerke mit dem ursprünglichen Lösungsverfahren vergleichbare untere Schranken sowie eine identische Anzahl Fahrzeuge bzw. Dienste berechnet werden. Dabei zeigt sich bei den beiden größeren Instanzen B und C der klare Vorteil der Strategien NR und FS gegenüber dem zufallsbasierten Vergleichsverfahren Rn zur Kantenauswahl. In der Tendenz ist bei diesen beiden Varianten eine schnellere Konvergenz im Verfahren bei weniger Column Generation Iterationen und vor allem deutlich weniger generierten Spalten zu erkennen. Dies führt nicht nur zu einer drastischen Laufzeitreduzierung im Pricing

zur Dienstgenerierung (sowie bei robuster Planung auch für die Umlaufgenerierung), sondern aufgrund der kleineren Modellgrößen auch zu einer deutlichen Laufzeitverringerung für das Lösen des beschränkten Master-Problems. Erwartungsgemäß schneidet die umfassenste Strategie FS am besten im Hinblick auf die Verringerung der Laufzeiten ab.

Pricing neuer Dienste und Umläufe

Das Pricing neuer Dienste und Umläufe kann bei Einsatz von Kantengenerierung in zwei Varianten durchgeführt werden. Diese unterscheiden sich hinsichtlich des jeweils genutzten Netzwerkumfanges und ihrer Aufgabe.

(A) Zum einen findet das Pricing von Diensten und (robusten) Umläufen gemäß des in diesem Abschnitt vorgestellten Verwendungszweckes von Kantengenerierung auf den eingeschränkten Netzwerken basierend auf der Kantenmenge $extA'^d$ je Depot-Fahrzeugtypkombination $d \in \mathcal{D}$ statt. Durch geschickte Kantenvorgaben zur Dienst- und Umlaufbildung soll so die Gesamtkonvergenz des Lösungsverfahrens zu einer guten Lösung beschleunigt werden.

(B) Zum anderen können zwischendurch in einzelnen Iterationen auch Dienste bzw. (robuste) Umläufe basierend auf den vollständigen Netzwerken erstellt werden, d.h. unter Einbezug der Kantenmenge $extA^d$ je Depot-Fahrzeugtypkombination $d \in \mathcal{D}$. So soll beispielsweise bei Einsatz der Pricing-Strategie FS ein stärkeres Einbeziehen der Robustheitsaspekte bzw. der Dienstsicht erreicht und eine zu starke Fokussierung auf den Modellteil zur (kostenminimierenden) Umlaufplanung für die Kantenwahl vermieden werden. Da die Dualwerte beim Lösen des beschränkten Master-Problems für alle Kanten des zugrunde liegenden Netzwerkmodells bestimmt werden, können die korrekten reduzierten Kosten für jede korrespondierende Kante in den einzelnen Pricing-Netzwerken angegeben werden. Erfolgt das Pricing auf vollständigen Netzwerken wird die in den erstellten Diensten bzw. Umläufen enthaltene Kantenmenge $extA''$ der Menge $extA'$ hinzugefügt und ab der nächsten Iteration als Erweiterung der eingeschränkten Netzwerke eingesetzt. Im Vergleich zum generell ausgeführten Multiple-Pricing bei dem pro Column Generation Iteration viele neue Spalten erzeugt werden (siehe für Details den späteren Abschnitt 11.1.3 über wesentliche Parametereinstellungen der Optimierungsverfahren), wird nur die Erstellung einer deutlich verringerten Anzahl neuer Dienste bzw. (robuster) Umläufe erlaubt, um die zusätzlich einzubeziehende Kantenmenge zu reduzieren.

Spalten bzw. Kantenmanagement

Wie beim Algorithmus zum integrierten Lösungsverfahren 8.3.1 in Abschnitt 8.3 beschrieben, können Spalten mit hohen reduzierten Kosten im Verfahrensverlauf ausgeschlossen werden. Im Vergleich dazu werden einmal aktivierte Kanten aber nicht wieder aus in den einzelnen Pricing-Netzwerken entfernt, so dass diese Netzwerke sukzessive wachsen. Wie zu Beginn des gegenwärtigen Abschnittes 9.4.3 erwähnt, existieren in den ersten Iterationen des Column Generation Verfahrens sehr viele mögliche Spalten bzw. Variablen mit negativen reduzierten Kosten, so dass ein hohes Potenzial für Laufzeitverbesserungen durch Beschleunigungstechniken vorhanden ist. In späteren Iterationen wird es allerdings schwieriger, eine Menge von zielfunktionswertverbessernden Spalten zu generieren. So findet kein endgültiger Ausschluß einzelner Kanten statt, da die jeweils durch einzelne Kanten repräsentierte Aufgabe bzw. Fahrt auch mit positiven reduzierten Kosten ein Teil eines zielfunktionswertverbessernden Dienstes oder Umlaufes sein kann. Darüber hinaus ist anzumerken, dass sich durch das Wachsen der Pricing-Netzwerke für das Lösen des (beschränkten) Master-Problem keine Laufzeitveränderungen ergeben, da die zu berücksichtigende Kantenanzahl beim Lösen der netzwerkflussbasierten Umlaufplanungs-Unterprobleme sowie für das Subgradientenverfahren (immer) gleich bleibt.

Terminierung

Das Column Generation Verfahren unter Einbeziehen von Kantengenerierung terminiert, wenn über mehrere Iterationen keine signifikante Verbesserung des Zielfunktionswertes durch neue Dienste bzw. Umläufe mehr erreicht werden kann, sowie wenn keine die Lösung verbessernden Dienste bzw. Umläufe mit negativen reduzierten Kosten mehr generiert werden. Letzteres ist aufgrund der eingeschränkten Netzwerkgrößen eher der Fall als beim Lösungsverfahren ohne Einsatz von Kantengenerierung. Wenn keine neuen Kanten mehr hinzugefügt werden, wird das Verfahren weiterhin solange durchgeführt bis die zuvor genannten Abbruchbedingungen eintreffen.

Ganzzahlige Lösung

Eine ganzzahlige Lösung wird unter Einbezug der vollständigen Kantenmenge erstellt, um in jedem Fall zulässige Lösungen garantieren zu können. Dabei dienen die in der letzten Iteration errechneten Lagrange-Multiplikatoren als Eingabe zur Bestimmung einer zulässigen Flusslösung für das Mehrdepot-Umlaufplanungsproblem auf dem durch Zeitfensterkanten erweiterten Netzwerken $extA^d$ für jede Depot-Fahrzeugtyp-

kombination $d \in \mathcal{D}$. Die weitere Ausgestaltung des Verfahrens entspricht dem in Abschnitt 8.3.4 beschriebenen Vorgehen.

9.5 Zusammenfassung und Implikationen

In diesem Kapitel wurde ein Lösungsansatz entwickelt, der in der Lage ist, durch das Berücksichtigen von Zeitfenstern bzw. Verschiebeintervallen für Servicefahrten Fahrplanungsentscheidungen in die robuste (und kosteneffiziente) integrierte Umlauf- und Dienstplanung einzubeziehen.

Als Vergleichsgrundlage für eine Ressourceneinsatzplanung mit Zeitfenstern wurde der Retiming-Ansatz von [Ahmadbeygi et al., 2010] aus dem Flugverkehr zum Zweck der Arbeit adaptiert. Dabei wurde vorgeschlagen, das auf bereits erstellten Umlauf- und Dienstplänen basierende nachträgliche Retiming-Verfahren als kostenneutrales Planungsverfahren einzusetzen, damit die Effekte der Fahrplananpassung in Bezug auf die Verringerung der Verspätungspropagation nicht durch ein einfaches Herausrechnen möglicher Verspätungen gegen zusätzliche Pufferzeit überschätzt werden. So werden in dieser Arbeit durch Retiming Pufferzeiten innerhalb von geplanten Dienststücken umverteilt ohne die Dauer des Ressourceneinsatzes zu ändern.

Mit dem integrierten Lösungsansatz zur robusten Umlauf- und Dienstplanung mit Zeitfenstern wurde ein Ansatz vorgestellt, der im Vergleich zu einem integrierten robusten Lösungsansatz mit anschließendem Retiming größere Freiheitsgrade aufweist. Da das Erstellen von Umlauf- und Dienstplänen sowie das geringfügige Verschieben und Fahrten simultan ausgeführt wird, verspricht der Ansatz eine weitere Erhöhung der Robustheit bei einer gleichzeitig möglichen Kostenreduzierung für den Ressourceneinsatzes.

Aufgrund explodierender Problemgrößen durch das Berücksichtigen von möglichen Fahrtenverschiebungen wurde neben einem problemspezifischem Preprocessing zur Einschränkung potenzieller Zeitfensterkanten im Optimierungsmodell mit der Kantengenerierung eine Erweiterung des integrierten Lösungsverfahrens für eine heuristische Einschränkung des Lösungsraumes vorgeschlagen. Diese Erweiterung führt durch die Einschränkung bzw. dynamische Ausweitung der zu betrachtenden Netzwerke zur Umlauf- und Diensterstellung zu schnelleren Pricing Phasen. Durch die gezielten Kantenvorgaben zur Spaltengenerierung, d.h. zur Generierung von Umläufen und Diensten, wird gleichzeitig auch eine schnellere Konvergenz des gesamten Lösungsprozesses ermöglicht. Für das Hinzufügen neuer Kanten wurden verschiedene Strategien miteinander verglichen, die bei drastischer Laufzeitreduktion zu vergleichsweise guten Lösungen führen. Das Vorgehen ist unabhängig der Zeitfensterbetrachtung und

daher auch für den integrierten Lösungsansatz zur Umlauf- und Dienstplanung aus Kapitel 8 einsetzbar.

Neben den zusätzlichen Freiheitsgraden für die Ressourceneinsatzplanung bietet die Anpassung des Fahrplanes als Schnittstelle zur Angebotsplanung (vgl. Abschnitt 2.1) einen Anknüpfungspunkt für weitere Forschung. Die Intention des vorgeschlagenen integrierten Ansatzes ist der Einsatz kleiner Verschiebeintervalle für eine geringe Modifikationen des Fahrplans mit großem Einfluss auf den Ressourceneinsatz – auch wenn der Lösungsansatz den Einsatz von großen Fahrtenverschiebungen ermöglicht und nicht ausschließt. Da die Fahrplanung selbst ein komplexes Planungsproblem darstellt, sollte allerdings bei größeren Eingriffen auch die Berücksichtigung von Passagieren bzw. Passagierbedürfnissen eine Rolle spielen und dann in das Lösungsverfahren einbezogen werden.

Kapitel 10

Ein hybrider Evolutionärer Algorithmus zum Erreichen Robuster Effizienz

In diesem Kapitel wird ein neuer hybrider evolutionärer Algorithmus für ein integriertes Umlauf- und Dienstplanungsproblem mit den Zielsetzungen Kosteneffizienz und Robustheit vorgeschlagen. Der Lösungsansatz verbindet mathematische Optimierung mit einem evolutionären Algorithmus. Der Ansatz basiert auf dem evolutionären Mehrzielalgorithmus Strength Pareto Evolutionary Algorithm 2 (SPEA 2) von [Zitzler et al., 2002] und den in den vorherigen Kapiteln 7 bis 9 entwickelten Lösungskomponenten zur gemeinsamen Berücksichtigung von Robustheit und Kosteneffizienz in der Ressourceneinsatzplanung. Durch diese Kombination kann im Lösungsverlauf eine breite Menge von Lösungen generiert werden, die die Pareto-optimale Menge von Lösungen bezüglich der beiden Zielsetzungen approximiert (siehe Abbildung 10.1). Der entwickelte Ansatz ist für eine im Vergleich zu den vorherigen Kapiteln vereinfachte Problemstellung der Umlauf- und Dienstplanung konzipiert und dient im Rahmen der Arbeit als Machbarkeitsstudie für eine andere Herangehensweise zum Erstellen (unterschiedlich) robust und kosteneffizient ausgeprägter Ressourceneinsatzpläne.

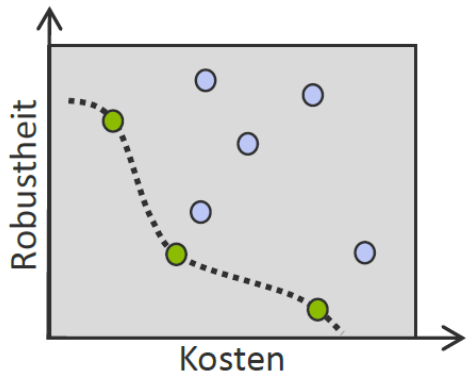

Abbildung 10.1: Pareto-optimale Lösungsmenge bezüglich Robustheit und Kosteneffizienz

Grundidee

Bei evolutionären Algorithmen (EA) werden mit Reproduktion, Variation und Selektion wesentliche Prinzipien der Evolutionstheorie nachgeahmt und zur Lösungsfindung benutzt. Der Evolutionsprozess wird ausgehend von einer initialen Population von Individuen nachgestellt. Dabei repräsentiert jedes Individuum eine potenzielle (zulässige) Lösung des betrachteten Optimierungsproblems und ihr Lösungswert wird als Fitness des Individuums bezeichnet. Besonders fitte Individuuen dürfen sich im Verfahren reproduzieren und dadurch neue Lösungen generieren. Für weiterführende Informationen und beispielhafte Anwendungsfälle sei an dieser Stelle auf [Dréo et al., 2006] verwiesen. Werden mehrere Ziele verfolgt, eignet sich der Einsatz Evolutionärer Mehrzielalgorithmen (MOEA). Diese können nicht sicherstellen, dass eine Menge von Pareto-optimalen Lösungen berechnet wird, sie sind jedoch in der Lage, eine Menge nicht-dominierter Lösungen zu finden, die die Pareto-optimale Menge von Lösungen approximiert (vgl. [Laumanns et al., 2002]). Für einen umfassenden Überblick über den aktuellen Stand der Forschung zu MOEAs inklusive den vielfältigen Anwendungsfeldern sei auf [Zhou et al., 2011] verwiesen. Exemplarische Anwendungen finden sich auch in [Gandibleux et al., 2004].

Im SPEA 2 nach [Zitzler et al., 2002] können im Lösungsverlauf potenzielle Lösungen so erstellt, modifiziert und selektiert werden, dass sich der Pareto-optimalen Menge angenähert wird. Im Verfahren wird zur Bewertung einzelner Lösungen deren Dominanz untereinander (vgl. Abschnitt 3.1.2), sowie ein Schätzmaß über die Fülle von Lösungen in deren Umkreis verwendet. Darüber hinaus wird mit Hilfe eines Archivs im Lösungsverlauf eine externe Population elitärer Individueen angelegt und gespeichert, um sicherzustellen, dass vielversprechende Lösungen im Verlauf nicht verloren gehen. Die Verknüpfung des SPEA 2 mit den Problemspezifika dieser Arbeit zu einem hybriden Ansatz für die Ressourceneinsatzplanung wird im Folgenden in diesem Kapitel erläutert. Dabei besteht eine Herausforderung darin, den mehrschichtigen Ressourceneinsatzes geeignet im Verfahren abzubilden.

Aufbau des Kapitels

Nach der Definition des betrachteten Planungsproblems in Abschnitt 10.1, gibt Abschnitt 10.2 eine Übersicht über bestehende Ansätze, die evolutionäre Mehrzielalgorithmen in der Problemdomäne Transport und Verkehr einsetzen. In Abschnitt 10.3 wird die gewählte Modellierung vorgestellt. Abschnitt 10.4 diskutiert das entwickelte Lösungsverfahren. Abschnitt 10.5 fasst das Kapitel zusammen und zeigt Implikationen auf.

10.1 Problemdefinition 2-SDIVCSP

Im Vergleich zu den vorangegangenen Kapiteln wird eine Problemstellung betrachtet, bei der nur im Depot Ablöse- und Pausenmöglichkeiten vorgesehen sind. Aus planerischer Sicht sind dann Fahrer an Fahrzeuge gebunden und Teilumläufe müssen so geplant werden, dass bei ihrer Ausführung Pausenzeiten für Dienste eingehalten werden können. Somit entsprechen Teilumläufe in der Planung einzelnen Dienststücken. Da der Fokus dieses Kapitels auf der Ausgestaltung des wesentlichen hybriden Lösungsprinzipes liegt, wird darüber hinaus nur der Eindepotfall diskutiert, wobei auf mögliche Erweiterungen für den Mehrdepotfall hingewiesen wird. Das *integrierte Umlauf- und Dienstplanungsproblem für den Eindepotfall mit der expliziten bikriteriellen Zielsetzung Kosteneffizienz und Robustheit* (engl. *bi-objective cost-efficient and robust single-depot integrated vehicle and crew scheduling problem, 2-SDIVCSP*) kann dann wie folgt definiert werden:

Gesucht werden Umlauf- und Dienstpläne, die für eine gegebene Menge an Servicefahrten und Leerfahrtmöglichkeiten, sowie eine gegebene Depot-Fahrzeugtypkombination bezüglich der Zielsetzungen Kosteneffizienz und Robustheit in Bezug auf (betrachtete) Verspätungsszenarien Pareto-optimal sind. Umlauf- und Dienstplan sollen jeweils zulässig und zueinander kompatibel sein, d.h. jede Servicefahrt des Fahrplans wird von genau einem Umlauf- und Dienst ausgeführt und jede zusätzliche Fahrzeugaktivität im Umlaufplan wird eindeutig von einem Dienst im Dienstplan abgedeckt.

10.2 Literaturübersicht – Evolutionäre Mehrzielalgorithmen in Fahrzeugeinsatz- und Personalplanung

In der Fahrzeugeinsatz- bzw. Personalplanung in den Domänen Busverkehr und Flugverkehr existieren einige Ansätze, die evolutionäre Mehrzielalgorithmen für verschiedenartige Problemstellungen verwenden. Dieser Abschnitt diskutiert die Ansätze im Hinblick auf den Einsatzbereich, die Zielstellung, das gewählte Lösungsverfahren, die Gestaltung der Lösungsrepräsentation, sowie die genutzten Selektions- bzw. Variationsoperatoren innerhalb des Verfahrens. Dabei werden Anknüpfungspunkte für diese Arbeit herausgestellt.

Busverkehr

[Lourenço et al., 2001] stellen für eine umlaufbasierte Dienstplanung ein mehrkriterielles Optimierungsproblem mit sechs Zielsetzungen auf. Dabei werden die Dienstkosten, die Anzahl der benötigten Dienste, die Anzahl der durch den Fahrereinsatz nicht abgedeckten sowie die Anzahl der mehrfach abgedeckten Dienststücke, die Anzahl von Diensten mit nur einem Dienststück, sowie die Anzahl von Umlaufwechseln minimiert. Bei der aufgestellten Set-Covering Formulierung handelt es sich um ein reines Auswahlproblem geeigneter Dienste. Voraussetzung ist daher, dass eine Menge zulässiger Dienste als Eingabe zur Verfügung steht. Deshalb kann als Lösungsrepräsentation ein binärer Vektor eingesetzt werden, dessen Position i beschreibt, inwiefern ein Dienst i für die Lösung ausgewählt ist. Im Lösungsverlauf erfolgt die Selektion zweier Individuen durch binäre Turnierselektion bei der zufällig eines der verschiedenen Zielkriterien zur Fitnessbewertung der Individuen bestimmt wird. Ein während des Verfahrens benutztes Archiv enthält alle im Verlauf gefundenen nicht dominierten Lösungen. Zur Erzeugung der Nachkommen für die beiden ausgewählten Lösungen werden zwei unterschiedliche Variationsoperatoren getestet. Ein Zwei-Punkt-Crossover und ein problemspezifischer Perfect-Offspring-Crossover. Bei diesem wird ein Unterproblem erneut gelöst, bei dem die abzudeckenden Dienststücke dem Ursprungsproblem entsprechen und als Variablen/Spalten alle Dienste aus beiden Elternteilen übernommen werden. Die Lösung des Untersproblems bildet den Nachkommen. Dabei werden kleine Instanzen exakt gelöst und für große Instanzen wird ein GRASP-Ansatz nach [Feo und Resende, 1995] verwendet. Als Mutationen werden zufällig bestimmte Dienste für die Lösung ausgewählt oder von dieser ausgeschlossen. Das entwickelte Verfahren wird auf sieben Instanzen mit 200 bis 350 abzudeckenden Dienststücken und ca. 19000 bis 74000 zuvor bereits erstellten Diensten getestet und mit der kostenoptimalen LP-Lösung der ursprünglichen Problemformulierung verglichen. ***Fazit:*** Die Ergebnisse zeigen jeweils erreichbare Verbesserungen bezüglich einzelner Zielkriterien. Dabei liefert der Einsatz des problemspezifischen Variationsoperators bessere Lösungen, allerdings zu deutlich höherer Laufzeit und für die größten Instanzen werden keine Lösungen berechnet. Er ist jedoch nur aufgrund erlaubter Unter- und Überdeckungen der Restriktionen (Dienststückmenge) brauchbar. Durch die zufallsbasierte Auswahl der betrachteten Zielfunktionsgröße zur Fitnessbewertung ist darüber hinaus zwar eine Diversifikation im Lösungsverfahren möglich, allerdings wird keine Information über die Güte der so approximierten Lösungsfronten gegeben. Da jeweils immer nur eine Zielfunktionsgröße und keine Pareto-basierte Fitnessbewertung eingesetzt wird, scheint es schwierig, dominierende Lösungen zu finden. Darüber hinaus erscheint eine explizite Formulierung aller möglichen Dienste im Vorfeld für

größere Problemstellungen nicht machbar. Es bleibt offen, inwiefern die mittels eines iterativen Verfahrens erzeugten Dienste die Lösungssuche vor dem evolutionären Verfahren bereits einschränken, da sie lediglich im Hinblick auf Dienstregelerfüllung nicht jedoch nach den verschiedenen Zielkriterien erstellt wurden.

[Steinzen, 2007] betrachtet ein bikriterielles Dienstplanungsproblem bei dem Fahrer und Fahrzeug eine Einheit bilden. Die Zielsetzungen bestehen darin, einen Dienstplan zu minimalen Kosten zu erstellen und gleichzeitig eine hohe Regelmäßigkeit gegenüber einem Referenzplan zu erreichen. Diese wird gemessen als die Anzahl unterschiedlicher Dienstelemente bezogen auf den gegebenen Referenzplan. Die Formulierung des Planungsproblem entspricht einem bikriteriellen Set-Partitioning Problem. Im Lösungsverfahren wird jedoch zunächst eine LP-relaxierte Set-Covering Formulierung mit Hilfe von Column Generation gelöst. Die generierten Dienste werden benutzt, um eine ganzzahlige Lösung zu finden. Dafür werden verschiedene multikriterielle Metaheuristiken getestet. Der Autor vergleicht das Verfahren SPEA2 und die Mehrziel-Varianten von Tabu Search, Simulated Annealing sowie Ant Colony Optimization insbesondere in Bezug auf die Qualität der approximierten Pareto-Front. Für die Implementierung des SPEA2 wird eine binäre Repräsentation genutzt, die angibt, ob ein bereits zuvor erstellter Dienst in der Lösung ausgewählt ist oder nicht (siehe [Lourenço et al., 2001]). Die Rekombination erfolgt für zwei zufällig bestimmte Lösungen durch einen Fusions-Operator nach [Beasly und Chu, 1996]. Die betrachteten Instanzen haben 300 bis 800 Servicefahrten, für die zwischen 60.000 und 160.000 Dienste generiert werden. Die Laufzeit der Metaheuristiken ist auf maximal zehn Minuten beschränkt. ***Fazit:*** Die Ergebnisse zeigen, dass die mit der Implementierung des SPEA2 errechneten Lösungen in der vorgebenen Laufzeit (fast) alle Lösungen der anderen Metaheuristiken dominieren und zum Abbilden der zweiten Zielsetzung am besten geeignet ist. Dabei werden keine Informationen über die Reparatur möglicher unzulässiger Lösungen gegeben. Es können weit verteilte, diversifizierte Lösungsfronten erstellt werden. Die Kosteneffizienz einer reinen Kostenminimierung kann aber deutlich nicht erreicht werden. Die für die Metaheuristiken zur Auswahl stehenden Dienste werden nur basierend auf Kostenminimierung erstellt. Um die erreichbare Regelmäßigkeit bzw. im Rahmen dieser Arbeit die Robustheit zu erhöhen, sollte das zusätzliche Zielkriterium bei der Dienstgenerierung auch Berücksichtigung finden bzw. die Dienstgenerierung im (evolutionären) Verfahren mit abgebildet werden.

[Moz et al., 2009] formulieren ein bikriterielles Dienstreihenfolgeproblem für den Fahrereinsatz als nicht lineares Programm. Als Ziele werden verfolgt eine Minimierung der Überstunden je Fahrer und eine Minimierung der Anzahl derjenigen Fahrer, deren Arbeitszeit geringer als vertraglich festgelegt ist. Das Optimierungsproblem

wird unter anderem mit dem SPEA2 gelöst. Dabei werden als Lösungsrepräsentation zwei Chromosomen verwendet. Ein Chromosom stellt als Liste eine Permutation aller zuvor erstellten Dienste und das andere als Liste eine Permutation aller Fahrer dar. Um die zahlreichen Restriktionen der Dienstreihenfolgeplanung im Lösungsverfahren berücksichtigen zu können, wird mit einem problemspezifischen Dekoderansatz gearbeitet. Mit Hilfe einer Heuristik wird aus den beiden Chromosomen pro betrachtetem Individuum in mehreren Iterationen ein zulässiger Dienstreihenfolgeplan erstellt. Die Rekombination zweier Individuen erfolgt durch den Order-Operator nach [Goldberg, 1989]. Als Mutation werden die Werte zweier zufällig in den Chromosomen ausgewählten Positionen miteinander vertauscht. Das Verfahren wird auf zwei Instanzensets getestet. Die Instanzen umfassen 350 bis 720 Dienste bei 45 Fahrern für einzelne Tage mit 80 bzw. 100 Servicefahrten über einen Planungszeitraum von einem Monat. Informationen zu Lösungslaufzeiten werden nicht gegeben. ***Fazit:*** Beim Verfahren werden begünstigt durch den heuristischen Dekoderansatz viele nahezu gleichwertige zulässige Lösungen erstellt. In den approximierten Lösungsfronten sind nur sehr wenige nicht dominierte Lösungen vorhanden (zwischen ein bis drei). Dennoch könnte der Einsatz eines Dekoders auch im Rahmen dieser Arbeit helfen, die umfassenden Problemspezifika geeignet im Lösungsverfahren abzubilden. Dabei ist jedoch eine Diversifikation der Lösungen im Verfahren zu beachten.

Flugverkehr

[Moudani et al., 2001] stellen ein bikriterielles Dienstreihenfolgeproblem für Crews von Flugzeugen auf. Die Zielsetzungen bestehen darin, die Kosten für den Personaleinsatz zu minimieren und gleichzeitig die Zufriedenheit der Crews zu maximieren. Die Problemformulierung entspricht einem Zuordnungsproblem der Crews zu zuvor festgelegten anonymen Diensten. Als Lösungsverfahren wird ein MOEA umgesetzt. Dabei wird als Lösungsrepräsentation eine Kodierung verwendet, die jedem Dienst i – gleichbedeutend mit der Position i im Chromosom – eine konkrete Crewnummer zuordnet. Die Auswahl zweier Individueen wird durch Rouletteselektion getroffen. Im Verfahrensverlauf wird ausgehend von mit einer Eröffnungsheuristik erstellten Lösungen, die allein die Zufriedenheit maximieren, versucht Lösungen zu generieren, die die Kosten für den Personaleinsatz weiter reduzieren. Neben einem nicht näher ausgewiesenen Crossover-Operator wird ein problemspezifischer Mutations-Operator verwendet. Ausgerichtet an der durchschnittlichen Arbeitsbelastung aller Crews der betrachteten Lösung werden zwei Gruppen gebildet. Crews mit geringerer und solche mit höherer Arbeitsbelastung als der Durchschnitt. Eine zufällig ausgewählte Crew-zu-Dienst Zuordnung wird durch die Zuordnung des Dienstes zu einer Crew

aus der Gruppe mit geringerer Belastung ersetzt, falls die neue Zuordnung zulässig ist. In einem Archiv werden alle nicht dominierten Lösungen gesammelt. Als Testinstanz wird eine Problemstellung mit 75 Crews und 275 Pairings verwendet. ***Fazit:*** Der Crossover-Operator trägt trotz höheren Zeitaufwandes nicht dazu bei, in einem gleichen Maß wie der problemangepasste Mutationsoperator die Lösungen zu verbessern. Für die Verbesserungen ist anzumerken, dass durch das evolutionäre Verfahren zwar die Kosten im Vergleich zur Ausgangslösung leicht reduziert werden und eine gleichmäßigere Verteilung der Arbeitsbelastung erfolgt, allerdings wird der Anteil nicht zufriedener Crews drastisch erhöht, da im Verfahrensverlauf keine Paretobasierte Fitnessbewertung erfolgt.

[Burke et al., 2010] betrachten ein bikriterielles Umlaufplanungsproblem mit Zeitfenstern (siehe auch Abschnitt 3.4). Als Zielsetzung soll die Robustheit für gegebene Flugzeugumläufe durch nachträgliche Anpassungen des Fahrplans bzw. der Umläufe erhöht werden. Zum Lösen wird ein MOEA mit einem lokalen Suchverfahren zu einem hybriden Verfahren kombiniert. Im Anschluss an die evolutionären Operationen sollen durch das lokale Suchverfahren die dann dekodiert betrachteten Lösungen zusätzlich im Hinblick auf die Zielsetzungen verbessert werden. Als Lösungsrepräsentation des MOEA werden die Flugsequenzen aller Umläufe als Permutationskodierung in Form einer Adjazenzliste hinterlegt. Damit die Permutationen zulässig sind, wird innerhalb der Kodierung der jeweils letzte Flug eines Umlaufes mit dem ersten Flug eines anderen verknüpft. Die Auswahl zweier Individuen wird durch Rouletteselektion getroffen. Die Fitnessbewertung der Individuen basiert dabei auf einer Kombination der (unterschiedlich) gewichteten Zielsetzungen. Die Rekombinationen der ausgewählten Lösungen erfolgt durch einen Ein-Punkt-Crossover. Da so aber nicht zwangsläufig Permutationen erzielt werden, wird zusätzlich eine Reparatur-Heuristik eingesetzt. Als Mutation werden zufällig zwei Indizes bestimmt, deren Werte getauscht werden. Im Verfahrensverlauf werden nicht dominierte Lösungen in einem Archiv gespeichert. Das Verfahren wird auf acht Testinstanzen mit 150 bis 500 Flügen und 5 bis 15 Flugzeugumläufen getestet. ***Fazit:*** Die Robustheit kann bezüglich der Zielkriterien sichtbar erhöht werden. Die gewählte Kodierung der Lösungen mittels einer Adjazenliste scheint auch für diese Arbeit geeignet, um Fahrten- bzw. Aufgabenketten im Lösungsverlauf abbilden und variieren zu können. Allerdings führt der einfache Rekombinationsoperator zu einem aufwändigen Reparaturprozess. Das lokale Suchverfahren wird nur in geringem Umfang eingesetzt, was laut Autoren gleichzeitig zu einer besseren Approximation der Pareto Fronten und geringerem Laufzeitaufwand führt.

Implikationen für diese Arbeit

In der Literatur werden zum einen kaum mehrkriterielle Problemstellungen bei unterschiedlichen Ressourcenebenen adressiert. Das Einbeziehen von Robustheitsaspekten wird nur in einem Fall im Flugverkehr auf Umlaufebene betrachtet. Zum anderen beschränken sich die Ansätze zum Großteil darauf, durch den evolutionären Prozess Auswahlprobleme bereits vorgenerierten Ressourceneinsatzes abzubilden. Dabei werden häufig erst bei der Zusammenstellung der bereits erstellten Umläufe, Dienste oder Dienstreihenfolgen unterschiedliche Zielsetzungen berücksichtigt, nicht aber schon bei der Generierung der einzelnen Ressourceneinsätze. Um keines der Ziele Kosteneffizienz und Robustheit bei der Planerstellung zu vernachlässigen, sollte in dieser Arbeit deshalb untersucht werden, inwiefern im evolutionären Lösungsprozess eine zielgerichtete Generierung der einzelnen Ressourceneinsätze mit abgebildet werden kann. Dabei ist herauszuarbeiten, auf welcher Ebene problemunspezifische evolutionäre Operationen und auf welcher zuvor entwickelte Lösungskomponenten eingesetzt werden können. Dekoderbasierte Ansätze ermöglichen es, Problemspezifika im evolutionären Lösungsprozess abzubilden. Vor dem Hintergrund der komplexen Problemstellung mit den zahlreichen Restriktionen zur Dienstbildung und verschiedenen Ressourcenebenen, sowie im Hinblick auf eine zielgerichtete Erstellung bzw. Bewertung der zu generierenden Individuen sollte in dieser Arbeit die evolutionäre Lösungssuche mit einem Dekoderansatz unterstützt werden. Dabei ist eine geeignete Form der Lösungsrespräsentation des mehrschichtigen Ressourceneinsatzes zu finden, die es ermöglicht sowohl zufallsbasierte, problemunspezifische evolutionäre Operationen als auch problemberücksichtigende Verfahrensschritte einzusetzen. Um eine Sequenz von miteinander verketteten, auszuführenden Aufgaben bzw. Fahrten im evolutionären Prozess abzubilden, scheint eine Adjazenzliste eine brauchbare Teil-Repräsentation einer Lösung zu sein. Überdies zeigt die Literatur in der Regel beim Einsatz des SPEA2 bzw. Pareto-basierten Fitnessbewertungen der Individuen eine gute und breit gefächerte Approximation der Lösungsfronten, weshalb SPEA2 als grundlegendes evolutionäres Verfahren in dieser Arbeit Verwendung finden sollte.

10.3 Modellierung 2-SDIVCSP

Das mathematische Modell für die in Abschnitt 10.1 beschriebene Problemstellung wird auf Basis eines Kompatibilitätsgraphen $\overline{G} = (N_T, A_y)$ nach [Gaffi und Nonato, 1999] formuliert. Die Knotenmenge $T \subset N_T$ repräsentiert hierbei die auszuführenden Servicefahrten und die Kanten A_y repräsentieren Warte- bzw. Leerfahrtaktivitäten, die kompatible Servicefahrten miteinander verknüpfen. Servicefahrten sind dann kom-

patibel, wenn sie zeitlich hintereinander (innerhalb eines Dienststückes) ausgeführt werden könnten. Die Knotenmenge N_T umfasst zusätzlich Quellknoten s und Senkeknoten t, die das Depot repräsentieren. Kantenverbindungen von der Quelle zu einem Serviceknoten stellen potenzielle Ausrückfahrten und Verbindungen von Serviceknoten zur Senke Einrückfahrten ins Depot dar. Jeder Pfad von s nach t entspricht einem Dienststück und besteht aus mindestens einer Servicefahrt. Da bei der Dienstplanung die Dauer von Dienststücken eine Rolle spielt, werden nur Fahrten innerhalb eines für die maximale Dienststückdauer passenden Zeitraumes miteinander verknüpft. Durch Kombination solcher Pfade können Dienste bzw. Umläufe erstellt werden. Die Wahl des beschriebenen Graphen als Grundlage für das Lösungsverfahren bietet eine einfache Prüfung der Zulässigkeit von Fahrtenverknüpfungen bei Durchführung der evolutionären Operatoren (siehe Abschnitt 10.4.3). Das Erstellen von zulässigen Dienst- und Umlaufplänen durch eine problemspezifische Dekoderkomponente erfolgt auf korrespondierenden TSN-Formulierungen, die denen in den vorherigen Kapiteln genutzten Netzwerkformulierungen entsprechen (siehe Abschnitt 10.4.4).

Es werden drei Mengen binärer Entscheidungsvariablen genutzt, um das mathematische Modell zu formulieren. Seien L die Menge aller zulässigen Umläufe und K die Menge aller zulässigen Dienste, dann repräsentiert die Variable z_l, ob Umlauf $l \in L$ für die Lösung ausgewählt wurde oder nicht, und x_k, ob Dienst $k \in K$ zur Lösung ausgewählt wurde oder nicht. Neben dieser pfadbasierten Formulierung von Umläufen und Diensten gibt die binäre Entscheidungsvariable y_{ij} für jede Kante $(i,j) \in A_y$ an, ob sie zur Lösung verwendet wird oder nicht. Ferner beschreibe $L(i,j) \subset L$ und $K(i,j) \subset K$ diejenige Menge von Umläufen bzw. Diensten, die die durch Kante (i,j) repräsentierte Aufgabe bedienen. Das *integrierte Umlauf- und Dienstplanungsproblem für den Eindepot-Fall mit der expliziten bikriteriellen Zielsetzung Kosteneffizienz und Robustheit* (engl. *bi-objective cost-efficient and robust single-depot integrated vehicle and crew scheduling problem, 2-SDIVCSP*) kann dann wie folgt formuliert werden.

2-SDIVCSP:

$$\sum_{l \in L} c_l z_l + \sum_{k \in K} f_k x_k \to \min \tag{10.1}$$

$$\sum_{l \in L} w_l z_l + \sum_{k \in K} v_k x_k \to \min \tag{10.2}$$

$$\text{s.t.} \sum_{i \in N_T} y_{ij} = 1 \qquad \forall j \in T \tag{10.3}$$

$$\sum_{i \in N_T} y_{ji} = 1 \qquad \forall j \in T \tag{10.4}$$

$$\sum_{l \in L(i,j)} z_l - y_{ij} = 0 \qquad \forall (i,j) \in A_y \tag{10.5}$$

$$\sum_{k \in K(i,j)} x_k - y_{ij} = 0 \qquad \forall (i,j) \in A_y \tag{10.6}$$

$$z_l \in \{0,1\} \qquad \forall l \in L \tag{10.7}$$

$$x_k \in \{0,1\} \qquad \forall k \in K \tag{10.8}$$

$$y_{ij} \in \{0,1\} \qquad \forall (i,j) \in A_y \tag{10.9}$$

Die Zielfunktion (10.1) minimiert Umlauf- und Dienstkosten. Mit der zweiten Zielfunktion (10.2) werden Verspätungspropagationen der Umläufe und Dienste minimiert. Die Verspätungspropagation wird dabei wie durch Formel 5.1 beschrieben berechnet. Nebenbedingungen (10.3) - (10.4) stellen sicher, dass jede Servicefahrt genau einmal abgedeckt wird. Nebenbedingungen (10.5) - (10.6) stellen sicher, dass gleiche Dienststückausprägungen verwendet werden und sorgen für einen kompatiblen Umlauf- und Dienstplan. Restriktionen (10.7) - (10.9) beschreiben den binären Charakter der Entscheidungsvariablen.

10.4 Lösungsverfahren

Da der SPEA2 problemunspezifisch agiert, müssen in das Verfahren problemspezifische Komponenten eingebaut werden. Insbesondere um die schwierigen und zahlreichen Restriktionen aus Umlauf- bzw. Dienstplanung berücksichtigen zu können, wird ein dekoderbasierter Ansatz gewählt. Der problemspezifische Dekoder wird eingesetzt, um während des Lösungsverfahrens anhand einer kodierten Lösung eine (vollständig) zulässige Lösung für das Umlauf- und Dienstplanungsproblem zu finden.

Abbildung 10.2 stellt die gewählte Problemdekomposition zur Erstellung eines konkreten Individuums dar. Als kodierte Lösung und somit als Eingabe für evolutionäre Operationen Rekombination und Mutation dient jeweils eine Menge an Dienststücken. Diese decken jeweils nach Individuum unterschiedlich alle zu bedienenden Servicefahrten genau einmal ab. Die evolutionären Operationen verändern also im Verfahrensverlauf allein die Ausgestaltung der einzelnen Dienststücke. Die (kodierten) Dienststücke werden dann mit Hilfe des Dekoders zu zulässigen und kompatiblen Umlauf- bzw. Dienstpläne zusammengesetzt. Da bei Anwendung der evolutionären Operationen Zulässigkeitskriterien für die Bildung von Dienststücken gegebenenfalls für die Problemstellung unzulässig werden können, wird vorher eine Dienststückreparatur durchgeführt. Unabhängig von der gewählten Problemdekomposition und der im Dekoder berechneten Lösung ist der evolutionäre Verfahrensablauf durch die vier

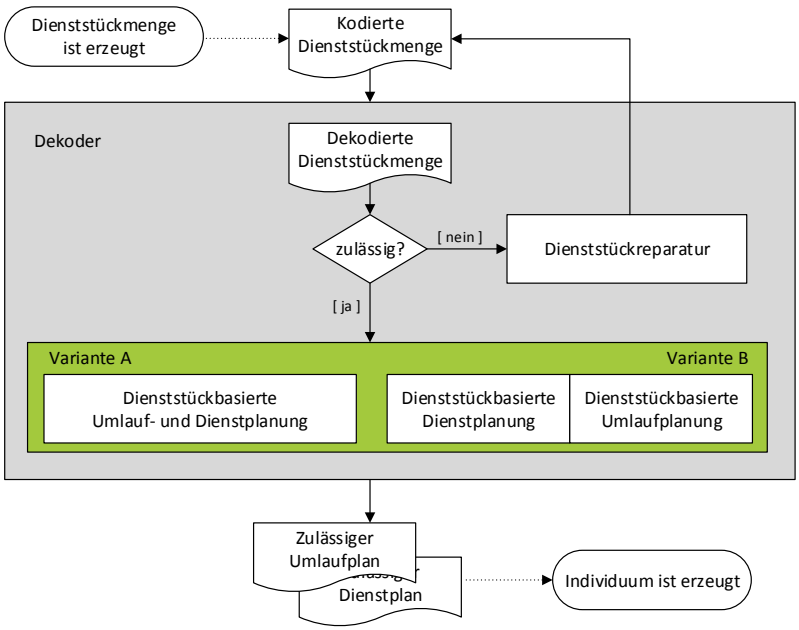

Abbildung 10.2: Dekoderbasierte problemspezifische Erzeugung eines Individuums

Parameter Archivgröße N, Populationsgröße P, Mutationswahrscheinlichkeit p_M und maximale Anzahl an Generationen G beeinflussbar.

Algorithmus 10.4 zeigt das gesamte problemangepasste Lösungsverfahren im Überblick. Es lässt sich grob in die folgenden drei Schritte einteilen: Initialisierung mit einer Startpopulation von Lösungen (1), dem eigentlichen evolutionärer Prozess zur Weiterentwicklung der Startpopulation (2), sowie gegebenenfalls einem abschließenden Lösen der im Archiv vorhandenen Individueen, falls zuvor zur Beschleunigung der Rechenzeit mit relaxierten Problemstellungen gearbeitet wurde (3). Im weiteren Verlauf dieses Abschnittes wird näher auf die einzelnen Komponenten Initialisierung des Verfahrens, Repräsentation einer Lösung, Variationsoperatoren, sowie den problembezogenen Dekoderansatz und die Terminierungskriterien des Verfahrens eingegangen.

10.4.1 Initialisierung

Um die evolutionsbasierte Suche nach geeigneten Lösungen beginnen zu können, wird eine initiale Population von Individuen benötigt. Es bieten sich mehrere Möglichkeiten an, eine solche Startpopulation zu erstellen:

- Einfache Konstruktionsheuristik (A)
- Eröffnungsverfahren angelehnt an [Ball et al., 1983] (B)
- Sequenzielle, teilintegrierte oder integrierte Verfahren (C)

Algorithmus 10.4.1 : Problemspezifischer SPEA2 für kosteneffiziente und robuste Umlauf- und Dienstplanung

(Schritt 1) **Initialisierung**
Setze Generationszähler g auf $g = 0$.
Erstelle (leeres) Archiv A_g.
Erstelle initiale Population P_g.
Zielfunktionswerte berechnen.

(Schritt 2) **Evolutionärer Prozess**
while *Abbruchbedingung nicht erfüllt* **do**
Fitnessberechnung aller Individuen in P_g und A_g.
Umgebungsbedingte Selektion: Kopiere alle nicht-dominierten Individuen aus P_g und A_g in das Archiv A_{g+1}.
if $|A_{g+1}| > N$ **then**
Entferne nicht-dominierte Individuen aus A_{g+1}.
else if $|A_{g+1}| < N$ **then**
Fülle A_{g+1} mit dominierten Individuen aus P_g und A_g auf.
Elternselektion: Suche durch binäre Turnierselektion Individuen aus A_{g+1} aus.
Variation: Erstelle durch Rekombination und Mutation die Population P_{g+1}.
Dekoder zur Umlauf- und Dienstplanung benutzen.
Zielfunktionswerte berechnen.
Erhöhe Generationszähler g auf $g = g + 1$.
end

(Schritt 3) **Ganzzahlige Lösung berechnen (optional)**
Berechne ganzzahlige Lösung für Individuen in A_g, falls bisher für Dienstplanungsteil nur mit LP-Relaxationen gearbeitet wurde.

(A) Die einfache Konstruktionsheuristik setzt so lange Fahrten zu Dienststücken zusammen, bis die maximal erlaubte Fahrtdauer erreicht ist. Jedes erstellte Dienststück entspricht dann einem Kurzdienst. Diese Kurzdienste werden dann Umläufen so zugeordnet, dass sie innerhalb der Umläufe zeitlich hintereinander ausführbar sind.

(B) Um die Lösungsqualität der Initiallösungen bezüglich der einzelnen Zielkriterien zu verbessern, kann bei der Dienststückkonstruktion ein Eröffnungsverfahren angelehnt an [Ball et al., 1983] verwendet werden. Dabei werden die einzelnen Servicefahrten gemäß ihrer Lage im Kompatibilitätsgraphen (vgl. Abschnitt 10.3) verschiedenen Hierarchieebenen zugeordnet. Im Verfahrensverlauf werden Fahrten der einzelnen Ebenen miteinander zu Dienststücken kombiniert, wobei so wenig Dienststücke wie möglich erstellt werden. Durch zufallsbasierte Gewichtung auf den Kanten zwischen den Ebenen können bei mehrfacher Anwendung verschiedene Partitionen der Servicefahrtenmenge zu Dienststücken generiert werden. Die erstellten Dienststücke

werden dann zu Umlauf- und Dienstplänen unter Verwendung der problemspezifischen Dekoderkomponenten 10.4.4 zusammengesetzt. Je nach parametrisierter Zielsetzung können auch hierbei unterschiedlich ausgeprägte Ressourceneinsätze erstellt werden.

(C) Darüber hinaus können mit den in den vorangegangenen Kapiteln 6 bzw. 7 und 8 beschriebenen sequenziellen, teilintegrierten und integrierten Verfahren erstellte Lösungen als Eingabe dienen, die ebenfalls mit unterschiedlichen Gewichtungen der Zielsetzungen berechnet wurden. Durch die verschiedenen Varianten zur Initialisierung können bereits im Vorfeld leicht unterschiedliche Teile des Suchraums als Ausgangspunkte adressiert werden, um durch eine breite Mischung an Ausgangsslösungen eine schnelle und weit gefächerte Approximation der pareto-optimalen Lösungsmenge im Verfahrensverlauf zu erreichen (vgl. z.B. [Shukla und Deb, 2007]).

10.4.2 Repräsentation

Jedes Individuum stellt eine zulässige Lösung des mit der Problemformulierung (10.1) - (10.9) abgebildeten Planungsproblems dar und besteht in dieser Arbeit aus den zur Lösung benutzten Dienststücken. Die Lösungsmenge der Dienststücke und somit jedes Individuum wird dabei repräsentiert durch ein Chromosom mit zwei Genen. Das erste Gen besitzt eine Permutationskodierung G_1 und das zweite Gen eine Binärkodierung G_2. Die *Permutationskodierung* erlaubt es, eine Reihenfolge von Elementen darzustellen. Die Permutation umfasst hier die Menge aller Servicefahrten. Die einzelnen Elemente der Permutation geben an, in welcher Reihenfolge Servicefahrten in den Dienststücken ausgeführt werden können. Um Vorgänger-Nachfolger Beziehungen abbilden zu können, wird eine Adjazenzliste mit natürlichen Zahlen gewählt. Vorgänger sind durch die jeweilige Position i angegeben und Nachfolger sind jeweils die Ausprägung an der Position i. Konkret repräsentiert also Ausprägung j an Position i, dass auf Servicefahrt i die Servicefahrt j innerhalb eines Dienststückes folgen kann. Da mit der Permutationscodierung mehr als ein Dienststück abgebildet wird, wird die zusätzliche *Binärkodierung* G_2 eingesetzt, um einzelne Dienststücke voneinander abzugrenzen. Der Vektor G_2 hat an Position i den Wert 1, wenn die Fahrt j an Position i in Vektor G_1 auf die Fahrt i im gleichen Dienststück folgt. Dienststücke werden im Gen G_2 durch die Ausprägung 0 voneinander getrennt. Dabei werden alle Dienststücke so verknüpft, dass ein Zyklus gebildet wird. Dafür wird die letzte Servicefahrt eines Dienststücks mit der ersten eines anderen Dienststückes verknüpft und im Vektor G_2 eine 0 eingetragen. Da nur die Dienststückmenge nicht aber die Reihenfolge der einzelnen Dienststücke zueinander in dieser Repräsentation abgebildet wird, können die Dienststücke beliebig verknüpft werden und es existieren eine

Vielzahl an Kombinationen, die dieselbe Dienststückmenge repräsentieren. Das in Abbildung 10.3 dargestellte Beispiel verdeutlicht die gewählte Kodierung.

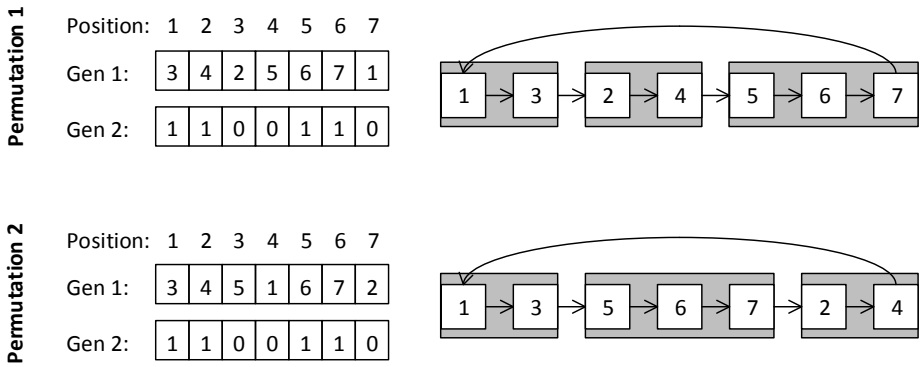

Abbildung 10.3: Repräsentation eines Individuums abbildbar als unterschiedliche Permutationen der gleichen Dienststückmenge

Die gewählte Repräsentation ermöglicht es, bei Bedarf weitere Gene hinzuzufügen. Mit einem zusätzlichen Gen könnten beispielsweise Dienststücke innerhalb eines Dienstes bzw. Umlaufes miteinander verknüpft werden. Dies wird im Rahmen dieser Arbeit allerdings nicht weiter verfolgt, da an dieser Stelle die Hybridisierung des Verfahrens und die Verknüpfung mit mathematischer Optimierung zum Tragen kommen soll (siehe Abschnitt 10.4.4). Mit problemspezifischem Modellen und Methoden sollen die Dienststücke so zu Umlauf- und Dienstplänen zusammen gesetzt werden, dass dabei die Zielkriterien Kosteneffizienz und Robustheit unter Einhaltung der Restriktionen zur Umlauf- und Dienstbildung berücksichtigt werden. So können im Vergleich zu evolutionären Operationen, die keine Zulässigkeitkriterien beachten, nachträglich durchzuführende Reparatur-Operationen vermieden werden.

Darüber hinaus könnte ein weiteres Gen eingeführt werden, um die Zuordnung von Dienststücken zu einer konkreten Depot-Fahrzeugtypkombination abzubilden. Diese Zuordnung müsste dann im Lösungsverfahren auch zusätzlich variiert werden und ist als Erweiterung des Verfahrens für eine Mehrdepot-Planung denkbar.

10.4.3 Variationsoperatoren

Im Lösungsverfahren werden ein *Rekombinations-* und ein *Mutationsoperator* zur Variation der Individuen eingesetzt. Beide Operatoren werden auf die Permutationskodierung G_1 angewandt, welche die Reihenfolge von Servicefahrten innerhalb von Dienststücken abbildet. Eine Anwendung von Variationsoperatoren auch auf die Binärkodierung G_2 ist an dieser Stelle nicht hilfreich, da die Binärkodierung lediglich

dafür eingesetzt wird, die ungeordnete Menge von Dienststücken voneinander abzugrenzen (siehe Abschnitt 10.4.2). Da die einzelnen Dienststücke in der Repräsentation nicht zeitlich sortiert sind, würden bei Variation der Binärkodierung häufig unkompatible Ketten von Servicefahrten entstehen, so dass im Nachhinein ein hoher Reparaturaufwand zu betreiben ist.

Rekombination

Als *Rekombinationsoperator* dient der von [Oliver et al., 1987] vorgeschlagene *Cycle-Crossover*. Dieser garantiert, dass nach der Rekombination die absoluten Positionen der abgebildeten Sequenz in den Nachkommen übernommen werden. Er eignet sich für die gewählte Adjazenzrepräsentation, da bei Anwendung immer eine gültige Permutation aller Servicefahrten in den Nachkommen erzeugt wird. Konkret werden die Werte an einer Position in G_1 entweder von Elternteil 1 oder Elternteil 2 vererbt. Wie in Abbildung 10.4 dargestellt, bleiben die absoluten Positionen einzelner Ausprägungen, die sich auf dem entstehenden Zyklus befinden, von Elternteil zu jeweiligem Nachkommen gleich. Für die anderen Positionen werden komplette Fahrtenketten von dem anderen Elternteil übernommen. Durch diese Art der Vererbung können zum einen Sequenzen von Servicefahrten in Dienststücken weitergereicht und zum anderen trotzdem Änderungen in der Ausprägung einzelner Dienststücke erzeugt werden.

Bei der Rekombination können aufgrund der gewählten Repräsentationsform von G_1 als Adjazenzliste statt einem, mehrere Zyklen entstehen (siehe ebenfalls Abbildung 10.4). Diese Eigenschaft lässt sich im Lösungsverlauf nutzen, indem bei den entstandenen Zyklen überprüft wird, ob sich die innerhalb eines Zyklus befindlichen Dienststücke zu größeren Fahrtensequenzen zusammenfassen lassen. Dafür wird unter Verwendung des Kompatibilitätsgraphen $\overline{G}$ geprüft, ob zwischen denen innerhalb eines Zyklus befindlichen Fahrten am Ende eines Dienststückes ($G_2 = 0$) eine Kantenverbindung zur in G_1 angegebenen Fahrt existiert, die Fahrten also kompatibel zueinander sind. Da die Prüfung unabhängig der tatsächlichen Fahrtendauern durchgeführt werden können Fahrtensequenzen mit zu langer Gesamtdauer entstehen. Diese werden mit der in Abschnitt 10.4.4 beschriebenen *Dienststückreparatur* problemspezifisch gekürzt. In der Abbildung 10.4 ist eine Zusammenführung der Fahrtensequenzen $(1, 3)$ und $(6, 7)$ im ersten Nachkommen bzw. $(1, 4)$ und $(5, 6, 7)$ im zweiten Nachkommen zu überprüfen. Werden Fahrtenfolgen zusammengefasst, muss nachträglich die (vererbte) Binärkodierung G_2 angepasst werden, so dass eine neu geschaffenen Verknüpfung von zwei Fahrten statt durch die Ausprägung 0 durch den Wert 1 repräsentiert wird. Um die gegebenenfalls noch vorhandene Zyklen in der Adjazenzrepräsentation der

Nachkommen wieder zu einer durchgängigen Adjazenzliste aller Fahrten aufzulösen, werden die Werte in G_1 in einem simplen Schritt angepasst. Beginnend an der ersten Position werden die weiteren Positionen des Gens anhand der Einträge besucht und jede Position dabei markiert. Weist ein Eintrag auf eine bereits besuchte Position (es liegt ein Zyklus vor) und die komplette Fahrtenmenge wurde noch nicht durchlaufen, dann wird der Eintrag so angepasst, dass auf die erste noch unmarkierte Position in G_1 verwiesen wird, die den Anfang eines Dienststücks repräsentiert. Von dieser ausgehend wird dann das Gen weiter durchlaufen.

Die Auswahl von Individuen zur Rekombination erfolgt wie von [Zitzler et al., 2002] für den SPEA2 vorgeschlagen durch eine binäre Turnierselektion geeigneter Individuen aus dem aktuellen Archiv. Zur Bewertung und Auswahl der Individueen werden Fitnesswerte berücksichtigt, die wie in Abschnitt 10.4.5 beschrieben berechnet werden.

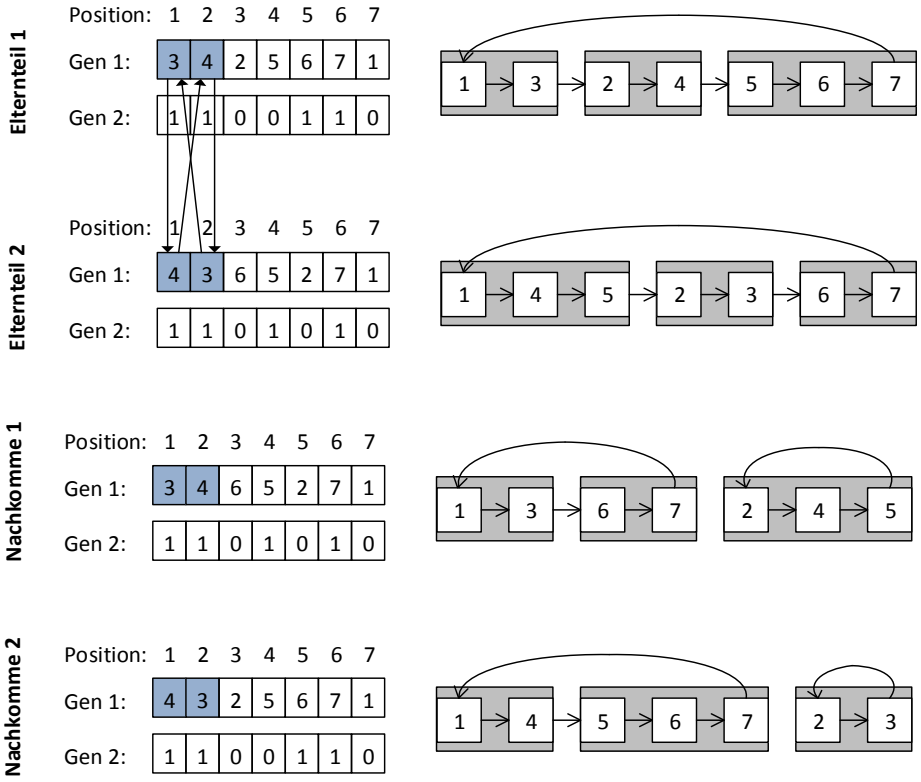

Abbildung 10.4: Cycle-Crossover-Operator zur Rekombination zweier Individuen und gegebenenfalls entstehende Zyklen, die pro Individuum zu einem Zyklus anzupassen sind

Mutation

Als *Mutationsoperator* wird ein zufälliger Tausch der Werte zweier Positionen i und j im Gen G_1 durchgeführt (siehe Abbildung 10.5). Danach muss die Kompatibilität der neu entstandenen Fahrtensequenzen geprüft werden. Die vertauschten Servicefahrtenfolgen sind weiterhin kompatibel, wenn für die Fahrtenfolge i und $G_1[i]$ eine Verbindung im Kompatibilitätsgraphen $\overline{G}$ existiert. In diesem Fall wird die Ausprägung $G_2[i]$ in Gen G_2 auf 1 gesetzt. Existiert keine Verbindung, wird $G_2[i]$ auf 0 gesetzt. Gleiches gilt für die Fahrtenfolge j und $G_1[j]$. Wie beim Variationsoperator werden durch die Mutation entstehende Zyklen darauf überprüft, ob die innerhalb des Zyklus vorliegenden Fahrtensequenzen zu längeren Dienststück(en) zusammengefasst werden können. Dies ist in der Abbildung 10.5 für die Fahrtenfolge $(2, 4, 5)$ der Fall. Im Beispiel erfolgt die Anpassung der Binärcodierung so, dass auf Servicefahrt 4 auch Fahrt 5 folgen darf. Über eine Mutationsrate p_M wird die Anzahl der zufälligen Tausche im Verfahrensablauf in Abhängigkeit der betrachteten Fahrtenanzahl T beeinflusst und ergibt sich als $\lfloor p_M * T \rfloor$.

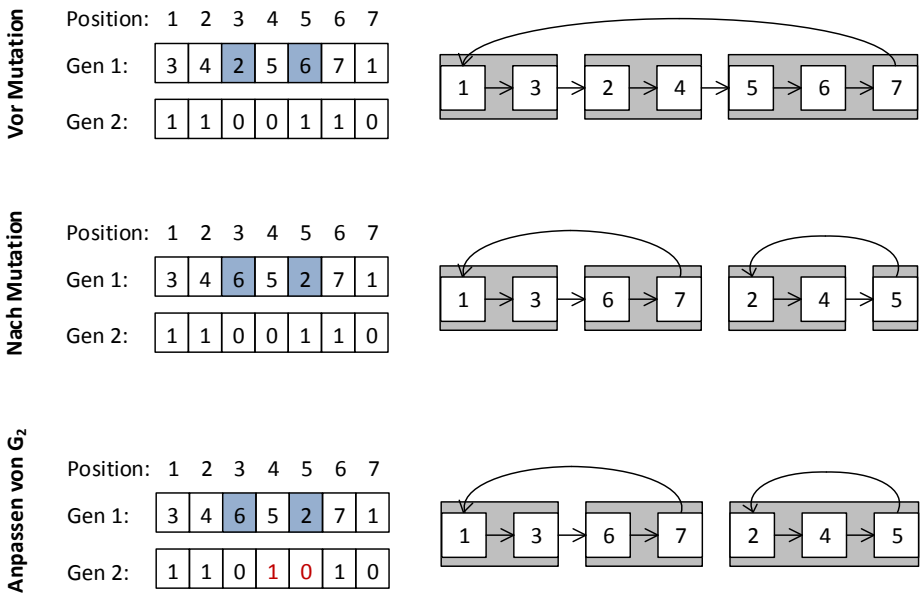

Abbildung 10.5: Zufälliger Tausch zweier Werte in der Adjazenzrepräsentation G1 als Mutationsoperator und anschließende Anpassung der Binärcodierung G2.

10.4.4 Dekoder

Die Lösungsmenge der Dienststücke muss noch zu zulässigen Dienst- und Umlaufplänen zusammengesetzt werden (siehe Abbildung 10.2). An dieser Stelle findet sich die Hybridisierung des Verfahrens wieder, indem zu den evolutionären Operatoren zusätzlich problemspezifische Lösungskomponenten basierend auf mathematischer Optimierung eingesetzt werden. Während die *Dienststückreparatur* mit einer einfachen Heuristik durchgeführt wird, werden die *Dienststückbasierte Dienstplanung* sowie *Umlaufplanung* mit Methoden der mathematischen Programmierung gelöst. Dafür werden zwei verschiedene Varianten vorgeschlagen: *Dekodervariante A* löst eine exakte Problemstellung und betrachtet dabei die dienststückbasierte Umlauf- und Dienstplanung gemeinsam. *Dekodervariante B* löst eine relaxierte Problemstellung und betrachtet dabei Dienstplanung und Umlaufplanung getrennt, um Laufzeiten bei sehr großen Problemstellungen zu reduzieren und den gesamten evolutionären Prozess in adäquater Zeit zu ermöglichen. Dabei verfolgen die Dienstplanung sowie die Umlaufplanung die beiden Ziele der Kostenminimierung und der Minimierung von Verspätungspropagation. Aufgrund der Verknüpfung über die gemeinsame Eingabemenge an Dienststücken können Dienstplanung und Umlaufplanung auch unabhängig voneinander durchgeführt werden. Die drei Komponenten Dienststückreparatur, sowie Dienstplanung und Umlaufplanung in den verschiedenen Varianten werden im Folgenden näher erläutert.

Dienststückreparatur

Eine *Dienststückreparatur* wird benötigt, da bei Variation der Individuen Dienststücke entstehen können, deren Länge die zulässige Maximaldauer überschreitet bzw. möglicherweise eine vorgegebene Mindestdauer unterschritten wird. Die Reparatur erfolgt durch ein einfaches heuristisches Verfahren, das aus vier Schritten besteht:

(1) Zunächst werden alle Dienststücke in drei Mengen unterteilt, je nachdem ob die Mindestdauer unterschritten ($Pows^-$), die Maximaldauer überschritten ($Pows^+$) oder keine Reparatur benötigt wird ($Pows$).

(2) Dann werden von den überlangen Dienststücken jeweils so lange Fahrten am Ende abtrennt, bis die Dienststücke eine zulässige Dauer besitzen und der Menge $Pows$ zugeordnet werden können. Die abgetrennten Fahrtensequenzen entsprechen jeweils neuen Dienststücken und werden anhand ihrer Ausprägung den drei Dienststückmengen zugeordnet. Solange noch Dienststücke der Menge $Pows^+$ existieren, wird dieser Schritt (2) wiederholt.

(3) Die durch Fahrtenabtrennungen neu entstandenen Dienststücke in den Mengen $Pows^-$ und $Pows$ werden dann versucht anderen Dienststücken zuzuordnen, wobei zunächst Anknüpfungsmöglichkeiten mit den anderen Dienststücken der Menge $Pows^-$ geprüft werden, dann mit denen der Menge $Pows$.

(4) Enthält die Menge $Pows^-$ noch Dienststücke, werden zunächst Anknüpfungsmöglichkeiten mit anderen Dienststücken der Menge $Pows^-$ und dann mit Dienststücken der Menge $Pows$ geprüft.

Bleiben nach Abschluss des Verfahrens unzulässige Dienststücke übrig, werden sie bei den im Folgenden beschriebenen Dekodervarianten zur Umlauf- und Dienstplanerstellung mit Strafkosten berücksichtigt.

Dekodervariante A - Exakte Problemstellung

Die mit dieser Dekodervariante abgebildete *Dienststückbasierte Umlauf- und Dienstplanung* entspricht in etwa den kombinierten Schritten der teilintegrierten Planung aus Kapitel 7. Der erste Schritt der teilintegrierten Planung zum Bestimmen eines zugrunde liegenden Umlaufplanbündels entfällt dabei, da das Bündel bereits implizit durch die erstellten Dienststücke vorgegeben ist. Die als Eingabe zur Verfügung stehenden Dienststücke werden gemäß den Zielsetzungen nach Robustheit und Kosteneffizienz zu einem zulässigen Umlauf- und Dienstplan zusammengesetzt.

Zum Lösen des *Dienstplanungsteiles* wird eine modifizierte Variante des in Abschnitt 7.5 beschriebenen Lösungsansatzes verwendet. Dabei erfolgt die Modellierung basierend auf Dienststücken, statt einzelne Dienstelemente bzw. Fahrzeugaktivitäten zu berücksichtigen. Bei der Initialisierung des Column Generations Verfahrens entspricht jedes Dienststück einem (Kurz-)Dienst. Für die Pricing Probleme werden Dienststück-basierte ressourcenbeschränkte Kürzeste-Wege-Probleme für das Erstellen von Diensten eingesetzt (statt Dienstelement-basierter). Der Aufbau der Dienstgenerierungsnetzwerke entspricht den in Abschnitt 7.5.2 beschriebenen Netzwerken mit dem Unterschied, dass Dienststück-Kanten statt Dienstelement-Kanten verwendet werden. Die Ressourcenverbräuche auf den Kanten umfassen kumuliert die jeweils durch die Dienstelemente des Dienststück vorgegebenen Verbräuche. Das Berechnen der genauen potenziellen Verspätungspropagation der Dienste basierend auf Dienststücken erfolgt analog der Phase II zum Erstellen robuster Umläufe bzw. der Dienststück-basierten Variante zur Umlaufgenerierung aus Abschnitt 7.6.2.

Zum Lösen des *Umlaufplanungsteiles* kann die in Abschnitt 7.6.1 beschriebene Modellierung zur globalen Flussdekomposition basierend auf Dienststücken verwendet werden. Die Formulierung R-GFD-VSP II Pow (7.53) - (7.56) wird dabei wie in Abschnitt 7.6.2 beschrieben gelöst. Im Rahmen des Column Generation Verfahrens

werden Pricing Probleme basierend auf Dienststück-basierten ressourcenbeschränkten kürzesten Wegen betrachtet. Um eine minimale Fahrzeuganzahl vorzugeben, wird in einem ersten Schritt das einfache Netzwerkflussproblem (7.57) - (7.60) gelöst, wobei sämtliche Kantenkosten auf 0 gesetzt werden und allein die Zirkulationsflusskante mit Fixkosten für den Umlaufeinsatz belegt wird.

Weitere Variation im Lösungsverfahren und damit eine zusätzliche Diversifikation der Lösungen im Hinblick auf die Approximation der Pareto-Front ist dadurch möglich, dass dieselbe Dienststückmenge unter Einbeziehen unterschiedlich gewichteter Zielsetzungen zu verschieden ausgeprägten Umlauf- und Dienstplänen zusammengestellt wird: fokussiert auf Kosteneffizienz, fokussiert auf Robustheit, sowie Kosteneffizienz und Robustheit gemäß der Parametervorgabe zur Steuerung in der Zielfunktion.

Dekodervariante B - Relaxierte Problemstellung

In dieser Variante werden zur Beschleunigung des gesamten Lösungsprozesses nur relaxierte Problemstellungen zur robusten und kosteneffizienten Umlauf- und Dienstplanung betrachtet. Im Einzelnen werden die *Dienststückbasierte Dienstplanung* und die *Dienststückbasierte Umlaufplanung* unabhängig voneinander gelöst.

Dienststückbasierte Dienstplanung: Zur Verringerung der Laufzeit wird allein die LP-Relaxation zur Dienstplanung betrachtet. Die Dienstplanung erfolgt also wie in Dekodervariante A, allerdings wird nach Abschluss des Column Generation Verfahrens keine ganzzahlige Lösung berechnet. Als Zielfunktionswert wird der Wert der LP-Relaxation zur Bewertung des Ressourceneinsatzes auf Dienstebene weitergegeben. Dieser setzt sich zu einem Teil aus den geplanten Kosten und zum anderen aus der abgebildeten Verspätungspropagation zusammen und liefert eine Abschätzung des Zielfunktionswertes einer ganzzahligen Lösung.

Dienststückbasierte Umlaufplanung: Wenn die Dienstplanung nur als LP-Relaxation der zugrunde liegenden Formulierung gelöst wird, ist eine explizite Verknüpfung zwischen Dienst- und Umlaufplan nicht möglich. Eine implizite Verknüpfung ist jedoch bereits durch die gemeinsam als Eingabe betrachtete Dienststückmenge vorgegeben. Ebenso wie die Dienstplanung in Dekodervariante B kann auch die Umlaufplanung zunächst gemäß Dekodervariante A als LP-Relaxation gelöst werden ohne aber im Anschluss an das Column Generation Verfahren eine ganzzahlige Lösung zu berechnen. Um jedoch den Einsatz des Column Generation Verfahrens für die Umlaufplanung zu vermeiden und so den Lösungsprozess noch weiter zu beschleunigen, wird ein Umlaufplanungsproblem basierend auf dem in Abschnitt 7.6.2 beschriebenen Netzwerk formuliert. Es entspricht der Formulierung (7.57) - (7.60) und unterschätzt so die

tatsächliche Verspätungspropagation innerhalb der Umläufe. Um Kosten für den Ressourceneinsatz einzubeziehen und die benötigte Umlaufanzahl minimieren zu können, wird die Zirkulationsflusskante mit Fixkosten für den Fahrzeugeinsatz versehen. Die operativen geplanten Kosten sind durch die Dienststücke/Teilumläufe vorgegeben und müssen daher nicht mit einbezogen werden. Die Fixkostenbetrachtung wird in der Zielfunktion zu der bereits abgebildeten, vereinfachten Betrachtung von Propagation zwischen zwei Dienststücken ergänzt. Alternativ zur vereinfachten Betrachtung von Verspätungspropagation zwischen zwei Dienststücken können die Kanten im zugrunde liegenden Netzwerk auch mit Werten losgelöst potenzieller Verspätungsszenarien belegt werden. Dabei werden die Pufferzeiten zwischen der Ausführung der einzelnen Dienststücke als Kehrwert abgebildet, wobei keine Pufferzeit dem Wert 1 entspricht. Eine Minimierung der Kantenkosten in der Zielfunktion entspricht dann einer Maximierung der Pufferzeiten zwischen Teilumläufen – allerdings ohne Propagationseffekte zu berücksichtigen. Als Lösung ergibt sich ein zulässiger Umlaufplan, der die geplanten Kosten minimiert und eine Absorption von Verspätungspropagation fördert. Die tatsächliche Verspätungspropagation wird (in beiden Varianten) basierend auf den auszuführenden Aufgaben pro Umlauf nachträglich gemäß Formel 5.1 berechnet.

Die jeweils errechneten Zielfunktionswerte zum Umlauf- und Dienstplanungsproblem gehen in die Fitnessbewertung der einzelnen Individueen ein (siehe nachfolgender Abschnitt 10.4.5). Dabei erfolgt bei Einsatz der Dekodervariante B die korrekte Bewertung des Ressourceneinsatzes bezüglich der Zielkriterien erst nach Terminierung des Verfahrens (siehe Terminierungskriterien in Abschnitt 10.4.6). Erst dann wird für jedes Individuum, das zum Archiv gehört, auch ein zulässiger Dienstplan basierend auf den zuvor generierten Spalten als eine ganzzahlige Lösung durch einen MIP-Solver berechnet.

10.4.5 Fitnessberechnung

Die anhand des Dekoders pro Lösung berechneten Zielfunktionswerte bezüglich Kostenminimierung und Minimierung von Verspätungspropagation für den jeweiligen Umlauf- und Dienstplan fließen in die Berechnung der Fitness eines Individuums ein. Nach [Zitzler et al., 2002] werden im Rahmen des SPEA2 für die Fitnessberechnung die Dominanz der Individuen untereinander und die Dichte im Lösungsraum in der Umgebung von Individuen berücksichtigt. Dieser Abschnitt erläutert das Vorgehen zur Fitnessberechnung. Er folgt dabei der Beschreibung von [Zitzler et al., 2002].

Bei der Betrachtung der Zielsetzungen nach Kosteneffizienz und Robustheit dominiert eine Lösung eine andere, wenn sie sowohl weniger geplante Kosten als auch we-

niger potenzielle Verspätungspropagation beinhaltet. Für jedes Individuum i werden deshalb zunächst diejenigen Lösungen j in der aktuellen Generation g der Population P_g und im Archiv A_g gezählt, die höhere Kosten besitzen und mehr Verspätungen propagieren. Für das Individuum i erhält man so die Anzahl der von i dominierten Lösungen $S(i)$ mit

$$S(i) = |\{j | j \in P_g \cup A_g \wedge i \succ j\}|. \tag{10.10}$$

Dann wird jedes Individuum i danach beurteilt, von wie vielen anderen Lösungen es dominiert wird und wie stark diese Lösungen im Vergleich zu den anderen Lösungen sind. Es wird quasi ein Rang für jedes Individuen festgelegt und so ergibt sich ein Wert $R(i)$ mit

$$R(i) = \sum_{j \in P_g \cup A_g, j \succ i} S(j). \tag{10.11}$$

Ein besonders fittes Individuen i wird von keiner anderen Lösung dominiert und besitzt daher einen Wert $R(i) = 0$. Damit sich die Fitnesswerte der Individuen voneinander unterscheiden können - auch dann, wenn es nur wenige dominierende Lösungen gibt - wird der Fitnesswert eines Individuums i noch verfeinert, indem ein Schätzmaß $D(i)$ für die Dichte im Lösungsraum in der Umgebung einer Lösung hinzugefügt wird. Dieses Schätzmaß wird auch zur Diversifikation im Lösungsverfahren benutzt. Existieren bereits viele Lösungen mit nahezu gleichwertigen Kosten und Verspätungspropagation, sollen diese eine schlechtere Fitness zugewiesen bekommen, damit andere Bereiche des Lösungsraums ausführlicher untersucht werden. Zur Abschätzung der Dichte nutzt SPEA2 im Allgemeinen den Euklidischem Abstand zum k-ten Nachbarn. Dabei wird k in Abhängigkeit der aktuellen Populations- sowie Archivgröße gewählt als $k = \lfloor \sqrt{|P_t| + |A_t|} \rfloor$. Die Distanz eines Individuums/Lösung i zum k-nächsten Nachbarn ergibt sich als euklidische Distanz σ_i^k unter Berücksichtigung der Umlauf- und Dienstplankosten ($Costs$) und Verspätungspropagation (SD) als

$$\sigma_i^k = \sqrt{(Costs(i) - Costs(k))^2 + (SD(i) - SD(k))^2}. \tag{10.12}$$

Das konkrete Schätzmaß $D(i)$ wird dann berechnet als

$$D(i) = \frac{1}{\sigma_i^k + 2}. \tag{10.13}$$

Die tatsächliche Fitness eines Individuums i ergibt sich aus dem Schätzmaß über die Dichte $D(i)$ und darüber wie stark es dominiert wird. Sie beträgt somit:

$$F(i) = R(i) + D(i) \tag{10.14}$$

Ein geringer Wert von F(i) kennzeichnet ein fittes Individuum. Nicht-dominierte Lösungen haben einen Wert, der kleiner als 1 ist, und dominierte Lösungen besitzen einen Wert, der größer als 1 ist. Im Lösungsverfahren werden also insbesondere diejenigen Individuen gesucht, die den Wert für F(i) minimieren.

10.4.6 Terminierung

Zur Terminierung des Verfahrens können verschiedene Kriterien gewählt werden. Neben einem Zeitlimit, lässt sich das Verfahren beispielsweise über die Anzahl G (mit $G \in \mathbb{N}$) möglicher Generationen steuern. Darüber hinaus kann auch die Lösungsqualität mitberücksichtigt werden. Werden über mehrere Generationen betrachtet keine neuen dominierenden Individueen in das Archiv aufgenommen, so ist dies ein Kennzeichen dafür, dass keine Verbesserungen zur Approximation der Pareto-Front im Lösungsverlauf erzielt wurden und der Suchprozess gegebenenfalls abgebrochen werden kann.

10.5 Zusammenfassung und Implikationen

In diesem Kapitel wurde ein hybrider Lösungsansatz entwickelt, um eine Menge an Umlauf- und Dienstplänen zu erzeugen, die den möglichen Trade-Off zwischen Kosteneffizienz und Robustheit bei der Ressourceneinsatzplanung abbilden. Dazu wurden die Vorzüge des evolutionären Mehrzielalgorithmus SPEA2 mit Teilkomponenten der in den vorherigen Kapiteln dieser Arbeit entwickelten Lösungsansätze kombiniert.

Im Vergleich zur Literatur bildet der Lösungsansatz erstmals die kombinierte Suche nach Robustheit und Kosteneffizienz sowie die Planung eines mehrschichtigen Ressourceneinsatz durch einen evolutionären Lösungsprozess ab. Eine Besonderheit besteht überdies darin, dass herausgearbeitet wurde, wie der evolutionäre Prozess auch zur Ausgestaltung einzelner Ressourceneinsätze genutzt werden kann. Bei vielen Anwendungen von MOEAs in der Literatur dient der evolutionäre Prozess dahingegen allein zur Zusammenstellung vorgenerierter Umläufe und Dienste bzw. Ressourceneinsätze. Im vorgeschlagenen Lösungsansatz werden die problemunspezifischen evolutionären Operatoren ebenso zur Lösungsgenerierung eingesetzt wie problemspezifische Lösungskomponenten. Dafür wurde als geeignete Repräsentationsebene und als

Ansatzpunkt zur Planung des (mehrschichtigen) Ressourceneinsatzes die zur Abdeckung der zu bedienenden Servicefahrten benötigte Dienststückmenge identifiziert. Als Lösungsrepräsentation werden die Dienststückmengen pro Individuum aus einer Kombination von Permutationskodierung und Binärkodierung dargestellt. Zur Variation der Individuen werden ein Cycle-Crossover zur Rekombination, sowie ein Tauschoperator zur Mutation eingesetzt. Die Hybridisierung des Verfahrens findet sich in problemspezifischen Dekoderkomponenten wieder, die die jeweils betrachtete Dienststückmenge durch Lösen von Optimierungsproblemen zu Umlauf- und Dienstplänen zusammensetzen. Die beiden vorgeschlagenen Dekodervarianten entsprechen in Teilen dem teilintegrierten Verfahren aus Kapitel 7, wobei zur Beschleunigung des Verfahrens die vereinfachte Dekodervariante beim Lösen zunächst eine relaxierte Problemstellung zur Umlauf- und Dienstbildung betrachtet und erst am Ende zulässige Ressourceneinsatzpläne erstellt.

Das vorgestellte hybride Verfahren basiert auf einer im Vergleich zu den vorherigen Kapiteln einfacheren Problemstellung für den Eindepot-Fall. Die Lösungsrepräsentation lässt jedoch eine Erweiterung für beispielsweise den Mehrdepot-Fall zu, ebenso wie das Lösungsverfahren Anpassungen dafür ermöglicht.

Mit dem Lösungsansatz wurde eine im Vergleich zu den in den vorangegangenen Kapiteln vorgestellten Ansätzen zusätzliche Form der Suche nach kosteneffizienten und robusten Ressourceneinsatzplänen im Lösungsraum entwickelt. Das vorgestellte Lösungsverfahren eignet sich im Rahmen eines Entscheidungsunterstützungssystems dafür, eine Abschätzung der möglichen Lösungsausprägungen zu liefern, indem parallel vielfältige Lösungsmöglichkeiten zur Umlauf- und Dienstplangestaltung aufgezeigt werden. Zusätzlich können bei Einsatz der entsprechenden Dekodervariante explizit zulässige Zwischenergebnisse (Umlauf- und Dienstpläne) während des Lösungsprozess zur Verfügung gestellt werden.

Teil III

Evaluation der neuen Ansätze und Erkenntnisse

Kapitel 11

Vergleichende Analysen numerischer Experimente

Dieses Kapitel diskutiert den robusten und kosteneffizienten Ressourceneinsatz anhand aufeinander aufbauender Studien, die auf Experimenten mit den in den vorangegangenen Kapiteln 7 bis 10 vorgestellten Optimierungsmodellen und Lösungsverfahren sowie der in Kapitel 5 beschriebenen Simulationsmodelle basieren. Die einzelnen Studien unterscheiden sich in denen bei der Planung und (simulierten) Ausführung des Ressourceneinsatzes einbezogenen Freiheitsgrade. Darüber hinaus wird bezüglich der Detailsicht auf die sich ergebenden Planstrukturen der errechneten Ressourceneinsatzpläne unterschieden. So umfassen die Studien jeweils in unterschiedlicher Ausprägung die drei Untersuchungsebenen *Planungsmethodik*, *Robustheitsaspekt* und *Planstruktur*. Die gesamte Analyse wird geleitet durch die in Kapitel 4 formulierten Hauptforschungsfragen nach quantitativem Einfluss der Planungsmethoden auf Robustheit und Effizienz sowie nach qualitativ ableitbaren Empfehlungen für einen als robust und kosteneffizient geplanten Fahrzeug- und Personaleinsatz. Als mögliche Mechanismen für Verspätungspropagation bzw. -absorption werden dabei speziell die *Vernetzung des Ressourceneinsatzes* und die *Verdichtung von Aufgabensequenzen* in den erstellten Plänen analysiert. Abbildungen 11.1 und 11.2 fassen die einzelnen Studien sowie die den Forschungsleitfragen untergeordneten Untersuchungsfragen der jeweiligen Teilkapitel als Übersicht zusammen.

Aufbau des Kapitels

In Abschnitt 11.1 werden zunächst wesentliche Parameter der Testumgebung sowie des Testaufbaus vorgestellt, die Grundlage für die in den folgenden Abschnitten dargelegten Experimente bilden. Anschließend werden in Abschnitt 11.2 Ergebnisse der Untersuchungen über den Einfluss der Planungsmethoden diskutiert. Dazu werden die unterschiedlichen aus Anwendung der Methoden resultierenden Effekte bezüglich Robustheit und Effizienz herausgearbeitet und mit verschiedenen Ausprägungen der

	F1) Welchen Einfluss haben unterschiedliche Planungsmethoden auf Robustheit und Effizienz der erstellten Pläne?	F2) Wieviel Robustheit kann in kostenoptimalen Plänen realisiert werden?	F3) Welche Planungsempfehlungen können für den robusten und kosteneffizienten Fahrzeug- und Personaleinsatz hergeleitet werden?
(11.2) Einfluss der Planungsmethodik auf Robustheit und Effizienz	***Berücksichtigen von Verspätungspropagation in der Planung (DEP)***		
	(DEP 1.) In welcher Ausprägung ist ein Trade-Off zwischen Kosteneffizienz und Robustheit in den Ressourceneinsatzplänen möglich? **(DEP 2.)** Welches Potenzial bietet die methodengeführte Verteilung von Pufferzeiten durch (globales) Abbilden von Verspätungspropagation gegenüber einer Planung mit fixen Pufferzeiten bzw. lokaler Betrachtung von Robustheit?	**(DEP 3.)** Kann bereits in kostenoptimalen/kosteneffizienten Ressourceneinsatzplänen Robustheit realisiert werden?	**(DEP 4.)** Inwiefern erklärt die Gesamtdauer der eingeplanten Pufferzeiten in den Lösungen Unterschiede der Lösungsqualität?
	Umlaufbasierte Planung des Ressourceneinsatzes (VCS)		
	(VCS 1.) Welche Auswirkungen haben kosteneffiziente Umlaufpläne auf die robuste Planung des gesamten Ressourceneinsatzes? / Welche Möglichkeiten zur Erhöhung der Robustheit bietet die Dienstplanung? **(VCS 2.)** Welche Auswirkungen haben robuste Umlaufpläne auf eine robuste und kosteneffiziente Planung des gesamten Ressourceneinsatzes? **(VCS 3.)** Welches Potenzial bietet eine teilintegrierte gegenüber einer sequenziellen Planung?		**(VCS 4.)** Inwiefern erklärt die Häufigkeit eingeplanter Dienstwechsel (zur Vernetzung des Ressourceneinsatzes) in den Lösungen Unterschiede der Lösungsqualität?
	Gemeinsame Planung von Umläufen und Diensten im Vergleich (IVCS)		
	(IVCS 1.) In welchem Ausmaß unterscheiden sich die einzelnen Planungsmethoden bezüglich des realisierbaren Trade-Off zwischen Kosteneffizienz und Robustheit? **(IVCS 2.)** Welches Potenzial bietet eine integrierte Planung gegenüber umlaufbasierter Planung des Ressourceneinsatzes?	**(IVCS 3.)** In welchem Ausmaß kann in kostenoptimalen/kosteneffizienten Pläne Robustheit realisiert werden?	**(IVCS 4.)** Inwiefern erklärt der Umfang der jeweils für einen Umlauf- und Dienst gemeinsam eingeplanten Aufgabensequenzen (zur Aggregation von Fahrten) in den Lösungen Unterschiede der Lösungsqualität?
(11.3) Einflussfaktoren in den Planstrukturen auf Verspätungspropagation	***Gemeinsamer Verdichtungs- und Vernetzungseffekt der Ressourcenebenen (VD-VN)***		
	(VD-VN 2.) In welchem Ausmaß unterscheiden sich die Anteile der Propagationseffekte in Abhängigkeit der eingesetzten Planungsmethode?		**(VD-VN 1.)** Was wiegt stärker auf die gesamte Verspätungspropagation: die Verdichtung (Aggregation) von Fahrtensequenzen auf Dienststücke oder die Vernetzung der Ressourcen durch eingeplante Umlauf- bzw. Dienstwechsel? **(VD-VN 3.)** Inwiefern unterscheiden sich die zur Verfügung stehenden Pufferzeiten bei gemeinsam von Umläufen und Diensten auszuführenden Aufgabensequenzen innerhalb von Dienststücken? **(VD-VN 4.)** Inwiefern unterscheiden sich die zur Verfügung stehenden Pufferzeiten bei eingeplanten Umlauf- und Dienstwechseln?

Abbildung 11.1: Übersicht der aufeinander aufbauenden Studien und Aufgliederung der drei Hauptforschungsfragen - Teil 1

Robustheitsaspekt

	Planungsmethode		Planstruktur
(11.3) Einflussfaktoren in den Planstrukturen auf Verspätungsprop.			***Planstruktur-Indikatoren/Kennzahlen zu Verdichtung und Vernetzung (KENN)***
			(KENN 1.) Welche Kennzahl ermöglicht allein anhand der Planstrukturen eine Bewertung des Ressourceneinsatzes hinsichtlich möglicher Verspätungs-propagation durch Verdichtung? **(KENN 2.)** Welche Kennzahl ermöglicht allein anhand der Planstrukturen eine Bewertung des Ressourceneinsatzes hinsichtlich möglicher Verspätungs-propagation durch Vernetzung?
(11.4) Einfluss der Ausweitung der betrachteten Planungsaufgabe	***Fahrplananpassungen basierend auf geplantem Ressourceneinsatz (RET)***		
	(RET 1.) Welcher Spielraum zur Verringerung der Verspätungspropagation ist durch Retiming gegeben, ohne Robustheit durch zusätzliche Pufferzeiten zu erkaufen? **RET 2.)** In welchem Ausmaß unterscheidet sich das Potenzial von Fahrtenverschiebungen in Abhängigkeit der ursprünglich eingesetzten Planungsmethoden?	**(RET 1.)** siehe Formulierung der Frage links	**(RET 3.)** Welche Eigenschaften im Hinblick auf die Verteilung von Pufferzeiten weisen die in Dienststücken verdichteten Aufgabensequenzen aus? **(RET 4.)** Welche Eigenschaften im Hinblick auf die Verteilung von Pufferzeiten weisen die angepassten Fahrpläne auf?
	Integrierte Planung mit Zeitfenstern im Vergleich (ITW)		
	(ITW 1.) Welches Potenzial bezüglich Robuster Effizienz besteht durch robuste integrierte Planung mit Zeitfenstern gegenüber sequenzieller, teilintegrierter und integrierter Ressourceneinsatzplanung? **(ITW 2.)** Welches Potenzial besteht darüber hinaus gegenüber dem Retiming von sequenziell, teilintegriert und integriert erstellten Plänen?		**(ITW 3.)** Inwiefern sind Mengen von Servicefahrten im Fahrplan identifizierbar, deren geringe Verschiebung sowohl Robustheit als auch Kosteneffizienz des Ressourceneinsatzes erhöhen können? **(ITW 4.)** Welche Ausprägungen weist der durch (robust) integrierte Planung mit Zeitfenstern erstellte Ressourceneinsatz im Hinblick auf die Verdichtung von Aufgabensequenzen (in Dienststücken) und die Vernetzung von Dienststücken auf?
(11.5) Einfluss der Ausweitung des betrachteten Robustheitsbegriffs	***Berücksichtigen von Ressourcentausch in der Ausführung (FLEX)***		
	(FLEX 2.) In welchem Ausmaß unterscheidet sich das Potenzial durch Tauschoperationen in Abhängigkeit der eingesetzten Planungsmethoden?	**(FLEX 1.)** Ist in den erstellten Plänen Potenzial zur Reduktion von Verspätungspropagation durch Tauschoperationen während der Ausführung vorhanden?	**(FLEX 3.)** Welche Planstrukturen begünstigen den Einsatz von Tauschopera-tionen?

Abbildung 11.2: Übersicht der aufeinander aufbauenden Studien und Aufgliederung der drei Hauptforschungsfragen - Teil 2

errechneten Planstrukturen erklärt. In Abschnitt 11.3 werden die bei Einsatz der verschiedenen Planungsmethoden entstehenden Planstrukturen vertiefend untersucht. Über die Struktur der entstehenden Lösungen werden Mechanismen für Verspätungspropagation bzw. -absorption eines kombinierten Ressourceneinsatzes hergeleitet, sowie darauf basierend einfache Kennzahlen zur Identifikation robuster Pläne entwickelt. In Abschnitt 11.4 wird die Sicht auf die betrachteten Planungsebenen erweitert, indem zusätzlich zur Umlauf- und Dienstplanung Aspekte der Fahrplanung und deren Potenzial zur Steigerung von Robustheit und Kosteneffizienz untersucht werden. Im darauf folgenden Abschnitt 11.5 wird die Evaluation der erstellten Ressourceneinsatzpläne ausgeweitet, indem dispositive Maßnahmen während der Ausführung als zusätzliches Mittel zur Verringerung von Verspätungspropagation in die Analyse einbezogen werden. Damit wird neben einer veränderten Bewertung von Robustheit zusätzlich ein Bezug zwischen den Robustheitsaspekten Stabilität und Flexibilität in der Ausführung untersucht. Abschnitt 11.6 fasst das Kapitel mit aus den Analysen abgeleiteten Konsequenzen für die Planung des Ressourceneinsatzes zum Erreichen Robuster Effizienz zusammen.

Genereller Aufbau der einzelnen Untersuchungen

Mit Blick auf die wesentlichen Erkenntnisse sind die einzelnen in den (Teil-)Abschnitten beschriebenen Untersuchungen in der Regel nach folgendem Aufbau gegliedert: (1) Zunächst werden die als Grundlage für die jeweiligen Experimente agierenden *Untersuchungsfragen* mit Bezug zu den drei Hauptforschungsfragen aufgeführt. (2) Dann werden die zur Untersuchung *genutzten Artefakte* beschrieben. (3) Anschließend wird die in grafischer und/oder tabellarischer Form aggregierte *Ergebnisdarstellung* erläutert. (4) Danach folgen die Beschreibung der Ergebnisse und die *Ergebnisinterpretation* zur Beantwortung der Untersuchungsfragen. (5) Im Anschluss werden ausgewählte Aspekte zur *Planstruktur* des Ressourceneinsatzes im Hinblick auf die zuvor beschriebenen Ergebnisse diskutiert. (6) Ist die Untersuchung Teil einer Untersuchungsgruppe, werden die *Kernaussagen* zusammengefasst.

11.1 Konfiguration der Testumgebung

Die in der gesamten Arbeit aufgeführten Tests wurden unter einer einheitlichen Testumgebung auf einem Dell OptiPlex 980 mit Intel Core i7 CPU mit 2,93GHz und 16 GB RAM unter Windows 7 Professional durchgeführt. Das Optimierungssystem ist in C# implementiert und kompiliert mit .NET Framework Version 3.5 SP1. Zum Lösen von LPs und MIPs wird ILOG Cplex in der Version 12.6.0.0 verwendet.

11.1.1 Instanzenauswahl

Als Testinstanzen werden Echtwelt-Problemstellungen betrachtet. Diese umfassen Fahrpläne mit 211, 339, sowie 661 Servicefahrten. Die Instanzen sind im Folgenden gemäß der aufgeführten Reihung mit A, B und C gekennzeichnet. Für die praxisnahe Dienstbildung werden fünf verschiedene Diensttypen mit unterschiedlichen Lenkzeit- und Pausenregelungen berücksichtigt. Im Einzelnen handelt es sich dabei um Diensttypen für *Kurzdienste*, drei Arten *zusammenhängender Dienste* und *geteilte Dienste*. Kurzdienste bestehen aus einem Dienststück mit einer maximalen Arbeitszeit von fünf Stunden. Zusammenhängende Dienste bestehen aus zwei bis fünf Dienststücken in Abhängigkeit der jeweils unterschiedlichen Pausenregeln, die Dauer und Anzahl der nötigen Pausen festlegt. Geteilte Dienste umfassen zwei Dienstteile mit einer langen Pause dazwischen, und sie müssen spätestens um 19:30 Uhr beendet sein. Insbesondere die unterschiedlich variable Anzahl an Dienststücken, sowie die Trennung von Arbeits- und Lenkzeit zur Pausenerstellung, machen die Dienstgenerierung zu einem sehr komplexen Problem. Tabelle 11.1 fasst die Eigenschaften der verwendeten Diensttypen zusammen.

Tabelle 11.1: Eigenschaften der verschiedenen Diensttypen (Zeiten in hh:mm)

	Kurzdienst		**Zusam. Dienst**		**Geteilter Dienst**	
	Min	**Max**	**Min**	**Max**	**Min**	**Max**
Dienstdauer	00:30	05:00	05:30	09:45	05:30	12:00
Arbeitszeit	00:30	05:00	05:30	09:45	05:30	09:00
Lenkzeit	00:00	04:30	04:30	08:45	00:00	08:45
Ununterbrochene Lenkzeit	00:00	04:30	00:00	04:30	00:00	04:30
Lenkzeitunterbrechung	00:00	00:00	00:15	00:45	00:15	00:45
Gesamtpausenzeit	00:00	00:00	00:45	01:30	00:00	–
Dienstunterbrechung	–	–	–	–	01:30	06:00
# Dienststücke/-teile	1	1	2	5	2	2

11.1.2 Wesentliche Parameter der Simulationsumgebung

Zur Evaluation der Robustheit unter verschiedenen Rahmenbedingungen wird das in Abschnitt 5.3.1 beschriebene Propagationsmodell in Kombination mit dem entwickelten Monte-Carlo Simulationsframework eingesetzt. So können Aussagen über die erwartete Verspätungspropagation der erstellten Pläne getroffen werden. Um eine Überanpassung zu vermeiden, unterscheiden sich die zur Optimierung genutzten und während der Simulation eingesetzten Verspätungsszenarien. Sie folgen bei den in diesem Kapitel aufgeführten Hauptergebnissen jedoch der gleichen stochastischen Verteilung. Als Eingabe für die Optimierungsverfahren werden jeweils 50 Szenarien

vorgesehen. Zur Evaluation werden jeweils ca. 1.000 unabhängige Simulationsläufe mit entsprechend vielen unterschiedlichen Szenarien durchgeführt. Die erwartete propagierte bzw. sekundäre Verspätung der Servicefahrten (*E(SD)*) wird als arithmetischer Mittelwert über alle Simulationsläufe bestimmt. Die Anzahl der Simulationsläufe ist so gewählt, dass sich die Werte zur Messung der Verspätungspropagation für jede einzelne Fahrt (und damit auch für die gesamte Verspätungspropagation) stabilisieren und gegen die jeweiligen Erwartungswerte konvergieren (vgl. Abschnitt 5.3.1 zur Konvergenz bei Monte-Carlo basierten Simulationen, sowie die exemplarisch aufgeführten Konvergenzverläufe in Anhang A).

Die stochastische Verteilung, aus der primäre Verspätungen gezogen werden, wird variiert. Die in den folgenden Abschnitten aufgeführten Hauptergebnisse werden mit Einsatz einer verschobenenen Weibull-Verteilung ermittelt. Bei dieser wurden die Parameter in Abhängigkeit der Uhrzeiten – zur Abbildung von zeitlich auftretenden Effekten wie Rush Hours – basierend auf der empirischen Verteilung historischer Verspätungsdaten geschätzt. Daneben findet eine Evaluation der Ergebnisse durch die Verwendung einer Exponentialverteilung sowie einer Gleichverteilung statt, die mögliche primäre Verspätungen systematisch unter bzw. deutlich überschätzen. Dem Einsatz letztgenannter Verteilung geht jeweils ein Bernoulli-Experiment voraus, um zu bestimmen, ob eine Fahrt verspätet ist oder nicht.

11.1.3 Wesentliche Parameter der Optimierungsverfahren

Um die Zielsetzungen nach Kosteneffizienz und Robustheit in den Lösungsverfahren zu verfolgen, werden folgende Parametereinstellungen vorgenommen. Da die Minimierung von einzusetzenden Fahrzeugen und Diensten eine Hauptzielsetzung dargestellt, werden fixe Kosten von 1.000 für jeden Umlauf bzw. Diensteinsatz vorgesehen. Variable, operationale Kosten für Fahrzeuge werden auf 1,0 für jede Minute gesetzt, die ein Fahrzeug außerhalb des Depots verbringt. Die variablen Dienstkosten entsprechen 0,1 für jede Minute Arbeitszeit, ähnlich wie bei [Steinzen et al., 2010] und [Kliewer et al., 2012]. Pausenzeit wird als bezahlte Arbeitszeit verrechnet, so dass überlange Pausen vermieden werden. Diese Einstellungen ermöglichen es, gute Umlauf- und Dienstpläne zu finden mit minimaler Anzahl von Umläufen und Diensten. Zusätzlich zu den geplanten Kosten werden mit Verspätungspropagation assoziierte Kosten als zweite Zielsetzung minimiert. In Abhängigkeit der Parameterwahl für α wird jede Minute möglicher Verspätungspropagation auf Servicefahrten bestraft. Bei Anwenden der verschiedenen Lösungsansätze wird der Parameter α zwischen 0 und 10 variiert. Diese Parametereinstellungen erlauben die Skalierung zwischen Kosteneffizienz und Robustheit.

Darüber hinaus werden folgende Hauptparameter für die Abbruchkriterien der Lösungsmethoden gesetzt:

- Column Generation terminiert, wenn eine Verbesserung des Zielfunktionswertes in den letzten 15 Iterationen kleiner als 0,1% ist. Dieses Terminierungkriterium basiert auf der Beobachtung, dass in der Regel keine signifikante Verbesserung der finalen, ganzzahligen Lösung erreicht werden, wenn Column Generation erst dadurch beendet wird, dass keine (robusten) Umläufe und/oder Dienste mit negativen reduzierten Kosten mehr erstellt werden können. Um dies zu erreichen, sind insbesondere für die integrierte Planung drastisch höhere Laufzeiten selbst für sehr kleine Instanzen nötig. Aus diesem Grund wird nicht notwendigerweise eine valide, bewiesene untere Schranke für das jeweils zu lösende Gesamtproblem durch Column Generation bestimmt.

- Die Pricing-Probleme werden separat pro Kombination aus Depot-Fahrzeugtyp und Diensttyp gelöst. Die Diensterstellung für eine Kombination wird beendet, wenn 2.000 Dienste für diese Kombination erstellt werden. Die erste Phase der Umlauferstellung für eine Depot-Fahrzeugtypkombination wird beendet, wenn 1.000 Umlaufblöcke erstellt wurden. Die zweite Phase der Umlauferstellung terminiert bei 20.000 generierten Umläufen. Diese Einstellungen bieten eine gute Balance zwischen Lösungsfortschritt und benötigter Laufzeit.

- Wird das Subgradientenverfahren im Rahmen von Column Generation eingesetzt, dann ist die maximale Anzahl der Iterationen pro Column Generation Iteration auf 1.000 beschränkt.

- Die Laufzeit für jeden Test ist limitiert auf 24 Stunden.

Bis auf die jeweils zusammengehörende erste und zweite Phase bei der Umlauferstellung erlauben die Lösungsansätze eine parallele Bearbeitung in den Pricing-Schritten, da die einzelnen Pricing-Probleme unabhängig voneinander formuliert sind. Die mögliche Anzahl parallel bearbeiteter Pricing-Probleme ist allein durch die Anzahl verfügbarer Threads des Testsystems limitiert. Im konkreten Fall stehen acht Threads zur Verfügung. Für eine Diskussion über die Laufzeitvorteile von parallelisiertem Pricing sei auf [Amberg, 2014] verwiesen.

11.2 Einfluss der Planungsmethodik auf Robustheit und Effizienz

In diesem Teilkapitel werden der in Kapitel 4 formulierten Forschungsleitfrage F1 nach dem Einfluss unterschiedlicher Planungsmethoden auf die entstehenden Umlauf- und Dienstpläne, sowie Forschungsleitfrage F2 nach Robustheit in kostenoptimalen bzw. kosteneffizienten Plänen nachgegangen. Dazu werden in den einzelnen Abschnitten den Forschungsleitfragen untergeordnete Fragestellungen untersucht. Mit Hinweisen auf die entstehenden Planstrukturen als Erklärungen für die Ergebnisse wird darüber hinaus eine vertiefende Auseinandersetzung mit Forschungsleitfrage F3 nach Empfehlungen für die Planung des Ressourceneinsatzes vorbereitet.

Um den Einfluss der Planungsmethoden auf Robustheit und Effizienz des Ressourceneinsatzes zu verdeutlichen, werden die durchgeführten Untersuchungen in unterschiedliche Gruppen von Experimenten eingeteilt. Zunächst werden das Abbilden von Verspätungspropagation über mehrere Fahrten beim Erstellen von Umlaufplänen sowie über mehrere Ressourcenebenen beim Erstellen von Umlauf- und Dienstplänen solchen Lösungsansätzen gegenübergestellt, die mit festen Pufferzeiten zwischen der Ausführung von Servicefahrten planen (Abschnitt 11.2.1). Davon ausgehend werden Testergebnisse analysiert, bei denen eine robuste und/oder kosteneffiziente Umlaufplanlösung als Eingabe für die Dienstplanung dient (Abschnitt 11.2.2). Diese Ergebnisse werden schließlich Planungsvarianten integrierter Planung von Umläufen und Diensten gegenübergestellt (Abschnitt 11.2.3). Abschließend werden die Untersuchungen zusammengefasst (Abschnitt 11.2.4).

11.2.1 Berücksichtigen von Verspätungspropagation in der Planung

In diesem Abschnitt wird aufgezeigt, welchen Einfluss das Berücksichtigen von Verspätungspropagation in Optimierungsverfahren auf die Lösungsqualität der erstellten Ressourceneinsatzpläne hat. Darüber hinaus wird die durch die vorgeschlagenen Methoden mögliche Skalierbarkeit der errechneten Pläne zwischen Robustheit und Effizienz herausgearbeitet. Inbesondere soll dabei im Vergleich mit rein kostenminimierender Planung geklärt werden, ob gleichzeitig sowohl kosteneffiziente als auch robuste Pläne erstellt werden können.

Untersuchungsfragen Die durchgeführten Tests basieren auf folgenden Fragestellungen:

DEP 1.) In welcher Ausprägung ist ein Trade-Off zwischen Kosteneffizienz und Robustheit in den Ressourceneinsatzplänen möglich?

DEP 2.) Welches Potenzial bietet die methodengeführte Verteilung von Pufferzeiten durch (globales) Abbilden von Verspätungspropagation gegenüber einer Planung mit fixen Pufferzeiten bzw. lokaler Betrachtung von Robustheit?

DEP 3.) Kann bereits in kostenoptimalen/kosteneffizienten Ressourceneinsatzplänen Robustheit realisiert werden?

DEP 4.) Inwiefern erklärt die Gesamtdauer der eingeplanten Pufferzeiten in den Lösungen Unterschiede der Lösungsqualität?

Genutzte Artefakte Zur Beantwortung der Fragestellungen und als Vergleichsbasis für die nachfolgenden Untersuchungen werden exemplarische Ergebnisse der Umlaufplanung und der integrierten Umlauf- und Dienstplanung betrachtet. Diese basieren jeweils auf einer Optimierung mit Berücksichtigen von Verspätungspropagation, sowie auf einer Optimierung, bei der feste Pufferzeiten zwischen der Ausführung von Servicefahrten verlangt werden, bei einer Minimierung der geplanten Kosten als Zielsetzung.

Für die Umlaufplanung wird das Erstellen robuster Umlaufpläne basierend auf einer kostenoptimalen Flusslösung und lokaler bzw. globaler Flussdekompositon gemäß Abschnitt 7.3 eingesetzt, sowie das Erstellen steuerbar robusterer Flusslösungen mit anschließender globaler Flussdekomposition gemäß Abschnitt 7.4. Gegenübergestellt wird eine robuste Umlaufplanung mit vorgegebenen Pufferzeiten zwischen Servicefahrten für die Ausführung der Pläne. Die festen Puffervorgaben sind zum einen von den gleichen Verspätungsszenarien hergeleitet, die auch innerhalb der robusten Umlaufoptimierung für eine potenzielle Abschätzung von Verspätungspropagation genutzt werden. Zum anderen werden fixe, für jede Servicefahrt gleiche Pufferzeitvorgaben im Bereich von eins bis vier Minuten verwendet.

Ergebnisse für die gemeinsame Umlauf- und Dienstplanung werden durch das Erstellen robuster Ressourceneinsatzpläne anhand integrierter Optimierung gemäß Abschnitt 8.3 berechnet. Dabei werden die entstehenden Umlaufpläne gemäß Abschnitt 8.3.4 bzw. 7.6 mit globaler Flussdekomposition basierend auf dem entstehenden Dienstplan generiert. Als Vergleich dient zum einen eine lokale Entscheidung bei der Flussdekomposition basierend auf dem jeweils errechneten Dienstplan. Zum anderen wird eine fixierte Vorgabe von Pufferzeiten für die Ausführung von Fahrtensequenzen geprüft. Wie bei der reinen Umlaufplanung basieren die Pufferzeitvorgaben

auf den in der Optimierung mit Verspätungspropagation genutzten Verspätungsszenarien, sowie auf für jede Servicefahrt gleiche Vorgaben im Bereich von eins bis vier Minuten. In der Simulation zur Robustheitsbewertung der Pläne wird die Stabilität des jeweiligen Ressourceneinsatzes herausgestellt, indem Verspätungen solange propagiert werden bis sie absorbiert worden sind.

Ergebnisdarstellung In den Abbildungen 11.3 ist die potenzielle Robustheit der Pläne gemessen in erwarteter Verspätungspropagation (pro Servicefahrt) während der Ausführung gegenüber den Plankosten abgetragen. Die Ergebnisse beziehen sich jeweils als Vergleichswert auf eine rein kostenminimierende Planung ohne Berücksichtigung von Robustheitsaspekten bzw. mit lokaler Flussdekomposition erstellten Ressourceneinsatzplänen (weiße Punkte mit schwarzem Rand). Dabei sind pro Zeile jeweils die zusammengefassten Ergebnisse einer Instanz abgebildet, hier beispielhaft veranschaulicht an den den Testinstanzen A, B und C mit 211 bis 661 Fahrten. Die Punkte in den Abbildungen der linken Spalte repräsentieren ausgewählte Umlaufpläne, die durch reine Umlaufplanung berechnet wurden. Die Punkte in den Abbildungen der rechten Spalte geben die Bewertung des gesamten Ressourceneinsatz von Umläufen und Diensten wieder, der durch integrierte Planung errechnet wurde. Dabei wird zwischen der Vorgabe von fixen Pufferwerten (blaue Rauten) sowie dem Berücksichtigen von Verspätungspropagation in der Planung (grüne Kreise) unterschieden. Lösungen mit Pufferzeiten basierend auf primären Verspätungsszenarien sind entsprechend dem zur Auswahl der Pufferzeiten gewählten Vorgehen gekennzeichnet. Mit w werden diejenigen Lösungen bezeichnet, die Pufferzeiten bereitstellen, um die größten primären Verspätungen (worst-case) aus den Szenarien für jede Servicefahrt absorbieren zu können. Mit a werden diejenigen Lösungen markiert, die Pufferzeiten beinhalten, um durchschnittliche primäre Verspätungen (plus Standardabweichung) aus den Szenarien pro Servicefahrt absorbieren zu können. Falls einheitliche Pufferzeiten genutzt werden, sind die Lösungen in Abhängigkeit der gewählten Puffergröße mit *1* bis *4* markiert.

Ergänzend zu den Abbildungen 11.3 schlüsselt Tabelle 11.2 den konkreten Ressourceneinsatz für die Lösungen mit fixen Pufferzeiten im Vergleich zu den Basislösungen der reinen Kostenminimierung auf. Neben den Varianten der Pufferberücksichtigung wird zwischen Umlaufplanung (VS) und integrierter Planung (INT) unterschieden, sowie zwischen Umlaufanzahl ($\#v$) und Dienstanzahl ($\#n$). Beachtenswert im Vergleich zu den Angaben in der Tabelle sind die benötigten Ressourcenanzahlen für die mit Berücksichtigung von Verspätungspropagation erstellten Lösungen. Im Fall robuster Umlaufplanung mit Berücksichtigung von Verspätungspropagation enthält jede in

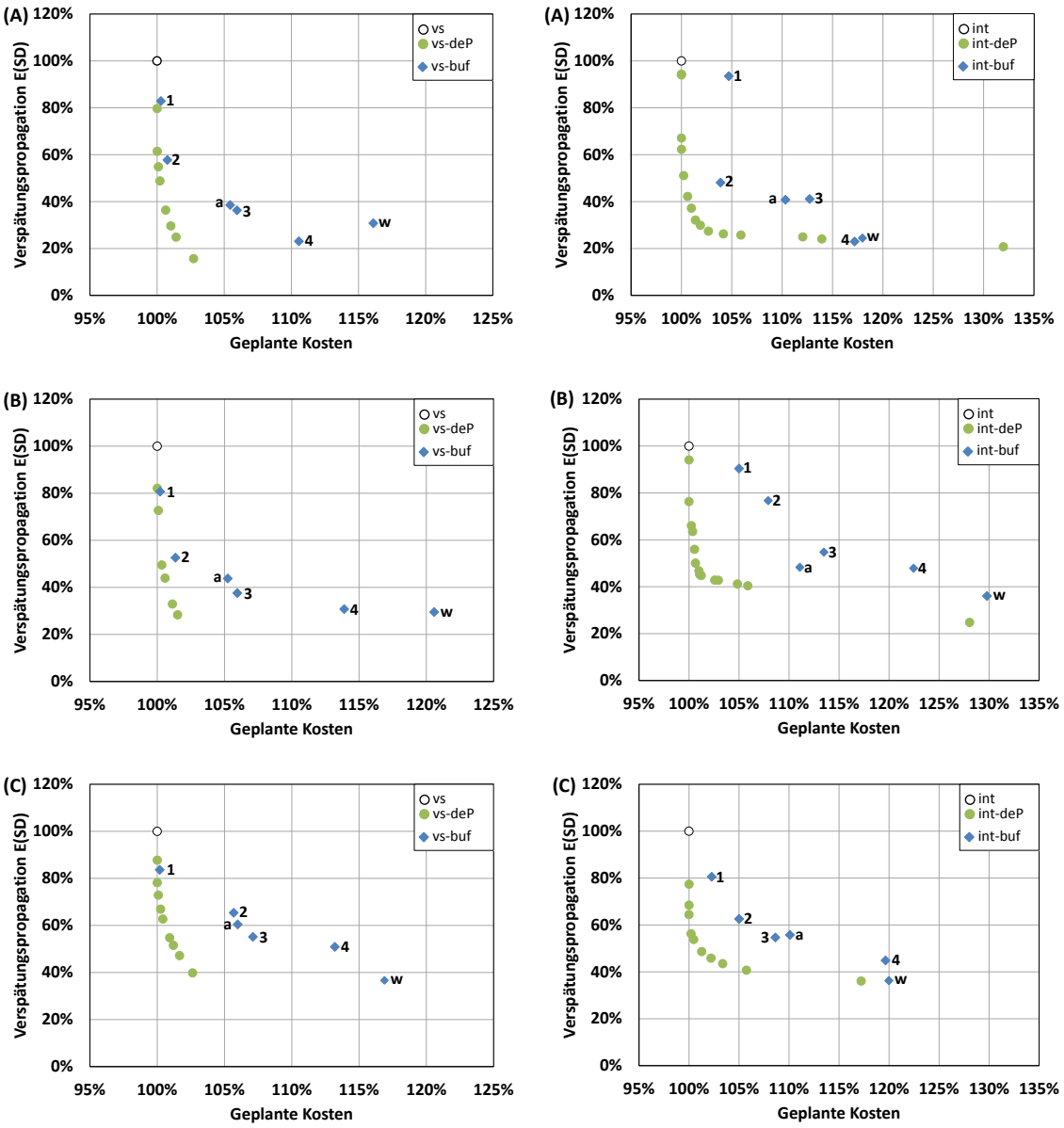

Abbildung 11.3: Gegenüberstellung geplanter Kosten und erwarteter Verspätungspropagation bei Umlaufplanung (vs) und integrierter Planung (int) mit Pufferzeiten basierend auf fixen Vorgaben (buf) und basierend auf der Berücksichtigung von propagierten Verspätungen (deP)

Tabelle 11.2: Anzahl Umläufe und Dienste bei Planung mit zuvor fixierten Pufferzeiten

	cost-opt		scenario based				global buffer time							
			avgStd		worst		1		2		3		4	
	#v	#d	#v	#d	#v	#d	#v	#d	#v	#d	#v	#d	#v	#d
Testinstanz A														
VS	17	-	18	-	20	-	17	-	17	-	18	-	19	-
INT	17	25	18	28	20	29	17	27	17	26	18	29	19	30
Testinstanz B														
VS	21	-	22	-	27	-	21	-	21	-	22	-	24	-
INT	21	29	23	32	27	37	21	32	21	33	22	34	24	37
Testinstanz C														
VS	33	-	35	-	39	-	33	-	35	-	35	-	37	-
INT	33	65	35	74	39	79	33	67	35	68	35	71	37	81

Abbildungen 11.3 dargestellte Lösung die minimale Fahrzeuganzahl. Im Fall robuster integrierter Planung mit Berücksichtigung von Verspätungspropagation enthält der Großteil der dargestellten Lösungen die minimale Umlauf- und Dienstanzahl. Genauer trifft dies für Lösungen im Bereich von 100% bis 101%, 102% der geplanten Kosten bei reiner Kostenminimierung zu. Die konkreten Umlauf- und Dienstzahlen für die einzelnen dargestellten integrierten Lösungen werden in Abschnitt 11.2.3 in Tabelle 11.4 aufgeführt.

Ergebnisinterpretation Die Abbildungen 11.3 zeigen, dass durch beide Vorgehensweisen, das Einplanen von Pufferzeiten basierend auf festen Vorgaben und das Bestimmen von Pufferzeiten basierend auf möglicher Verspätungspropagation, robuste Ressourceneinsatzpläne erzeugbar sind – sowohl allein auf Umlaufebene als auch kombiniert auf Umlauf- und Dienstebene.

(ad DEP 1.) Im Vergleich jeweils zur kostenminimierenden Planung als Basisfall gilt tendenziell, je teurer die Pläne werden, desto mehr Robustheit kann in der Ausführung erkauft werden. Insbesondere gilt mit Blick auf die ergänzenden Ergebnisse in Tabelle 11.2, je mehr Ressourcen eingesetzt werden, desto einfacher sind robustere Lösungen gestaltbar. Sowohl durch Vorgabe verschiedener Pufferzeiten aber insbesondere durch die Variation der Parameter zur Gewichtung der Berücksichtigung von Verspätungspropagation in den vorgeschlagenen Ansätzen kann also eine Skalierung des Ressourceneinsatzes zwischen Robustheit und Effizienz abgebildet werden. Dabei kann bei Berücksichtigung von Verspätungspropagation ein sehr steiler Trade-Off hergestellt werden. Ergänzend zu dieser Beobachtung zeigt Anhang B die Approximation einer

bezüglich der Zielkriterien Robustheit und Kosteneffizienz pareto-optimalen Lösungsmenge durch das in Kapitel 10 vorgeschlagene hybride Lösungsverfahren.

(ad DEP 2.) Im Vergleich der in Abbildung 11.3 dargestellten Lösungen wird deutlich, dass es für das Erstellen gleichzeitig kosteneffizient und robuster Ressourceneinsatzpläne nicht ausreicht, feste Pufferzeiten für die Verknüpfung von Servicefahrten einzuplanen. Bei der Umlaufplanung liegen zwar einzelne Lösungen mit geringen Puffervorgaben nah an den durch Berücksichtigen von Verspätungspropagation erstellten Lösungsfronten. Insgesamt ermöglicht das Berücksichtigen potenzieller Verspätungspropagation allerdings eine feinere Vorplanung zu geringeren Kosten, so dass diese Lösungen die Lösungen mit fester Puffervorgabe bezüglich der Ergebnisqualität dominieren. Das Festhalten an fixen Pufferzeiten führt zwar immer auch zu robusteren Plänen, allerdings wird dafür häufig der Einsatz zusätzlicher Ressourcen erzwungen (vgl. Tabelle 11.2). So zeigt die Tabelle 11.2, dass beispielsweise bei der reinen Umlaufplanung sowie bei der integrierten Planung eine globale Pufferzeit von jeweils mindestens zwei Minuten die Propagation von Verspätungen deutlich reduzieren kann. Im Fall der integrierten Planung werden dafür allerdings bei jeder aufgeführten Instanz zusätzliche Ressourcen benötigt. Globale Pufferzeiten von drei oder vier Minuten, genauso wie Pufferzeiten, welche die Worst-Case oder gemittelte primäre Verspätungen je Servicefahrt absorbieren können, können die potenzielle Propagation von Verspätungen drastisch reduzieren, allerdings klar zu Lasten des Ressourceneinsatzes. So werden beispielsweise für das Worst-Case Szenario bei Instanz B sechs zusätzliche Umläufe eingeplant, um einen robusten Umlaufplan zu erstellen. Für die integrierte Planung werden in diesem Fall vierzehn zusätzliche Ressourcen eingeplant, um einen robusten Umlauf- und Dienstplan zu erhalten. Im Vergleich dazu ist durch die Planung mit Berücksichtigen von Verspätungspropagation die gleiche Robustheit bei deutlich weniger Ressourceneinsatz zu erreichen. Im Beispiel der Instanz B wird jeder durch Planung mit festen Pufferzeiten erzielte Umlaufplan von einer Lösung dominiert, die durch Planung mit Berücksichtigen von Verspätungspropagation erstellt wurde und die minimale Umlaufanzahl enthält. Für die integrierte Planung trifft dies ebenso für alle Umlauf- und Dienstpläne zu, mit Ausnahme der Lösung für das Worst-Case Szenario. Ähnliche Beobachtungen gelten für die anderen Instanzen. Denn wie zuvor erwähnt, werden in den abgebildeten Lösungen mit Berücksichtigung von Verspätungspropagation im Bereich von 100% bis ca. 102% der Plankosten die gleichen Ressourcenanzahlen eingesetzt wie jeweils bei der rein kostenminimierenden Lösung vorgesehen.

(ad DEP 3.) Darüber hinaus wird deutlich, dass auch im Vergleich zu rein kostenminimierend geplantem Ressourceneinsatz robustere Lösungen bei nahezu iden-

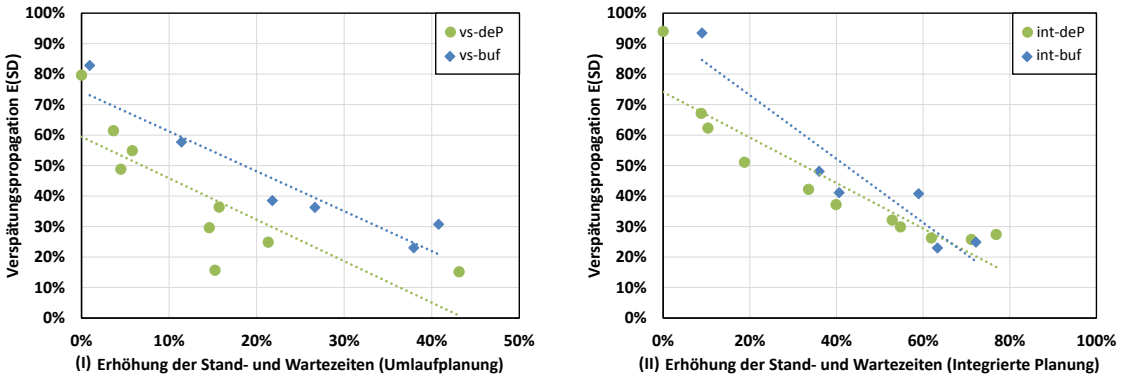

Abbildung 11.4: Eingeplante Pufferzeiten und potenzielle Verspätungspropagation im Vergleich zu rein kostenminimierender Planung am Beispiel der Testinstanz A

tischen Plankosten gefunden werden können. So sind bei der Umlaufplanung bzw. bei der integrierten Umlauf- und Dienstplanung mit geringfügigen Mehrkosten ohne zusätzliche Fahrzeuge, deutlich robustere Pläne im Vergleich zu kostenminimierender Planung möglich. Das in den Abbildungen dargestellte Verbesserungspotenzial liegt im Bereich zwischen 45-75% möglicher Propagationsreduktion. Mehr noch, in jeder Grafik entspricht die erste Lösung – der erste (grüne) Punkt – unter dem Basisfall der kostenminimierenden Planung jeweils einem Umlaufplan bzw. einem Umlauf- und Dienstplan mit exakt identischer Aufgaben bzw. Kostenstruktur wie beim Basisfall. Lediglich eine andere Zuordnung von Aufgabensequenzen bzw. Dienststücken und Diensten zu Umläufen führt zu einem deutlichen Verbesserungspotenzial im Rahmen von ca. 5 bis 20%. Womit aus methodischer Sicht eine Überlegenheit globaler Flussdekomposition zur Verteilung von Pufferzeiten im Vergleich zu lokaler Flussdekomposition auf Umlauf- bzw. Dienststückebene verdeutlicht wird.

Planstrukturen (ad DEP 4.) Ein vertiefender Blick auf die Strukturen der errechneten Ressourceneinsatzpläne zeigt, dass in der Tendenz zusätzliche Robustheit durch zusätzliche Pufferzeiten erkauft wird. Abbildung 11.4 stellt exemplarisch für Testinstanz A die Erhöhung der Pufferzeit, die gleichzeitig für sowohl Fahrer als Fahrzeuge eingeplant wird, und die potenzielle Verspätungspropagation im Vergleich zur kostenmininierenden Lösung gegenüber. Dabei betrachtet Bild I robuste Planung durch Berücksichtigen von Verspätungspropagation und durch fixe Pufferzeiten in der Umlaufplanung, und Bild II in der integrierten Umlauf- und Dienstplanung.

In beiden Bildern der Abbildung 11.4 wird deutlich, dass einerseits die Pufferzeit wesentlich entscheidend für zusätzliche Robustheit ist. Andererseits bedarf es aber

auch einer geschickten Verteilung dieser Pufferzeit, um robustere Lösungen zu generieren. Nahezu durchweg dominieren Ressourceneinsatzpläne, die mit Berücksichtigen von Verspätungspropagation erstellt wurden, die mit fester Puffervorgabe erstellten Pläne. Deutlich erkennbar ist dies an den Lösungen der Umlaufplanung, in denen speziell in der Planung mit vogegebenem Puffer zwar viele Warte- und Standzeiten vorgesehen werden, diese aber ungeschickter verteilt werden. Bei der integrierten Planung ist dieser Effekt auch vorhanden, wird aber nicht ganz so deutlich herausgestellt. Die Robustheit durch Planung mit festen Pufferzeiten wird hier dadurch begünstigt, dass (deutlich) mehr Ressourcen eingesetzt werden und somit die auszuführenden Aufgabensequenzen auf mehr Ressourcen verteilt werden können. So haben bis auf zwei Lösungen, alle anderen mindestens ein Fahrzeug mehr im Einsatz, sowie alle Lösungen mindestens einen Dienst mehr zur Verfügung (vgl. Tabelle 11.2).

Kernaussagen Robustheit im Sinne von Stabilität in der Ausführung kann bei der Planung des Ressourceneinsatzs von Umläufen und Diensten leicht durch zusätzliche Ressourcen und zusätzliche Pufferzeit erkauft werden. Als ein wesentlicher Indikator für die Robustheit kann dabei die eingeplante Dauer von (gemeinsamen) Stand- bzw. Wartezeiten in Umläufen und Diensten betrachtet werden.

Zugleich wurde gezeigt, dass mit den vorgeschlagenen Lösungsansätzen eine deutliche Robustheitserhöhung im Vergleich zu kostenminimierender Planung möglich ist, ohne bzw. mit nur geringfügiger Erhöhung der geplanten Kosten. Der gemeinsame Ressourceneinsatz von Umläufen und Diensten kann folglich sowohl gleichzeitig kosteneffizient als auch robust gestaltet werden. Dabei führt das Abbilden von Verspätungspropagation in den Lösungsverfahren zu einer Verbesserung der Lösungen. Unter Einbeziehung der durchzuführenden Fahrtensequenzen erfolgt eine passendere Verteilung der Pufferzeiten bei der Planerstellung, so dass die gleiche Robustheit zu wesentlich geringeren Kosten erreichbar ist im Vergleich zu einer festen Pufferzeitvorgabe bei der Planung.

11.2.2 Auswirkungen umlaufbasierter Planung des Ressourceneinsatzes

Im vorangegangenen Abschnitt 11.2.1 wurde gezeigt, dass mit den vorgeschlagenen Optimierungsmethoden allein für die Umlaufplanungsebene sowie für eine integrierte Planung von Umläufen und Diensten sowohl robuste als auch kosteneffiziente Lösungen erreicht werden können. Als Frage bleibt offen, wie entscheidend die Planungsmethodik für die Robustheit (im Sinne von Stabilität) und die Kosteneffizienz der

erstellten Pläne ist, wenn sowohl Umlauf- als auch Dienstebene betrachtet werden. Deshalb werden in diesem Abschnitt zusätzlich die Auswirkungen einer der Dienstplanung vorgelagerten Umlaufplanung auf die Planung des gesamten Ressourceneinsatz untersucht, bevor diese Ergebnisse im anschließenden Abschnitt 11.2.3 mit den Ergebnissen integrierter Planung verglichen werden. Für die Untersuchung werden zwei umlaufbasierte Planungsmethoden mit unterschiedlichen Freiheitsgraden miteinander verglichen. Zum einen eine sequenzielle Umlauf- und Dienstplanung. Und zum anderen eine teilintegrierte Umlauf- und Dienstplanung.

Dabei sei darauf hingewiesen, dass in dieser Arbeit eine sequenzielle Folge der Planungsstufen in der Reihung zunächst Dienstplanung dann Umlaufplanung nicht näher aufgeführt wird. Bei einer solchen Planweise können Dienste bequem kostenminimal eingeplant werden, häufig zu Lasten der benötigten Fahrzeuganzahl. Darüber hinaus unterliegen Umläufe allein – im Gegensatz zu Fahrerdiensten – keinen Pausenzeitregelungen oder einschränkenden Dienstdauern. Sie umfassen deshalb einen größeren Zeitraum und beinhalten dadurch in der Regel mehr Servicefahrten auf denen sich Verspätungen kumulieren können.

Untersuchungsfragen Grundlage für die durchgeführten Tests bilden folgende Fragestellungen:

VCS 1.) Welche Auswirkungen haben kosteneffiziente Umlaufpläne auf die robuste Planung des gesamten Ressourceneinsatzes? / Welche Möglichkeiten zur Erhöhung der Robustheit bietet die Dienstplanung?

VCS 2.) Welche Auswirkungen haben robuste Umlaufpläne auf eine robuste und kosteneffiziente Planung des gesamten Ressourceneinsatzes?

VCS 3.) Welches Potenzial bietet eine teilintegrierte gegenüber einer sequenziellen Planung des Ressourceneinsatzes?

VCS 4.) Inwiefern erklärt die Häufigkeit eingeplanter Dienstwechsel (zur Vernetzung des Ressourceneinsatzes) in den Lösungen Unterschiede der Lösungsqualität?

Genutzte Artefakte Um die Fragestellungen zu untersuchen, wird die teilintegrierte Planung gemäß dem in Abschnitt 6.3.3 zusammengefassten Vorgehensschema mit einer sequenzieller Planung gemäß Abschnitt 6.3.2 verglichen. Die sequenzielle Planung wird derart ausgeführt, dass nachträglich auf einem gemäß Abschnitt 7.3 aus einer (robusten) Flusslösung dekomponiertem Umlaufplan ein Dienstplan erstellt wird. Dabei wird die Dienstplanung als Set-Partitioning Problem basierend auf den konkreten Aufgabenzuordnungen einzelner Umläufe formuliert und gelöst. Im Unterschied

dazu werden teilintegrierte Lösungen basierend auf einer (robusten) Flusslösung berechnet. Diese Flusslösung ist identisch mit derjenigen, die bei sequenzieller Planung und globaler Flussdekomposition für die Umlaufplanung verwendet wird. Die globale Flussdekomposition beim teilintegrierten Verfahren erfolgt jedoch erst, nachdem ein Dienstplan erstellt wurde. In beiden Fällen basiert der Dienstplan auf einem Umlaufplan mit exakt gleicher Kostenstruktur und durchzuführenden Aufgaben in der Dienstebene. Aufbauend auf Abschnitt 11.2.1 bilden dabei speziell die kosteneffizienten Umlaufpläne bzw. die korrespondierenden kostengleichen Umlaufplanbündel mit bis zu ca. 101,5% Kosten im Vergleich zu reiner kostennminimierenden Planung die Aufgabenvorgabe für die Dienstplanung. In diesen Umlaufplänen sind keine zusätzlichen Fahrzeuge eingeplant, so dass der Fokus der Untersuchung auf sowohl Robustheit als auch Kosteneffizienz des gesamten Ressourceneinsatzes liegen kann. Wie im Abschnitt zuvor gezeigt, kann Robustheit bei Vernachlässigen der Kosteneffizienz leicht durch zusätzliche Ressourcen erkauft werden.

Ergebnisdarstellung Tabelle 11.3 fasst Ergebnisse einer umlaufbasierten Gesamtplanung des Ressourceneinsatzes zusammen. Sie stellt exemplarisch die errechneten Ressourceneinsatzpläne für verschiedene Testinstanzen aufgeschlüsselt in Umlaufebene (*VS*), Dienstebene (*CS*) sowie Gesamtressourceneinsatz (*VCS*) mit geplanten Kosten und möglicher Verspätungspropagation gegenüber. Dabei werden sequenzielle Planungsmethoden und teilintegrierte Planungsmethoden hinsichtlich ihrer Lösungsqualität und Laufzeiten miteinander verglichen.

Aufgeführt werden zum einen die verschiedenen Lösungsansätze (*approach*). Für die Teilabschnitte zu sequenzieller Planung (*seq*) bzw. teilintegrierter Planung (*pInt*) erfolgt der Aufbau gleich. In der ersten Zeile wird jeweils eine rein kostenminimierende Planung aufgeführt, basierend auf lokaler Flussdekomposition auf der Umlaufebene (*seq, pint*). Im direkten Vergleich folgt dann eine Lösung mittels globaler Flussdekomposition zur identischen Flusslösung auf Umlaufebene (*seqGFD, pIntGFD*). Nachfolgend werden weitere Lösungen aufgeführt, die mit aufsteigendender Robustheit der zugrunde liegenden Umlaufpläne bzw. Bündel erstellt wurden (*seqR, pIntR*).[1] Zusätzlich werden Lösungen aufgeführt, die durch teilintegrierte Planung erstellt wurden, bei

[1] Sofern nicht explizit anders ausgewiesen umfassen die in den Tabellen dieses Kapitels mit „R“ gekennzeichneten Lösungsansätze immer auch die zum Lösungsansatz korrespondierende Version globaler Flussdekomposition „GFD“. In den verschiedenen Abbildungen dieses Kapitels werden die durch diese Ansätze berechneten Lösungen generell gemeinsam mit „deP“ gekennzeichnet, wodurch das explizite Berücksichtigen von (globaler) Verspätungspropagation verdeutlicht wird. Die Unterscheidung zwischen durch „GFD“ und mit umfassenderer Berücksichtigung von Verspätungspropagation im Lösungsprozess durch „R“-Ansätze erstellten Lösungen wird dann in den zugehörigen Erläuterungen getroffen.

der für die Dienstplanung die gleichen Kosten der jeweils korrespondierenden sequenziellen Dienstplanlösung erlaubt bzw. erzwungen wurden. Der Einsatz teilintegrierter Methoden ist in solchen Fällen mit „*“ gekennzeichnet. Für die Umlaufebene werden die benötigte Laufzeit (*runtime*), die Anzahl benötigter Fahrzeuge ($\#v$), die Plankosten (*costs*) in Bezug auf den als Vergleichsbasis dienenden, rein kosteniminierend erstellten Umlaufplan, sowie die Verspätungspropagation innerhalb der Umläufe als Erwartungswert in Sekunden pro Servicefahrt ($E(SD_V)$) angegeben. Analog werden für die Dienstebene die Anzahl benötigter Dienste ($\#d$), die Plankosten in Bezug auf den als Vergleichsbasis dienenden, sequenziell erstellten Dienstplan, sowie die Verspätungspropagation innerhalb der Dienste als Erwartungswert pro Servicefahrt ($E(SD_C)$) aufgelistet. Zusätzlich werden ebenso für den gesamten Ressourceneinsatz die benötigte Rechenzeit (*runtime*), die Plankosten im Vergleich zum sequenziell errechneten Ressourceneinsatz bei reiner Kostenminimierung (*costs*) und die potenzielle Propagation von Verspätungen unter Einbeziehen aller Ressourcenverknüpfungen als Erwartungswert pro Servicefahrt ($E(SD)$) aufgeführt.

Ergebnisinterpretation Die Tabelle 11.3 verdeutlicht im Allgemeinen, dass robustere Umlaufpläne als Vorgabe in den beiden Planungsvarianten sequenzielle und teilintegrierte Planung zu einem robusteren Gesamtressourceneinsatz führen. Es sind variantenübergreifend zwei gemeinsame Beobachtungen festzustellen.

(ad VCS 1.) Erstens zeigt die Tabelle in Ergänzung zu den in Abschnitt 11.2.1 aufgeführten Ergebnissen zur Beantwortung der zugehörigen Frage (DEP 3.), dass wesentlich robustere Pläne zu identischen Kosten im Vergleich mit einer rein kostenminimierenden Planung berechnet werden können, sofern beim Lösen Robustheitsaspekte (speziell Verspätungspropagation) global für den Ressourceneinsatz betrachtet werden – sowohl allein für die Umlaufplanung als auch bei der gemeinsamen Umlauf- und Dienstplanung. Folglich ist beim Lösen der Planungsprobleme eine gewisse Robustheit für das Zusammenstellen der Aufgabensequenzen eigentlich schon vorhanden, die bei rein kostenminimierenden Lösungsansätzen allerdings ignoriert wird. Zu erkennen ist dies jeweils beim Vergleich der ersten mit der zweiten Zeile eines Tabellenabschnittes. Der Vergleich zwischen lokaler und globaler Flussdekomposition mit Robustheitsbetrachtung gilt bei teilintegrierter Planung für den gesamten Ressourceneinsatz, und bei sequenzieller Planung eingeschränkt für die Spalten zur Umlaufplanung. In den exemplarisch aufgeführten Testinstanzen liegt das Verbesserungspotenzial auf Umlaufebene im Bereich von 10-20% und beim gesamten Ressourceneinsatz im Bereich von 13-24% im Vergleich zur potenziellen Verspätungspropagation der rein kostenmi-

Tabelle 11.3: Vergleich von Kosteneffizienz und Robustheit bei sequenzieller und teilintegrierter Planung auf Umlauf- (VS) und Dienstebene (CS), sowie beim gemeinsamen Ressourceneinsatz (VCS)

approach	**VS**				**CS**			**VCS**		
	runtime	#v	costs	E(SD_V)	#d	costs	E(SD_C)	runtime	costs	E(SD)
					Testinstanz A					
seq	00:00:01	17	100,0	53,3(- 0%)	31	100,0	42,1(- 0%)	00:00:21	100,0	61,4(- 0%)
seqGFD	00:00:04	17	100,0	42,4(-20%)	31	99,9	38,6(- 8%)	00:00:47	99,9	44,9(-27%)
seqR	00:01:02	17	100,1	29,2(-45%)	35	112,5	24,6(-42%)	00:01:24	106,9	32,1(-48%)
	00:01:00	17	100,2	26,0(-51%)	36	115,8	22,7(-46%)	00:01:24	108,8	29,0(-53%)
	00:01:11	17	100,6	19,4(-64%)	39	124,9	16,2(-62%)	00:01:43	114,0	22,3(-64%)
	00:01:09	17	101,1	15,8(-70%)	39	125,0	11,7(-72%)	00:01:35	114,2	19,4(-68%)
	00:01:23	17	101,4	13,3(-75%)	39	125,1	11,2(-73%)	00:01:41	114,4	15,4(-75%)
pInt	00:00:01	17	100,0	47,5(- 0%)	30	96,9	42,5(- 0%)	00:00:30	98,3	50,9(- 0%)
pIntGFD	00:00:01	17	100,0	43,3(- 9%)	30	96,9	42,5(- 0%)	00:00:31	98,3	43,9(-13%)
pIntR	00:01:02	17	100,1	36,2(-24%)	30	96,9	35,9(-16%)	00:01:34	98,4	37,4(-26%)
	00:01:00	17	100,2	33,9(-29%)	29	94,1	33,5(-21%)	00:01:28	96,8	34,4(-32%)
	00:01:11	17	100,6	24,7(-48%)	32	103,1	23,8(-44%)	00:01:42	102,0	25,3(-50%)
	00:01:09	17	101,1	26,2(-45%)	32	103,2	24,9(-41%)	00:01:57	102,2	28,0(-44%)
	00:01:23	17	101,4	19,2(-60%)	32	103,3	18,4(-57%)	00:02:03	102,4	19,7(-61%)
pInt*	00:00:01	17	100,0	47,4(- 0%)	31	100,1	42,3(- 0%)	00:00:30	100,1	50,1(- 0%)
pIntGFD*	00:00:01	17	100,0	43,1(- 9%)	31	100,1	42,3(- 0%)	00:00:34	100,1	43,6(-13%)
pIntR*	00:01:02	17	100,1	29,4(-38%)	34	110,4	27,2(-36%)	00:01:29	105,8	29,6(-41%)
	00:01:00	17	100,2	26,3(-45%)	35	112,7	24,5(-42%)	00:01:33	107,0	26,6(-47%)
	00:01:11	17	100,6	19,7(-58%)	37	118,7	17,4(-59%)	00:01:43	110,6	20,1(-60%)
	00:01:09	17	101,1	15,9(-66%)	37	119,2	14,3(-66%)	00:01:56	111,0	16,5(-67%)
	00:01:23	17	101,4	13,3(-72%)	37	119,4	12,4(-71%)	00:02:10	111,3	13,9(-72%)
					Testinstanz B					
seq	00:00:03	21	100,0	120,2(- 0%)	37	100,0	88,3(- 0%)	00:00:22	100,0	132,3(- 0%)
seqGFD	00:00:05	21	100,0	108,7(-10%)	37	100,0	85,8(- 3%)	00:00:27	100,0	121,0(- 9%)
seqR	00:01:10	21	100,1	87,4(-27%)	35	94,8	81,1(- 8%)	00:01:42	97,2	91,5(-31%)
	00:04:01	21	100,3	59,5(-50%)	44	117,7	53,3(-40%)	00:04:24	109,7	63,8(-52%)
	00:03:55	21	100,6	52,8(-56%)	44	117,8	44,6(-50%)	00:04:17	109,9	58,3(-56%)
	00:06:12	21	101,1	39,6(-67%)	43	115,4	29,9(-66%)	00:06:47	108,8	45,8(-65%)
	00:10:31	21	101,5	34,0(-72%)	47	125,8	24,6(-72%)	00:11:10	114,7	39,8(-70%)
pInt	00:00:03	21	100,0	116,1(- 0%)	32	87,0	97,6(- 0%)	00:03:50	93,0	120,2(- 0%)
pIntGFD	00:00:05	21	100,0	99,4(-14%)	32	87,0	97,6(- 0%)	00:03:53	93,0	99,9(-17%)
pIntR	00:01:10	21	100,1	86,4(-26%)	32	87,1	84,7(-13%)	00:01:49	93,1	88,0(-27%)
	00:04:01	21	100,3	69,2(-40%)	39	105,0	63,6(-35%)	00:04:53	102,8	74,1(-38%)
	00:03:55	21	100,6	52,7(-55%)	40	107,4	47,3(-52%)	00:04:56	104,3	55,6(-54%)
	00:06:12	21	101,1	42,6(-63%)	40	107,6	36,5(-63%)	00:08:44	104,6	47,3(-61%)
	00:10:31	21	101,5	40,5(-65%)	41	110,2	32,0(-67%)	00:12:03	106,2	46,1(-62%)
					Testinstanz C					
seq	00:00:14	33	100,0	31,6(- 0%)	70	100,0	23,2(- 0%)	00:02:15	100,0	34,9(- 0%)
seqGFD	00:00:24	33	100,0	27,2(-14%)	70	100,1	23,0(- 1%)	00:05:59	100,0	31,4(-10%)
seqR	00:06:51	33	100,1	22,6(-28%)	71	101,4	18,8(-19%)	00:10:06	100,8	26,6(-24%)
	00:10:08	33	100,2	20,1(-36%)	70	100,1	18,3(-21%)	00:15:21	100,2	23,8(-32%)
	00:10:37	33	100,4	18,1(-43%)	73	104,2	16,5(-29%)	00:14:38	105,5	22,1(-37%)
	00:14:50	33	101,2	16,0(-49%)	78	111,0	14,0(-40%)	00:18:05	106,5	18,9(-46%)
	00:14:33	33	101,6	14,9(-53%)	77	109,7	14,1(-39%)	00:23:45	106,0	19,2(-45%)
pInt	00:00:14	33	100,0	33,7(- 0%)	69	98,7	23,9(- 0%)	00:06:33	99,3	40,0(- 0%)
pIntGFD	00:00:24	33	100,0	27,7(-18%)	69	98,7	23,9(- 0%)	00:06:36	99,3	30,2(-24%)
pIntR	00:06:51	33	100,1	23,0(-32%)	71	101,4	18,8(-21%)	00:12:28	100,8	26,2(-35%)
	00:10:08	33	100,2	21,1(-37%)	69	98,8	18,5(-23%)	00:30:20	99,4	22,5(-44%)
	00:10:37	33	100,4	18,5(-45%)	71	101,4	17,9(-25%)	00:18:01	101,0	21,8(-46%)
	00:14:50	33	101,2	16,3(-52%)	75	106,9	14,7(-38%)	00:20:37	104,3	17,7(-56%)
	00:14:33	33	101,6	15,0(-55%)	75	106,5	14,2(-41%)	00:27:43	104,9	17,1(-57%)

nierenden Ausgangslösung, bei der lediglich lokale Entscheidungen über Robustheit getroffen werden.

(ad VCS 2.) Zweitens wird veranschaulicht, dass mit geringfügig zunehmenden Kosten und gleichzeitig zunehmender Robustheit auf Umlaufebene auch die zugehörigen Dienstpläne mit größerer Robustheit ausgestattet werden. Je nach Planungsvariante können die Dienstpläne aber in der Gesamttendenz dabei auch deutlich teurer gestaltet sein als die rein kostenminimierende Vergleichslösung. Dabei ist zu beachten, dass teilweise teurere Umlaufpläne zu kostengünstigeren Dienstplänen und somit zu einem günstigerem (und robusterem) Gesamtressourceneinsatz führen, da diese Umlaufpläne zufällig besser an die bei der Dienstplanung zu berücksichtigten Dienstregeln angepasst sind.

(ad VCS 3.) Im Vergleich des ermittelten Ressourceneinsatzes zeigen sich Unterschiede zwischen den beiden Planungsvarianten sequenzielle und teilintegrierte Planung. Bei teilintegrierter Planung kann durch die Auswahl des am besten zum Dienstplan passenden Umlaufplans der Ressourceneinsatz aufgrund weniger benötigter Dienste deutlich kostengünstiger gestaltet werden. Im Hinblick auf die Robustheit ist dabei festzustellen, dass somit zu geringeren Kosten robustere Gesamtlösungen als durch sequenzielle Planung erstellt werden können. Dies ist auch in einigen Fällen beim direkten Vergleich des Ressourceneinsatzes zu beobachten, der durch Planungsvarianten basierend auf der aufgabenidentischen Umlauflösung als Vorgabe erstellt wurde (siehe insbesondere Testinstanz C). Allerdings zeigt sich bei diesem Vergleich auch, dass bei gleicher Vorgabe der Umlaufplanaufgaben, die sequenzielle Planung durch die Inkaufnahme eines größeren Ressourceneinsatz zu einer robusteren Gesamtlösung führen kann (siehe insbesondere Testinstanz A). Dazu sei angemerkt, dass mit der teilintegrierten Planung ein zu diesen sequenziellen Lösungen mindestens gleichwertig robuster Ressourceneinsatz bestimmt werden kann, wenn beispielsweise durch Anwenden des Dienstplanungsproblems (7.47) - (7.51) mit Elastic-Cost-Constraint gemäß Abschnitt 7.5.1 eine einzuhaltende Kostenvorgabe bei der Dienstplanung erzwungen wird. Dies zeigen die exemplarisch aufgeführten Ergebnisse (siehe mit „*“ markierte Verwendung teilintegrierter Methoden). Der hierbei zugestandene und erzwungene Einsatz eines größeren Kostenrahmens für die Dienstplanung führt aufgrund der Freiheitsgrade teilintegrierter Planung zu robusteren Gesamtlösungen bei größtenteils immer noch geringerem Ressourceneinsatz als es die sequenzielle Planung ermöglicht.

Wird allein die Umlaufebene im Vergleich der Planungsmethoden mit Berücksichtigung von Verspätungspropagation betrachtet, wird deutlich, dass die mögliche Verspätungspropagation in den sich ergebenden Umlaufplänen bei sequenzieller Planung

methodengemäß in der Regel kleiner gleich der Propagation in den aufgabengleichen Umlaufplänen der teilintegrierten Planung ist. Dies ist auf die globale Flussdekomposition zurückzuführen, die bei sequenzieller Planung allein eine Minimierung der Verspätungsprogation in den Umlaufplänen zum Ziel hat ohne Dienste einzubeziehen – mit der Folge, dass durch mögliche Unangepasstheit der enstehenden Aufgabensequenzen für die Dienstbildung auch mehr Dienste für die Bedienung der vorgegebenen Umlaufpläne nötig werden. Ausnahmen bei der Verspätungspropagation auf Umlaufebene bilden hier eine durch seqGFD bzw. seqR berechnete Umlauflösung für Testinstanz B. Bei diesen Lösungen ist das Column Generation Verfahren vergleichsweise früh terminiert, so dass in beiden Fällen das aufgabengleiche durch pIntGFD bzw. pIntR auf der Dienstlösung berechnete Umlaufplan-Gegenstück eine geringere Propagation aufweist.

(ad VCS 1.) Wird die Verspätungspropagation auf den einzelnen Ressourcenebenen verglichen, zeigen sich unabhängig der eingesetzten Planungsmethode unterschiedliche Propagationseffekte in Umlaufplänen und Dienstplänen sowie im gesamten Ressourceneinsatz. Zum einen ist zu erkennen, dass Umläufe isoliert betrachtet mehr Verspätungspropagation verursachen als Dienste. Dies ist auf die bei Umläufen in der Regel längeren Verkettungen von Fahrtensequenzen während eines Betriebstages zurückzuführen, da beim Fahrzeugeinsatz im Gegensatz zu Fahrerdiensten keine Pausen oder Arbeitszeitenvorgaben eingehalten werden müssen. Aus diesem Grund gestaltet es sich schwierig, durch robuste Dienstplanung nachträglich einen insgesamt stabilen/robusten Ressourceneinsatz basierend auf einer rein kosteneffizienten Umlauflösung auszuarbeiten (wie es z.B. auch trotz der zusätzlichen Freiheitsgrade teilintegrierter Planung bei pInt* bzw. pIntGFD* im Vergleich zu pInt bzw. pIntGFD verdeutlicht wird). Als Konsequenz ergibt die Messung der Verspätungspropagation auf Umlaufebene eine bessere Abschätzung der potenziell erreichbaren Robustheit im Sinne von Stabilität des gesamten Ressourceneinsatzes.

(ad VCS 3.) Zum anderen wird beim Vergleich der Verspätungspropagation in Bezug auf die verschiedenen Ressourcenebenen deutlich, dass der Ressourceneinsatz von Umläufen und Diensten kombiniert zu größerer Verspätungspropagation führt als beispielsweise allein durch die Propagation auf Umlaufebene abgebildet wird. Durch die Vernetzung der beiden Ressourcenebenen über den geplanten Diensteinsatz entsteht also ein zusätzlicher Propagationseffekt. Dies führt dazu, dass im Vergleich zwischen sequenziell und teilintegrierter Planung ein robuster bewerteter Umlaufplan als Grundlage für die Planung nicht zwangsläufig zu einem robusteren Gesamtressourceneinsatz führen muss (siehe auch vorherige Erläuterungen zur Beantwortung

der Frage VCS 3.), da die teilintegrierte Planung verfahrensgemäß besser mit der Vernetzung der beiden Ressourcenebenen umgehen kann.

Beim Blick auf die zur Lösung benötigten Gesamtlaufzeiten offenbaren sich erst bei den größeren Instanzen Laufzeitunterschiede zwischen sequenzieller und teilintegrierter Planung. Die teilweise höhere Laufzeit teilintegrierter Planung resultiert aus den für die Dienstplanung zur Verfügung stehenden größeren Freiheitsgrade zur Dienstbildung. Darüber hinaus wird deutlich, dass das explizite Einbeziehen von Verspätungspropagation zu höheren Laufzeiten führt. Treibende Kraft ist hierbei insbesondere die robuste Umlaufplanung (siehe Ansätze mit „R“ im Vergleich mit den anderen).

Planstrukturen (ad VCS 4.) Die erstellten Pläne weisen aufgrund der eingesetzten Lösungsmethodik bei sequenzieller und teilintegrierter Planung exakt die gleichen gemeinsam für Fahrer und Fahrzeuge eingeplanten Stand- und Wartezeiten auf. Die Strukturen der Ressourceneinsatzpläne offenbaren allerdings Unterschiede bei den Aufgabenzuordnungen zu den einzelnen Ressourcen und somit auch in der Verknüpfung von Umlauf- und Dienstebene. Abbildung 11.5 stellt exemplarisch für Testinstanz B den im Vergleich zur isolierten Umlaufbetrachtung durch Einbeziehen der Dienstebene zusätzlich entstehenden Verspätungspropagationseffekt bei sequenzieller und teilintegrierter Planung dar. Bild I zeigt den Effekt im direkten Vergleich der Ressourceneinsatzpläne, die basierend auf von Kostenstruktur und Aufgaben identischen Umlaufplänen erstellt wurden. Die Ergebnisse sind gemäß der Reihung in Tabelle 11.3 angeordnet. In Bild II ist die in diesen Ressourceneinsatzplänen vorgesehenen Anzahl an Dienstwechseln innerhalb von Umläufen dem entstehenden Effekt gegenübergestellt.

Der Vergleich in Bild I verdeutlicht, dass bei teilintegrierter Planung der zusätzliche Effekt auf die Verspätungspropagation in Bezug zur Verspätung auf Umlaufebene geringer als bei sequenzieller Planung gehalten werden kann. Dies ist zum Teil der aufgrund des höheren Ressourceneinsatzes häufig leicht höheren Anzahl Dienstwechsel in den sequenziell erstellten Plänen geschuldet. Zum anderen ist dies allerdings auf eine geschicktere Verknüpfung des Ressourceneinsatzes bei teilintegrierter Planung zurückzuführen. Bild II verdeutlicht exemplarisch den Einfluss der geplanten Dienstwechsel auf den Propagationseffekt. Tendenziell steigt dieser mit zunehmender Anzahl der Dienstwechsel. Aber auch bei nahezu gleicher Anzahl Dienstwechsel sind deutliche Unterschiede bezüglich des Zusatzeffektes zwischen sequenziell und teilintegriert erstellten Plänen zu erkennen.

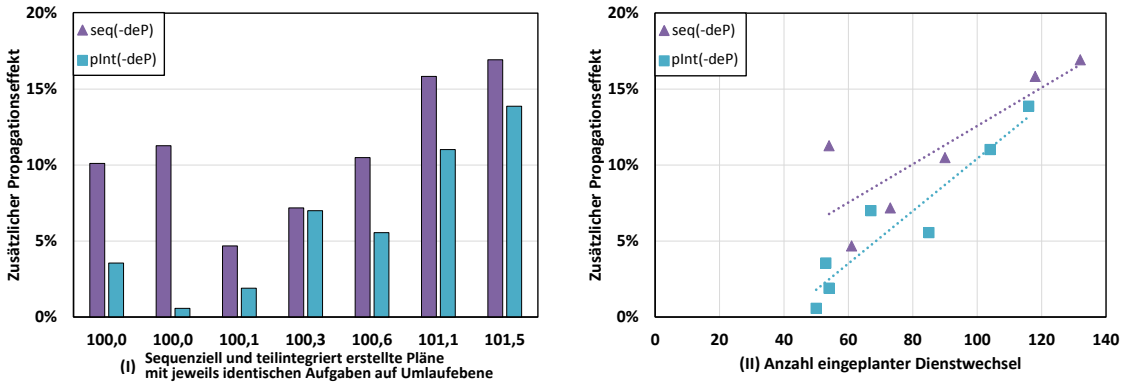

Abbildung 11.5: Effektgröße zusätzlicher Verspätungspropagation durch Einbeziehen der Dienstebene gegenüber der Propagation auf Umlaufebene im Vergleich von sequenziell und teilintegriert erstellten Plänen am Beispiel der Testinstanz B

Kernaussagen Zusammengefasst lässt sich festhalten, dass die Vorgaben in der Umlaufebene einen starken Einfluss auf die insgesamt erreichbare Robustheit (im Sinne von Stabilität), sowie bekannterweise auf die Kosteneffizienz des Ressourceneinsatzes haben. Unabhängig vom Lösungsansatz wird der Spielraum für Robustheit hauptsächlich durch die eingeplanten Aufgabensequenzen der Umlaufebene vorgegeben. Allerdings ist bei der Planung eine zusätzlich durch das Einbeziehen von Diensten entstehende Unrobustheit zu berücksichtigen. Im Vergleich zwischen sequenzieller und teilintegrierter Planung zeigt sich, dass robustere Umlaufpläne (einer sequenziellen Planung) nicht zwangsläufig zu robusterem Gesamtressourceneinsatz führen. Die zusätzlichen Freiheitsgrade in der teilintegrierten Planung ermöglichen den Umgang mit durch Vernetzung der Ressourcen zusätzlich auftretenden Propagationseffekten.

11.2.3 Gemeinsame Planung von Umläufen und Diensten im Vergleich

Im vorangegangenen Abschnitt 11.2.2 wurde gezeigt, dass eine Teilintegration der Planungsstufen zur Umlauf- und Dienstplanung aufgrund der größeren Freiheitsgrade gegenüber einer sequenziellen Planung zu einer besseren Bewertung der Pläne hinsichtlich Robuster Effizienz führen kann. In diesem Abschnitt werden die durch das Einbeziehen weiterer Freiheitsgrade durch eine integrierte Planung von Umläufen und Diensten erzielten Ergebnisse mit den zuvor beschriebenen Ergebnissen der umlaufbasierten Planung verglichen.

Untersuchungsfragen Grundlage für die Untersuchungen bilden folgende Fragestellungen:

IVCS 1.) In welchem Ausmaß unterscheiden sich die einzelnen Planungsmethoden bezüglich des realisierbaren Trade-Off zwischen Kosteneffizienz und Robustheit?

IVCS 2.) Welches Potenzial bietet eine integrierte Planung gegenüber umlaufbasierter Planung des Ressourceneinsatzes?

IVCS 3.) In welchem Ausmaß kann in kostenoptimalen/kosteneffizienten Plänen Robustheit realisiert werden?

IVCS 4.) Inwiefern erklärt der Umfang der jeweils für einen Umlauf- und Dienst gemeinsam eingeplanten Aufgabensequenzen (zur Aggregation von Fahrten) in den Lösungen Unterschiede der Lösungsqualität?

Genutzte Artefakte Für den Vergleich der Planungsmethoden werden die in den beiden vorangegangenen Teilkapiteln aufgeführten Ergebnisse zusammengeführt. Neben den im vorherigen Abschnitt 11.2.2 adressierten Lösungsverfahren und den damit erstellten Plänen werden Ressourceneinsatzpläne betrachtet, die durch eine integrierte Planung mit Berücksichtigen von Verspätungspropagation bzw. durch rein kostenminimierende integrierte Planung erstellt werden (vgl. Abschnitt 11.2.1). Durch Variationen in der Gewichtung der Zielsetzungen wurden jeweils im Hinblick auf Robustheit und Kosteneffizienz unterschiedlich ausgeprägte Umlauf- und Dienstpläne erstellt. Wie zuvor wird bei der Simulation des Ressourceneinsatzes allein die Stabilität der Pläne evaluiert.

Ergebnisdarstellung Die Abbildungen 11.6 stellen die potenzielle Robustheit der mit sequenziellen, teilintegrierten und integrierten Planungsmethoden erstellten Umlauf- und Dienstpläne ihrer geplanten Kosteneffizienz gegenüber. Der Einfluss bzw. die Unterschiede der Planungsmethoden werden exemplarisch veranschaulicht an den Instanzen A, B und C. Als Basisvergleichswert dient pro Instanz eine integrierte, rein kostenminimierende Planung gegen die geplante Kosten und potenzielle Verspätungsprogagation der anderen Ressourceneinsatzpläne abgetragen werden. Für jede Planungsmethode ist eine rein kostenminimierende Planung ohne Berücksichtigung von Robustheitsaspekten aufgeführt (als weißes Symbol mit schwarzem Rand). Dem jeweils gegenüber gestellt wird die Ergebnisqualität der Ressourceneinsatzpläne, die durch Berücksichtigung von Verspätungspropagation in der Planung erstellt wurden.

Ergänzend gibt Tabelle 11.4 einen Überblick über die Anzahl eingeplanter Umläufe (*#v*) und Dienste (*#d*) für (ausgewählte) Lösungen aus Abbildung 11.6. Wie in der Abbildung wird zwischen sequenziell (*SEQ*), teilintegriert (*PINT*) und integriert (*INT*) erstellten Lösungen unterschieden. Zusätzlich zu dem eingeplanten Ressourcenumfang wird für jede aufgeführte Lösung die erwartete Verspätungspropagation (*E(SD)*) mit Bezug zum jeweils korrespondierenden Vergleichswert einer rein kostenminierenden integrierten Planung aufgeführt. Für jede Probleminstanz zeigt die erste Zeile in Tabelle 11.4 jeweils Lösungen, die durch rein kosteneffiziente Planung erstellt wurden. Alle anderen Zeilen umfassen Lösungen, die durch Planung mit Einbeziehen von Verspätungspropagation erstellt wurden.

Zusätzlich fasst Tabelle 11.5 die Rechenzeiten der einzelnen Planungsmethoden in Lösungsgruppen pro Instanz zusammen. Für jedes Planungsschema (*SEQ*, *PINT*, *INT*) werden die Zeiten aggregiert, die benötigt wurden, um die in Abbildung 11.6 aufgeführten Lösungen mit Berücksichtigung von Verspätungspropagation zu erstellen. Für jede so gebildete Gruppe von Lösungen wird die durchschnittliche (*cpu_avg*), die minimale (*cpu_min*) und die maximale (*cpu_max*) Gesamtlaufzeit in Stunden, Minuten und Sekunden angegeben. Die Laufzeiten werden je Planungsschema mit der Laufzeit zur Berechnung der entsprechenden Basislösung verglichen (*cpu_baseline*), die durch rein kostenminimierende Planung ohne Berücksichtigung von Verspätungspropagation erstellt wurde.

Ergebnisinterpretation (ad IVCS 1.) Die Abbildungen 11.6 zeigen, dass sich die mit unterschiedlichen Planungsmethoden berechneten Umlauf- und Dienstpläne jeweils in ihrer Gesamtheit gesehen bezüglich der Ergebnisqualität deutlich voneinander unterscheiden. So dominieren sich die erstellten Pläne in Abhängigkeit der genutzten Freiheitsgrade in der Planung: Aufgrund geringerer (Dienst-)Kosten durch größere Freiheitsgrade und besserer Verteilung der Aufgaben bzw. Pufferzeiten dominieren teilintegrierte Lösungen die sequenziell errechneten Ressourceneinsatzpläne. Gleiches gilt für integrierte Lösungen gegenüber teilintegriert errechneten Plänen.

Insbesondere bei vollständig integrierter Planung von Umläufen und Diensten lassen sich mit nur marginalen Mehrkosten kontrolliert wesentlich robustere Ressourceneinsatzpläne erstellen, ohne an der grundlegenden Kosteneffizienz zu ändern. Die sehr steil abfallende Form der entstehenden Pareto-Fronten für integriert erstellte Umlauf- und Dienstpläne unterstreicht dies. Im Gegensatz dazu erscheinen die abgebildeten Fronten bzw. Lösungen von sequenziell bzw. teilintegriert erstellten Plänen ungeordnet. Als Grundlage für deren Erstellung dienen die ebenfalls sehr steilen Lösungsfronten der Umlaufpläne aus Abbildung 11.3 bzw. die korrespondierenden Umlauf-

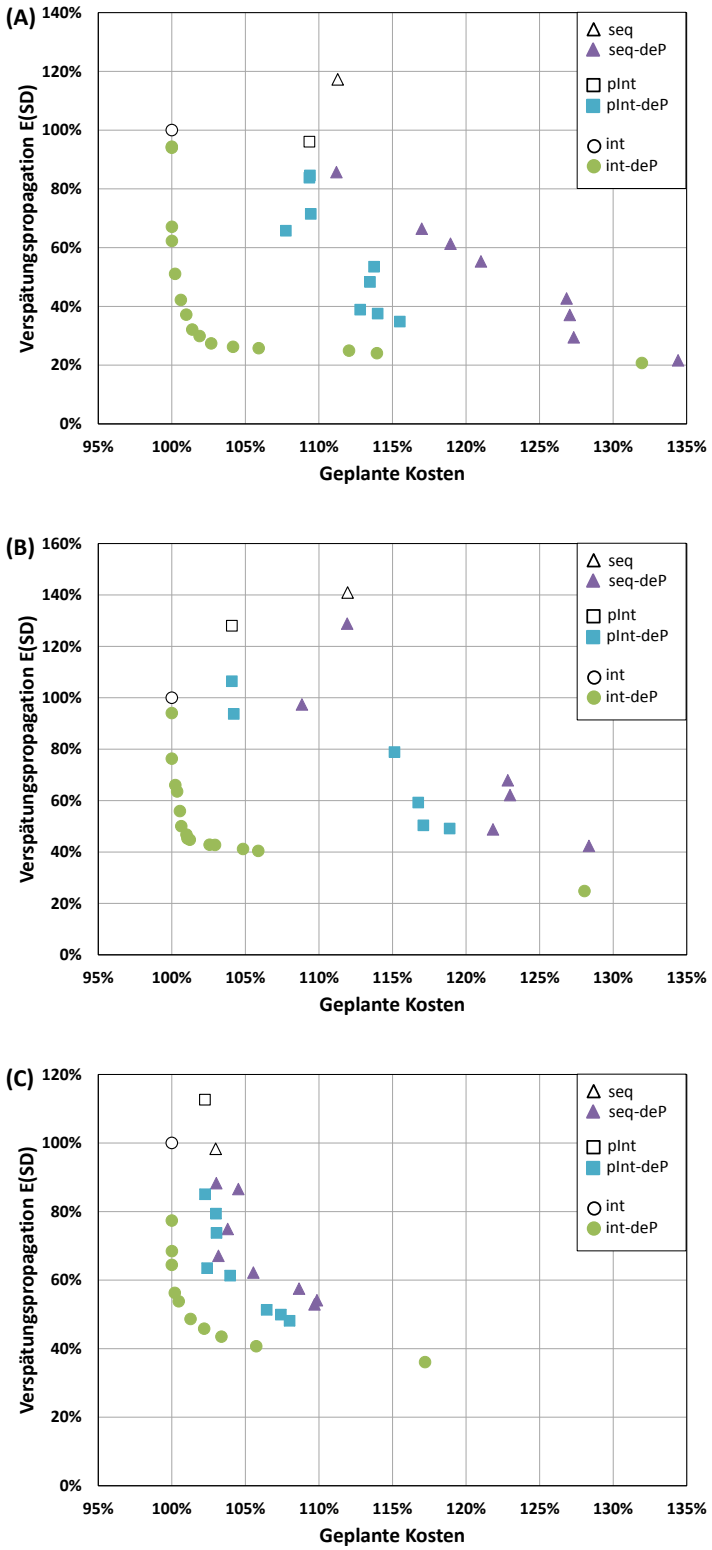

Abbildung 11.6: Gegenüberstellung geplanter Kosten und erwarteter Verspätungspropagation bei sequenzieller (seq), teilintegrierter (pInt) und integrierter Planung (int) mit Berücksichtigen von Verspätungspropagation (deP)

Tabelle 11.4: Gegenüberstellung von Umlauf- und Dienstanzahl und erwarteter Verspätungspropagation für ausgewählte Lösungen aus Abbildung 11.6

	—SEQ—			—PINT—			—INT—		
instance	#v	#d	E(SD) in %	#v	#d	E(SD) in %	#v	#d	E(SD) in %
A	17	31	117.3	17	30	96.1	17	25	100.0
	17	31	85.7	17	30	83.8	17	25	94.0
	17	34	66.5	17	29	65.8	17	25	62.3
	17	36	55.3	17	32	48.3	17	25	51.1
	17	39	42.7	17	32	38.9	17	25	42.4
	17	39	37.1	17	32	37.6	17	25	37.2
	17	39	29.5	17	32	34.8	17	26	26.2
B	21	37	141.0	21	32	128.0	21	29	100.0
	21	37	128.9	21	32	106.5	21	29	94.1
	21	35	97.4	21	32	93.7	21	29	76.3
	21	44	67.9	21	39	78.9	21	29	55.9
	21	44	62.1	21	40	59.2	21	29	45.3
	21	43	48.8	21	40	50.4	21	30	42.8
	21	47	42.2	21	41	49.1	21	31	41.2
C	33	70	98.3	33	69	112.6	33	65	100.0
	33	70	88.3	33	69	85.1	33	65	77.4
	33	71	74.9	33	70	73.8	33	65	68.5
	33	70	67.1	33	69	63.5	33	65	64.4
	33	73	62.2	33	71	61.3	33	65	53.8
	33	77	54.1	33	75	49.9	33	66	48.7
	33	78	52.9	33	75	48.1	33	69	43.5

Tabelle 11.5: Rechenzeiten der verschiedenen Lösungsschemata im Überblick (in hh:mm:ss)

instance	scheme	cpu_avg	cpu_min	cpu_max	cpu_baseline
A	SEQ	00:02:11	00:00:47	00:05:03	00:00:21
	PINT	00:02:29	00:00:32	00:06:07	00:00:30
	INT	00:15:33	00:08:13	00:31:31	00:07:44
B	SEQ	00:04:17	00:00:27	00:08:10	00:00:22
	PINT	00:06:35	00:01:49	00:12:04	00:03:50
	INT	00:41:04	00:28:17	01:07:56	00:31:59
C	SEQ	00:11:20	00:02:26	00:23:45	00:02:15
	PINT	00:18:25	00:06:36	00:30:20	00:06:33
	INT	01:48:23	01:02:49	03:24:11	01:09:01

planbündel. Aufgrund der Vorgaben aus der Umlaufplanung lässt sich allerdings nicht immer ein deutlicher Anstieg von Robustheit bei gleichzeitig nur marginaler Erhöhung der geplanten Gesamtkosten verwirklichen.

(ad IVCS 2.) Zählt allein die Robustheit der Pläne, lassen sich mit einem robusten Umlaufplan bzw. einem Bündel von Umlaufplänen als Grundlage ähnlich robuste Pläne erstellen wie mit integrierter robuster Planung – nur zu teureren Gesamtkosten aufgrund eines erhöhten Diensteinsatzes. Wie im Vergleich zwischen den Umlaufplänen aus sequenzieller und teilintegrierter Planung in Abschnitt 11.2.2 ist auch im Vergleich zwischen Umlaufplänen aus sequenzieller und integrierter Planung ein sequenzieller Umlaufplan mit gleichen Kosten wie ein integriert erstellter Umlaufplan isoliert für sich betrachtet aufgrund der Planungsmethodik robuster. In Kombination mit dem zugehörigen Dienstplan verschiebt sich diese Bewertung zugunsten der integriert erstellten Lösung.

Mit Blick auf Tabelle 11.4 wird deutlich, dass durch das Berücksichtigen von Verspätungspropagation in den Lösungsansätzen jedes Planungsschema (SEQ, PINT, INT) in der Lage ist, Robustheit zu erhöhen ohne zusätzliche Ressourcen im Vergleich zu einer rein kostenminimierenden Planung in Anspruch zu nehmen. Allerdings schränkt die zugrunde liegende Umlaufplanlösung die Freiheitsgrade bei einer sequenziellen und teilintegrierten Planung ein. Infolgedessen wird eine drastische Reduktion von Verspätungspropagation in der Regel bei sequenzieller und teilintegrierter Planung durch zusätzliche Dienste erreicht. Im Gegensatz dazu ist die robuste integrierte Planung in der Lage, Verspätungspropagation drastisch zu reduzieren ohne zusätzliche Ressourcen einzusetzen.

Ergänzend zeigt Tabelle 11.5, dass das Berücksichtigen von Verspätungspropagation im Allgemeinen zu höheren Laufzeiten führt – verglichen zu rein kostenminimierender Planung. Verfahrensgemäß führt die integrierte Planung ebenfalls zu einer Laufzeiterhöhung im Vergleich zu sequenzieller und teilintegrierter Planung. Trotz der zusätzlichen Berücksichtigung von Verspätungspropagation bleiben die vorgeschlagenen Lösungsansätze allerdings immer noch rechentechnisch attraktiv und performant. Im Vergleich zu Echtzeit-Anwendungen (wie z.B. dem Umplanen im Fall einer schwerwiegenden Störung), wird die Umlauf- und Dienstplanung gewöhnlicherweise mehrere Wochen vor Ausführung der Pläne durchgeführt, so dass auch Laufzeiten von mehreren Stunden annehmbar wären. Für zusätzliche Techniken zur Beschleunigung der Laufzeiten sei auf den Abschnitt 9.4.3 über Kantengenerierung während des Lösungsverfahrens zur integrierten Planung mit Zeitfenstern hingewiesen. Die Kantengenerierung ist gleichermaßen für die normale integrierte Umlauf- und Dienstplanung einsetzbar (siehe Tabelle 9.2 für exemplarische Ergebnisse).

(ad IVCS 3.) Darüber hinaus wird nochmals deutlich, dass in allen Planungsvarianten zu der kostenoptimalen bzw. kostenminimierenden Lösung ein robustes Gegenstück mit absolut identischen Aufgaben bzw. Kostenstruktur berechnet werden kann. Dies gilt insbesondere bei teilintegriert und integrierter Planung für die Kombination aus Umlauf- und Dienstplan. Aufgrund einer anderen Verteilung der Aufgaben auf die einzelnen Ressourcen ist diese Lösung bereits (deutlich) robuster als der rein kostenminimierend erstellte Ressourceneinsatz.

(ad IVCS 1.) Vergleicht man die Lösungen instanzenübergreifend, fällt auf, dass die Abstände zwischen den Fronten bzw. den Lösungen der verschiedenen Planungsmethoden über die einzelnen Instanzen gesehen variieren können. Dies liegt zum einen darin begründet, dass die Bemessungsgrundlage für die Bewertung des gesamten Ressourceneinsatzes unterschiedlich ist. So fallen für einige Instanzen zusätzliche Dienste bzw. Umläufe bei der aggregierten Betrachtung in der Abbildung 11.6 nicht so stark ins Gewicht (siehe Tabelle 11.4, welche die jeweils benötigte Fahrzeug- und Dienstanzahl exemplarisch für die Instanzen aufschlüsselt). Zum anderen können in Abhängigkeit der Servicefahrten und vorhandener Auslastungsspitzen bei einigen Instanzen Vorgaben aus der Umlaufplanung bei sequenziellen und teilintegrierten Planungsmethoden eher den Einsatz zusätzlicher Dienste erfordern, so dass die Plankosten rascher steigen (Exemplarisch sei hier auf die Unterschiede der Ganglinien der Instanzen B und C in den späteren Abbildungen 11.15 und 11.17 verwiesen). Überdies ist das Potenzial zur Reduktion der Verspätungspropagation durch die Fahrplanvorgaben verschieden ausgeprägt. Auf die Beschaffenheit der durch die Fahrplanung für die Verknüpfung von Servicefahrten je Instanz unterschiedlich vorhandenen Pufferzeiten wird in Abschnitt 11.3.2 speziell Abbildung 11.12 mit Blick auf die Planstrukturen des jeweils erstellten Ressourceneinsatzes näher eingegangen. Die Veränderung der potenziellen Robustheit und Kosteneffizienz durch Eingriffe in die Fahrplanung wird in Teilkapitel 11.4 diskutiert.

(ad IVCS 2.) Insgesamt zeigt sich, dass bei jeder Instanz integrierte Verfahren bessere Lösungen liefern können als teilintegrierte bzw. sequenzielle Lösungsverfahren. Was zeichnet also nun den integriert geplanten Ressourceneinsatz gegenüber den anderen Lösungen aus?

Planstrukturen (ad IVCS 4.) Die Strukturen der errechneten Ressourceneinsatzpläne verdeutlichen, warum integriert erstellte Umlauf- und Dienstpläne sequenziell bzw. teilintegriert erstellten Ressourceneinsatz in der Ergebnisqualität bezüglich Robuster Effizienz dominieren. Abbildung 11.7 stellt exemplarisch für Testinstanz C in Bild I die Warte- und Standzeit, die gemeinsam für sowohl Fahrer und Fahrzeug

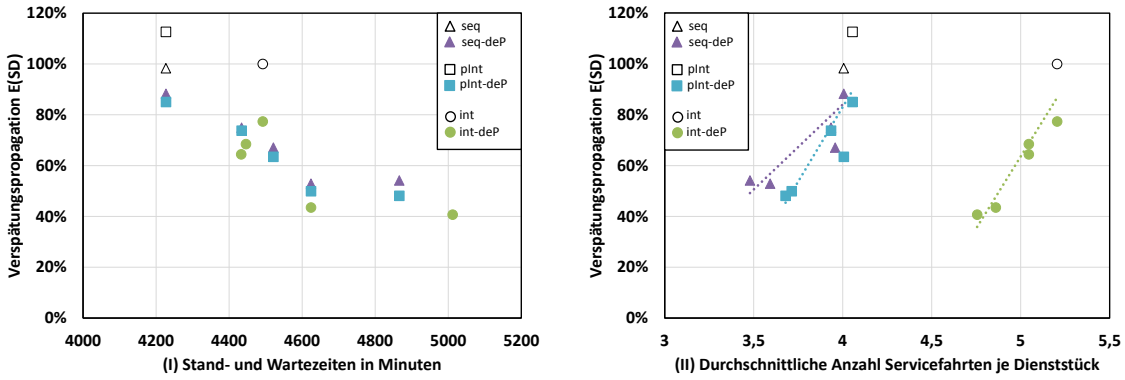

Abbildung 11.7: Eingeplante Pufferzeiten, Anzahl Servicefahrten vor Dienstpausen und potenzielle Verspätungspropagation im Vergleich der Planungsvarianten am Beispiel der Testinstanz C

eingeplant werden, und die potenzielle Verspätungspropagation einzelner Lösungen gegenüber. Bild II ergänzt die Analyse dieser Lösungen durch einen Vergleich über die durchschnittliche Anzahl Servicefahrten, die ein Fahrer bis zu einer Pause innerhalb eines Dienststückes auf einem Fahrzeug auszuführt, und der potenziellen Verspätungspropagation.

Der Vergleich in Bild I zeigt, dass mehr Pufferzeit tendenziell zu robusteren Plänen führt. Die summierten Gesamtpufferzeiten in den betrachteten sequenziell, teilintegriert und integriert erstellten Lösungen sind darüber hinaus ähnlich zueinander. Gleichzeitig enthalten die integriert erstellten Pläne weniger Dienste, wodurch die Anzahl der pro Dienst zu erledigenden Aufgaben steigt. Bild II verdeutlicht, dass in den integriert erstellten Lösungen deutlich längere Fahrtensequenzen von Servicefahrten innerhalb eines Dienststückes auszuführen sind. Erkennbar steigt die Robustheit der einzelnen Pläne mit abfallender Tendenz auszuführender Fahrtenfolgen, da somit weniger Verspätungen kumuliert und propagiert werden können. Auffallend ist, dass bei integriert erstellten Lösungen trotz längerer Fahrtensequenzen, robustere bzw. gleich robuste Ressourceneinsatzpläne möglich sind im Vergleich zu sequenziell oder teilintegriert erstellten Lösungen.

Kernaussagen Die Ausmaße von Robustheit und Kosteneffizienz des Ressourceneinsatzes werden deutlich durch die Freiheitsgrade der Planungsmethoden bestimmt. Eine Verbesserung hinsichtlich sowohl Robustheit als auch Kosteneffizienz wird dabei durch zusätzliche Freiheitsgrade in der Planung erreicht. Robuste Umlaufpläne – erstellt in sequenzieller (bzw. teilintegrierter Planung) – eignen sich auch für eine robuste Gesamtplanung, wobei höhere Dienstplankosten in Kauf genommen wer-

den müssen. Diese Kosten können durch teilintegrierte Planung reduziert werden, bei gleichbleibender Robustheit. Integrierte Planung mit Berücksichtigung von Verspätungspropagation liefert in Bezug auf beide Zielgrößen die besten Lösungen. Dazu werden aufgrund der steigenden Problemkomplexität deutlich höhere Laufzeiten benötigt. Die integriert erstellten Ressourceneinsatzpläne zeichnet darüber hinaus eine höhere Aufgaben- bzw. Servicefahrtendichte pro Dienststück aus. Dies erklärt zum einen den geringeren Ressourceneinsatz. Zum anderen wird verfügbarer Puffer zwischen den Aufgaben besser verteilt, da trotz längerer Aufgabensequenzen geringere Verspätungspropagation erzielt werden kann.

11.2.4 Zusammenfassung der Untersuchungen

In diesem Abschnitt wurde der Einfluss einer sequenziellen, teilintegrierten und integrierten Planungsmethodik auf die Robustheit (Stabilität) und Kosteneffizienz der resultierenden Umlauf- und Dienstplanlösungen untersucht. Durch die Variation auf Ebene der Planungsmethoden und die jeweilige Variation mit verschiedenen Gewichtungen der Zielsetzungen bei Ausführung einzelner Planungsmethoden wurden unterschiedlich ausgeprägte Ressourceneinsatzpläne erstellt. Neben der Evaluation der in Teil II dieser Arbeit vorgeschlagenen Optimierungsverfahren zur Planerstellung wurden so verschiedene Effekte zum Einfluss der Planungsmethodik auf Robustheit und Kosteneffizienz der ertellten Ressourceneinsatzpläne aufgezeigt.

(ad F1.) Dabei stellt das Abbilden von Verspätungspropagation in den einzelnen Lösungsansätzen eine Verbesserung zu existierenden Lösungsmethoden dar. Durch die methodengeführte, geschickte Verteilung der Pufferzeiten in den Umlauf- und Dienstplänen lassen sich im Vergleich zu rein kostenminierender Planung zu nahezu identischen Kosten wesentlich robustere Pläne erstellen. Durch die in Abstimmung mit dem Ressourceneinsatz passgenauere Verteilung von Pufferzeiten bietet dieses Vorgehen einen zusätzlichen Vorteil gegenüber dem Berücksichtigen von (historischen) Verspätungsinformationen durch feste Pufferzeitenvorgaben bei der Planung.

Wie bei reiner kostenminierender Ressourceneinsatzplanung ist es auch bei robuster Planung vorteilhafter, Dienstplanung und Umlaufplanung kombiniert zu betrachten. So werden einseitig zur Erhöhung von Robustheit auf einer Planungsstufe getroffene Entscheidungen vermieden, die zu deutlich höherem Ressourceneinsatz auf der anderen Planungsebene führen können. Zum anderen können bei der Planung die gegenseitigen Abhängigkeiten von Umlauf- und Dienstplan berücksichtigt werden. Insbesondere können so die unterschiedlichen Verknüpfungen des Ressourceneinsatzes und daraus resultierende potenzielle Verspätunspropagation bzw. -absorption beeinflusst werden. Dabei ist die integrierte Planung als Planungsmethode mit den größten

Freiheitsgraden der sequenziellen und teilintegrierten Planung bezüglich der erreichbaren Robustheit im Verhältnis zum Ressourceneinsatz überlegen. Bedingt durch die Freiheitsgrade zeigt sich eine abstufende Dominanz der Planungsmethoden untereinander.

(ad F2.) Es zeigt sich, dass aufgrund der gewählten Modellierung mit zugrunde liegender TSN-Formulierung und daraus folgender Nutzung von Lösungsflüssen bereits kostenoptimale bzw. -effiziente und robuste Lösungen erstellt werden können, die allerdings mit bisherigen Verfahren nicht aus der entstehenden Lösungsmenge herausgearbeitet wurden. Die spezifische Berücksichtigung von Robustheitsaspekten – speziell Verspätungspropagation – in den während der Lösungsverfahren eingesetzten Dekompositionsstragien erlaubt aus der Vielzahl möglicher, von den geplanten Kosten gleichwertiger Lösungen diejenige auszuwählen, die bei gleichen Kosten die potenzielle Verspätungspropagation minimiert.

(ad F3.) Bei der Analyse der Lösungen wurden als Ursachen für die unterschiedlichen Ergebnisse der verschiedenen Planungsmethoden sowie die unterschiedlich robust ausgeprägten Lösungen Veränderungen in den unterliegenden Planstrukturen verdeutlicht. Dabei wurde ein Bezug zu den eingeplanten Pufferzeiten, der Verknüpfung der Umlauf- und Dienstebene durch eingeplante Dienstwechsel, sowie der Dichte der gemeinsam von Dienst und Umlauf auszuführenden Aufgabensequenzen bzw. Servicefahrten herausgestellt. Die Untersuchung der sich in den einzelnen Lösungen ergebenden Planstrukturen sowie ihr Einfluss auf Verspätungspropagation bzw. -absorption wird im folgenden Teilkapitel 11.3 im Hinblick auf eine Beantwortung der Forschungsfrage F3 weiter vertieft.

11.3 Einflussfaktoren in den Planstrukturen auf Verspätungspropagation

Im vorangegangenen Teilkapitel 11.2 wurde der unterschiedliche Einfluss der sequenziellen, teilintegrierten und integrierten Planungsmethoden auf Robustheit und Effizienz der entstehenden Umlauf- und Dienstpläne herausgearbeitet. Neben der Dominanz der Planungsmethoden untereinander in Bezug auf die erreichbare Lösungsqualität wurde dabei auf die entstehenden Planstrukturen eingegangen. So wurde zum einen gezeigt, dass die verschiedenen Planungsvarianten die Umlauf- und Dienstebene unterschiedlich miteinander verknüpfen (vgl. Abschnitt 11.2.2). Zum anderen wurde gezeigt, dass gemeinsame Aufgabensequenzen für Fahrer und Fahrzeuge in unterschiedlichen Ausmaß auf Umläufen und Diensten eingeplant werden (vgl. Ab-

schnitt 11.2.3). Daran anknüpfend werden in diesem Abschnitt die planimmanenten Ursachen für Verspätungspropagation bzw. -absorption im Ressourceneinsatz näher bestimmt. Mit Fokus auf Forschungsfrage F3 wird das Zusammenspiel von eingeplanten Aufgabensequenzen, Dienst- bzw. Umlaufwechseln und vorgesehenen Pufferzeiten auf Umläufen und Diensten in robusten und kosteneffizienten Plänen untersucht.

Für die Untersuchung werden zwei wesentliche Struktureigenschaften der entstandenen Ressourceneinsatzpläne herausgestellt: Die *Vernetzung* der Ressourcen durch eingeplante Umlauf- bzw. Dienstwechsel und die *Verdichtung* (Aggregation) von Fahrten in Dienststücken. Ihr jeweiliger Einfluss auf die Weitergabe von Verspätungen wird in Abschnitt 11.3.1 analysiert und der resultierenden Gesamtverspätungspropagation gegenübergestellt. Im darauf folgenden Abschnitt 11.3.2 werden anhand der zuvor beschriebenen Effekte mit Bezug zu den zugrunde liegenden Planstrukuren Kennzahlen für die Beurteilung von robusten (und kosteneffizienten) Ressourceneinsatzplänen herausgearbeitet. Abschnitt 11.3.3 fasst die Untersuchungen zusammen.

11.3.1 Verdichtungs- und Vernetzungseffekt der Ressourcenebenen

In diesem Abschnitt wird der Einfluss der *Vernetzung* der Ressourcen durch vorgesehene Fahrer- und Fahrzeugwechsel, sowie der Einfluss der *Verdichtung* von gemeinsam auszuführenden Aufgabensequenzen bzw. Fahrten in Dienststücken auf Verspätungspropagation genauer untersucht. Gegenstand der Untersuchungen ist daher sowohl die Weitergabe von Verspätungen zwischen als auch innerhalb von Dienststücken.

Mit Dienststücken und den darin zusammengefassten Aufgaben werden die kleinsten Aufgabensequenzen betrachtet, für die Fahrer und Fahrzeuge – im Rahmen dieser Arbeit – eine Einheit bilden und die sie gemeinsam ausführen (vgl. Abschnitt 2.3). Die allein innerhalb der Dienststücke gemessene Verspätungspropagation gibt daher eine untere Schranke der potenziellen gesamten Verspätungspropagation von Umläufen und Diensten wieder. Zusätzlich gibt sie einen Hinweis auf eine mögliche Grundrobustheit des gesamten Ressourceneinsatzes, da sie verdeutlicht, inwiefern die gemeinsamen Aufgabensequenzen einzelner Umläufe und Dienste in der Lage sind, Verspätungen zu absorbieren. Die Differenz zur gesamten Verspätungspropagation kann dem Einfluss der Vernetzung unterschiedlicher Dienststücke bzw. Ressourcen zugerechnet werden.

Untersuchungsfragen Die folgenden Fragestellungen bilden die Grundlage für die durchgeführten Experimente:

VD-VN 1.) Was wiegt stärker auf die gesamte Verspätungspropagation: die Verdichtung (Aggregation) von Fahrtensequenzen auf Dienststücke oder die Vernetzung der Ressourcen durch eingeplante Umlauf- bzw. Dienstwechsel?

VD-VN 2.) In welchem Ausmaß unterscheiden sich die Anteile der Propagationseffekte in Abhängigkeit der eingesetzten Planungsmethoden?

VD-VN 3.) Inwiefern unterscheiden sich die zur Verfügung stehenden Pufferzeiten bei gemeinsam von Umlauf- und Diensten auszuführenden Aufgabensequenzen innerhalb von Dienststücken?

VD-VN 4.) Inwiefern unterscheiden sich die zur Verfügung stehenden Pufferzeiten bei eingeplanten Umlauf- und Dienstwechseln?

Genutzte Artefakte Zur Untersuchung der Fragestellungen werden die Lösungen miteinander verglichen, die durch sequenzielle, teilintegrierte und integrierte Planungsmethoden erstellt wurden und in Abschnitt 11.2.3 aufgeführt sind. Um unterschiedliche Propagationseffekte nachzuweisen, werden bei der Simulation der Ausführung der Servicefahrten jeweils die einbezogenen Ressourcenebenen variiert. Abbildung 11.8 schlüsselt die Varianten der Simulation sowie die einzelnen in dieser Arbeit herausgearbeiteten und definierten Propagationseffekte beispielhaft auf.

Dabei gibt die gemeinsame Simulation von Umlauf- und Dienstplan die gesamte Verspätungspropagation des Ressourceneinsatzes unter Einbeziehen sämtlicher Verknüpfungen verschiedener Ressourcen wieder (siehe Bild III in Abbildung 11.8).

Dem gegenübergestellt wird eine Simulation, bei der einzelne Aufgabensequenzen quasi getrennt voneinander durchgeführt werden (siehe Bild I). Jedes Dienststück und die darin enthaltenen Sequenzen von Servicefahrten werden dann in der Simulation als in sich abgeschlossene Einheit behandelt, die nicht mit anderen Aufgabensequenzen verknüpft ist. Es wird also bei der Simulation angenommen, dass keine Verspätungen zwischen Dienststücken weitergegeben werden und die Ausführung jedes Dienststückes wie geplant starten kann. Die gemessene Verspätungspropagation bezieht sich somit allein auf einzelne, gemeinsam von einem Umlauf- und Dienst ausgeführte Aufgabensequenzen bzw. Servicefahrten, d.h. allein auf die Dienststückebene. Sie ergibt den *Verdichtungspropagationsseffekt* im Vergleich zur gesamten Verspätungspropagation.

Die Differenz zwischen dem Verdichtungspropagationsseffekt und der gesamten Verspätungspropagation kann der Weitergabe von Verspätungen aufgrund der Vernetzung von Dienststücken zugeordnet werden. Dabei kann ein solcher *Vernetzungspropagationseffekt* noch weiter aufgeschlüsselt werden. Es können noch Propagationseffekte

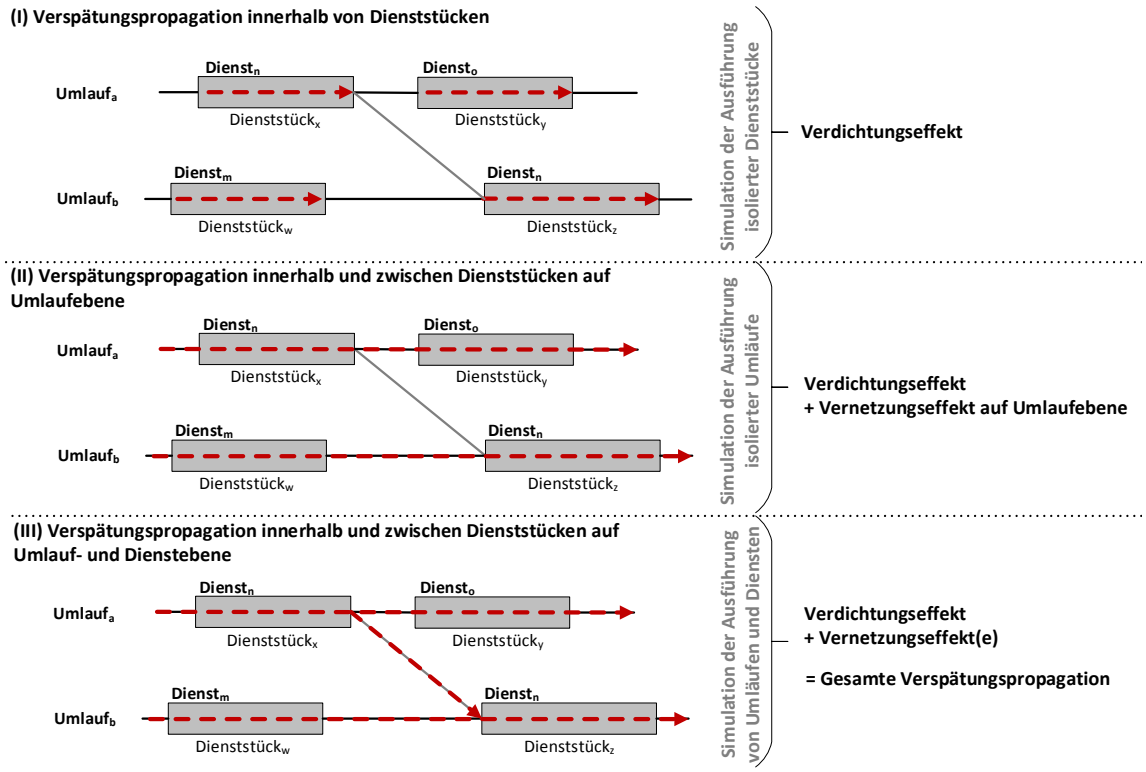

Abbildung 11.8: Aufschlüsselung der gemessenen Verspätungspropagation (*rot gestrichelte Linien*) sowie der nachweisbaren Propagationseffekte in Abhängigkeit der durchgeführten Simulationsvarianten

durch Vernetzung innerhalb eines Umlaufes sowie Propagationseffekte durch Vernetzung über Dienstwechsel zu einem anderen Umlauf unterschieden werden (siehe Bilder II und III in Abbildung 11.8). Der erstgenannte Propagationseffekt zur Vernetzung von Dienststücken innerhalb einzelner Umläufe lässt sich durch eine Simulation evaluieren, bei der die Aufgabensequenzen der Umläufe getrennt voneinander durchgeführt werden. Jeder Umlauf und die darin enthaltenen Sequenzen von Servicefahrten werden als in sich abgeschlossene Einheit behandelt. Die Simulation findet also allein auf Umlaufebene statt. Der zweitgenannte Propagationseffekt zur Vernetzung ergibt sich als Differenz aus der allein auf Umlaufebene gemessenen Verspätungspropagation und der gesamten Verspätungspropagation. [2]

[2] Zu beachten bei der im Rahmen dieser Arbeit getroffenen Unterscheidung der einzelnen Propagationseffekte ist, dass die allein auf Umlaufebene gemessene Verspätungspropagation aus einem anderen Blickwinkel betrachtet ebenso einer *Verdichtung* von Aufgabensequenzen in den jeweiligen Umläufen zugerechnet werden kann. Fokussiert auf reine Umlaufplanung und lösgelöst von der Dienstebene könnte der entstehende Propagationseffekt also ebenso als ein *Verdichtungspropagationseffekt* bezeichnet werden. Da allerdings der gemeinsame Ressourceneinsatz von Fahrzeugen und Fahrern Untersuchungsgegenstand der Arbeit ist, erfolgt die Sicht auf den Ressourceneinsatz aus Perspektive der sowohl Umlauf- als auch Dienstaufgaben gemeinsam verbindenden

Ergebnisdarstellung Die Abbildungen 11.9 stellen den Anteil unterschiedlicher Propagationseffekte an der Geamtverspätungspropagation in den errechneten Ressourceneinsatzplänen dar. Sie vergleichen exemplarisch verschiedene Lösungen der Umlauf- und Dienstplanung in fünf Dimensionen: (1) Pro Zeile sind Ergebnisse einer Testinstanz zusammengetragen. Abgebildet werden Lösungen der Instanzen A, B und C. (2) Die unterschiedlichen Spalten repräsentieren von links nach rechts durch integrierte, teilintegrierte und sequenzielle Umlauf- und Dienstplanung berechnete Lösungen. (3) Pro Planungsschema sind auf der horizontalen Koordinatenachsen für jede abgebildete Lösung die Gesamtkosten des Ressourceneinsatzes im Vergleich zum Basisfall der integrierten Planung mit reiner Kostenminimierung abgetragen. Die Lösungen sind pro Planungsschema von links nach rechts nach aufsteigender Robustheit und eingesetzter Lösungsmethode sortiert. Beginnend mit einer Lösung, die mit der Zielsetzung reiner Kostenminimierung errechnet wurde. Der daran anschließende Eintrag repräsentiert eine Lösung, die im Vergleich dazu aufgabenidentische Einsatzpläne aufweist, allerdings unter Einbeziehen (globaler) Verspätungspropagation bei der Flussdekomposition erstellt wurde. Bei integrierter und teilintegrierter Planung jeweils beim Verknüpfen von Umlauf- und Dienstplan und bei sequenzieller Planung beim Erstellen der Umlaufplanlösung als Grundlage für die Dienstplanung. Es folgen vier Lösungen, bei denen bereits während der Dienst- und Umlauferstellung neben der Kostenminimierung die Minimierung von Verspätungspropagation berücksichtigt wird. (4) Auf der vertikalen Achse ist die Verspätungspropagation der einzelnen Lösungen angegeben. Dabei bezieht sich die Angabe zur besseren Vergleichbarkeit pro Instanz jeweils auf den Basisfall der rein kostenminimierenden, integrierten Planung. (5) Darüber hinaus wird die gesamte Verspätungspropagation aufgeschlüsselt in Anteile, die der Verdichtung (grau) sowie der Vernetzung (Blautöne) des Ressourceneinsatzes zugewiesen werden können. Dabei ist der Anteil der Verdichtung, also der innerhalb von Dienststücken gemessenen Verspätungspropagation, mit weißer Schrift angegeben. Die Differenz zu 100% bildet den sich zusätzlich durch die Vernetzung ergebenden Anteil an der gesamten Verspätungspropagation. Dieser kann wie oben bei Abbildung 11.8 beschrieben in eine Weitergabe von Verspätungen durch Verknüpfung von Dienststücken innerhalb der gebildeten Umläufe (blau) sowie eine Weitergabe von Verspätungen durch Dienstwechsel zu anderen Umläufen (hellblau) unterteilt werden.

Konstrukte, d.h. der Dienststücke. Eine Sequenz von Dienststücken auf der Umlaufebene repräsentiert die Verknüpfung verschiedener Ressourcen bzw. unterschiedlicher Dienste. Deshalb wird in der Analyse diese Verbindung einzelner Dienststücke auf der Umlaufebene einem *Vernetzungspropagationseffekt* zugerechnet. Zumindest bei integrierter Planung werden darüber hinaus Umläufe und Dienste gemeinsam erstellt, weswegen die entstehenden Aufgabensequenzen auf Umlaufebene nicht lösgelöst von der Dienstebene gesehen werden sollten.

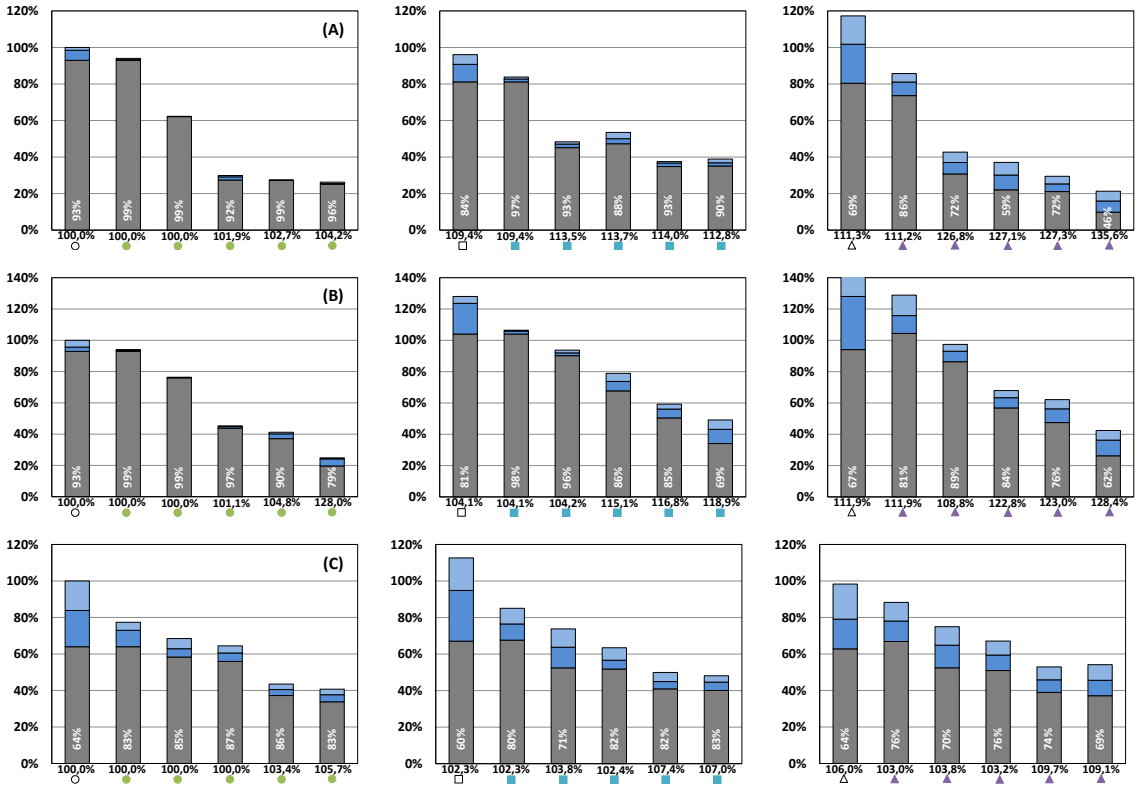

Abbildung 11.9: Gegenüberstellung und Aufschlüsselung von Verspätungspropagation (vertikale Achse) verschiedenartig erstellter Lösungen (horizontale Achse) unterschiedlicher Instanzen (Zeilen) und Planungsschemata (Spalten) in Verdichtungseffekt (*grau*) und Vernetzungseffekte (*hellblau, blau*)

Ergebnisinterpretation (ad VD-VN 1.) Die Abbildungen 11.9 zeigen, dass sich die jeweils den Effekten von Verdichtung und Vernetzung zuordnenbaren Anteile an der gesamten Verspätungspropagation deutlich unterscheiden. Instanzen- und planungsmethodenübergreifend ist die allein auf Dienststückebene gemessene Verspätungspropagation wesentlich größer als die Verspätungspropagation, die zusätzlich durch die Verknüpfung verschiedener Dienststücke bzw. Aufgabensequenzen induziert wird. Auch bei rein kostenminimierend erstellten Lösungen ohne Berücksichtigen von Robustheitsaspekten bei der Planung sind die einer Verdichtung zurechenbaren Anteile an der Verspätungspropagation in der Tendenz etwas größer als die der Vernetzung zurechenbaren. Insbesondere jedoch bei Planung unter Einbeziehen von Robustheitsaspekten bestimmt die Verdichtung auf Dienststückebene erkennbar die Ausrichtung der Robustheit des jeweiligen Ressourceneinsatzes. Ein Großteil der gesamten Verspätungspropagation kann somit jeweils auf einen Verdichtungseffekt zurückgeführt werden. Dabei ist anzumerken, dass die gesamte Ausprägung der Verspätungspropagation nicht allein durch die Verdichtung festgelegt wird, sondern auch durch die Art der Vernetzung der einzelnen Aufgabensequenzen. So kann eine geschickte Vernetzung eine zusätzliche Weitergabe von Verspätungen zwischen Dienststücken und ein daraus resultierendes weiteres Aufschaukeln innerhalb der Aufgabensequenzen von Dienststücken vermeiden. Speziell beim Vergleich von reiner kostenminimierender Planung und den aufgabengleichen Lösungen durch Planung mit Berücksichtigung von globaler Verspätungspropagation fällt auf, dass offensichtlich genügend Freiheitsgrade vorhanden sind, um durch eine andere Zusammenstellung und Verknüpfung der Dienststücke mögliche zusätzliche Verspätungseffekte zu absorbieren, so dass kein bzw. ein deutlich geringerer Vernetzungseffekt auftritt.

(ad VD-VN 2.) Die Anteile und Ausmaße der einzelnen Effekte variieren erkennbar in Abhängigkeit der eingesetzten Planungsmethoden. Bei robuster integrierter Planung ist hauptsächlich eine Propagation von Verspätungen durch Verdichtung zu erkennen, wohingegen kaum bzw. nur geringe Propagation als Vernetzungseffekt auftritt. Der Anteil des gesamten Vernetzungspropagationseffekts ist in der Regel kleiner als bei (robust erstellten) teilintegrierten oder sequenziellen Lösungen.

Darüber hinaus ist im Vergleich der Planungsmethoden bei integrierter Planung zu deutlich geringeren Kosten ein absolut gesehen kleinerer Verdichtungseffekt möglich. Dies zeigt sich im Vergleich integriert erstellter Pläne zu kostenähnlichen Plänen, die durch teilintegrierte oder sequenzielle Planung berechnet wurden (siehe z.B. bei Instanz B die integrierte Lösung mit 104,8% Kosten im Vergleich zu 104,2%, sowie die integrierte Lösung mit 128,0% Kosten im Vergleich zu 128,4%, oder bei Instanz

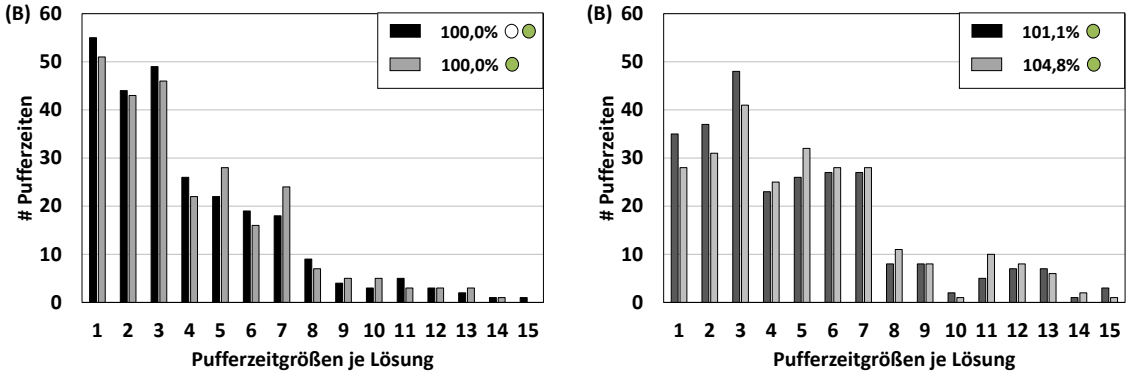

Abbildung 11.10: Häufigkeiten der Pufferzeiten in allen Dienststücken für (robust) integriert erstellte Lösungen exemplarisch für Testinstanz B

C die integrierte Lösung mit 103,4% Kosten im Vergleich zu Lösungen mit 103,8% und 103,2%).

Bei durch teilintegrierte Planung erstellten Lösungen sind wiederum im Vergleich zu sequenzieller Planung mögliche Vernetzungspropagationseffekte in der Regel geringer ausgeprägt (siehe dazu auch die bereits in Abschnitt 11.2.2 durchgeführte Untersuchung zum Einfluss der Verknüpfung von Umlauf- und Dienstebene auf die Robustheit des gesamten Ressourceneinsatzes).

Besonders bei integrierter und teilintegrierter Planung werden folglich gemäß der vorgeschlagenen Lösungsverfahren die Aufgaben so eingeplant, dass geringere Verspätungseffekte von Dienststücken innerhalb von Umläufen (auf andere Dienste) bzw. (durch gleiche Dienste) zu anderen Umläufen weitergegeben werden. Durch integrierte Planung sind darüber hinaus die einzelnen Aufgabensequenzen so gestaltbar, dass geringere Verspätungseffekte innerhalb von Dienststücken weitergegeben werden (siehe zur Verdeutlichung die jeweils dritte aufgeführte integrierte Lösung mit 100,0% geplanter Kosten im Vergleich zur zweiten bzw. ersten).

Planstrukturen Die Planstrukturen der miteinander verglichenen Pläne erklären die unterschiedlichen Ausmaße der in Abbildung 11.9 dargestellten Verdichtungs- und Vernetzungseffekte.

(ad VD-VN 3.) Die Länge der einzelnen Dienststücke sowie die Anzahl der Aufgaben steigt im Vergleich von integrierter, teilintegrierter und sequenzieller Planung bzw. sinkt jeweils mit zunehmender Robustheit (vgl. Erläuterungen der Planstrukturen in Abschnitt 11.2.3 bzw. zu Abbildung 11.7). Neben der insgesamt zur Verfügung stehenden Pufferzeit ist dabei entscheidend für die Weitergabe von Verspätungen innerhalb der Dienststücke die passende Verteilung von Pufferzeiten in Bezug zu

den durchzuführenden Aufgaben bzw. eingeplanten Aufgabenwechseln. Abbildung 11.10 zeigt exemplarisch für Lösungen der Testinstanz B unterschiedliche Größen von Pufferzeiten innerhalb von Dienststücken mit ihren Häufigkeiten aufgeschlüsselt von einer bis fünfzehn Minuten. Verglichen werden mit (robust) integrierter Planung erstellte Lösungen. Dabei werden im linken Bild I die aufgabenidentischen Lösungen rein kostenminimierender Planung sowie kostenminimierender Planung mit anschließender Berücksichtigung von (globaler) Verspätungspropagation bei der Flussdekomposition einer integrierten Planung gegenübergestellt, bei der während der Umlauf- und Diensterstellung sowohl Kostenminimierung als auch eine Minimierung von Verspätungspropagation berücksichtigt wird. Die Gesamtdauer der gemeinsam innerhalb von Dienststücken für Umläufe und Dienste eingeplanten Pufferzeiten ist in diesem Fall fast gleich mit der Gesamtdauer der Pufferzeiten der kostenminimierenden Planung. Den Lösungen in Bild I werden im rechten Bild II Lösungen gegenübergestellt, die insgesamt mehr Pufferzeiten besitzen und deshalb einen (absolut gesehen) deutlich verringerten Verdichtungspropagationseffekt aufweisen.

Während der Unterschied zwischen der Verteilung von Pufferzeiten im Vergleich von Bild II zu Bild I offensichtlich ist, ist der Unterschied der Lösungen in Bild I nicht so stark ausgeprägt. Die insgesamt fast gleiche Pufferzeit ist in den Lösungen ähnlich verteilt. Der Unterschied liegt darin, dass die robustere Lösung etwas weniger Pufferzeiten im Bereich von 1 bis 4 Minuten aufweist, dafür aber ein paar größere Pufferzeiten um 5 bzw. 7 Minuten zur Verfügung stellt. Diese kleinere Abweichung einzelner Pufferzeitlängen führt zu einem (absolut gesehen) deutlich verringerten Verdichtungspropagationseffekt (siehe Abbildung 11.9).

(ad VD-VN 4.) Im vorangegangenen Abschnitt 11.2.2, speziell Abbildung 11.5, wurde bereits der Einfluss der eingeplanten Dienstwechsel(anzahl) im Zusammenhang mit einer zusätzlichen durch Vernetzung der Ressourcen entstehenden Verspätungspropagation thematisiert. Entscheidend für das Ausmaß an Verspätungspropagation ist dabei die Einbettung der Wechsel in um sie eingeplante Wartezeiten bzw. Pufferzeiträume. Die Abbildung 11.11 zeigt exemplarisch die Verteilung von Pufferzeit für im Ressourceneinsatz vorgesehene Dienst- bzw. Umlaufwechsel, die im Folgenden konkretisierend auch als Dienststückwechsel bezeichnet werden. So werden die Pufferzeiten eingeplanter Dienstwechsel auf Umläufen (gleichbedeutend mit Pufferzeiten zwischen Dienststücken auf einem Umlauf) und die bei Umlaufwechsel eines Dienstes in Pausen zusätzlich zur Verfügung stehende Pufferzeit (gleichbedeutend mit Pufferzeit zwischen Dienststücken eines Dienstes auf unterschiedlichen Umläufen) zusammengefasst und in verschiedene Kategorien gemäß der vorhandenen Pufferzeit aufgeschlüsselt. Dargestellt werden Dienststückwechsel ohne zusätzlichen Puffer (*0*), bis

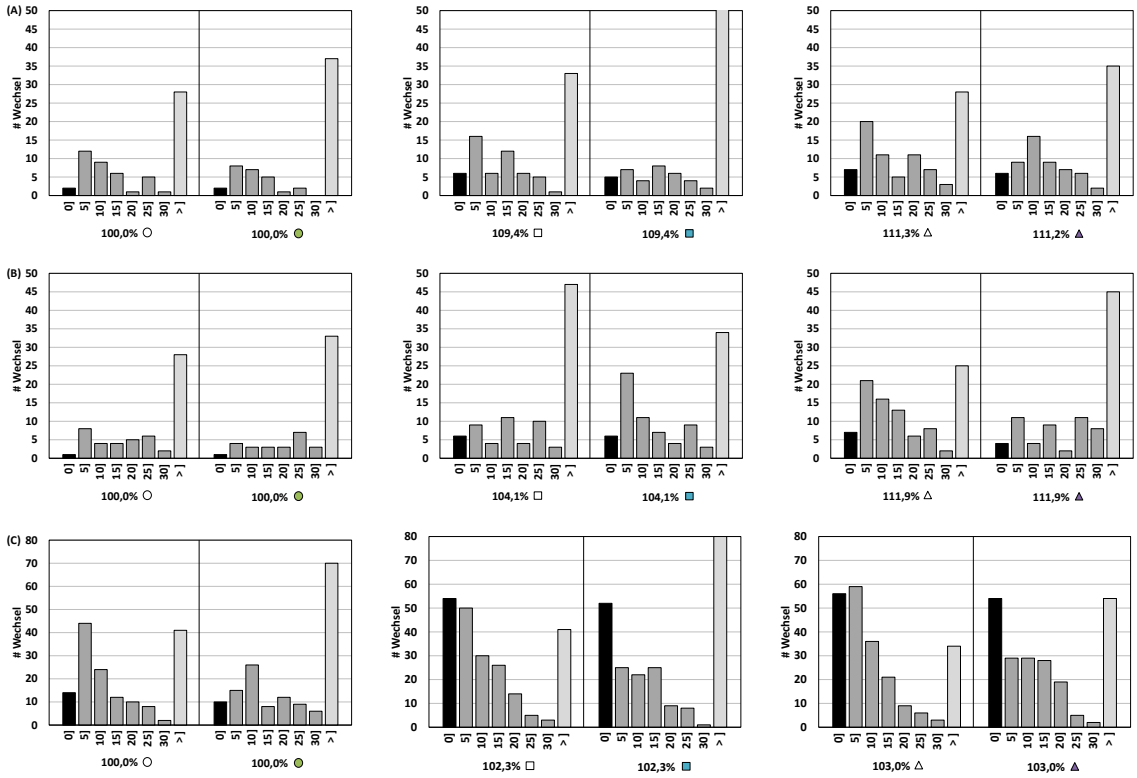

Abbildung 11.11: Anzahl von Dienststückwechseln kategorisiert nach zur Verfügung stehenden Pufferzeiten für integriert, teilintegriert und sequenziell erstellte Lösungen für Testinstanzen A, B, C

zu fünf Minuten Puffer (*5*), sowie in Schritten von fünf Minuten in weitere Zeiträume zwischen fünf und dreißig Minuten (*10, 15, 20, 25, 30*). Zusätzlich werden alle übrigen Wechsel mit mehr als dreißig Minuten Pufferzeit aufgeführt (>).

Eine direkte Vergleichbarkeit der Ergebnisse ist bei den aufgabenidentischen und kostengleichen Plänen gegeben. Daher werden als Beispiele eine rein kostenminimierende Planung gegenüber einer Planung mit Berücksichtigung von Verspätungspropagation durch Anwenden globaler Flussdekomposition aufgeführt. In den Fällen der teilintegrierten und integrierten Planung bleiben die Aufgabensequenzen bzw. Dienststücke gleich, nur der Puffer zwischen den einzelnen Dienststücken verändert sich gemäß der geplanten Verknüpfungen der Ressourcen.

Die Abbildungen 11.11 verdeutlichen den Grund für die Verringerung des Vernetzungseffektes und damit einhergehend für eine Verringerung der gesamten Verspätungspropagation. Erkennbar variiert bei Optimierung mit Berücksichtigen von (globaler) Verspätungspropagation die Anzahl von Dienststückwechseln mit wenig Pufferzeitraum im Vergleich zu Optimierung mit reiner Kostenminimierung als Zielsetzung. Die Pläne werden so erstellt, dass Dienststückwechsel mit wenig Pufferzeit

nach Möglichkeit vermieden werden. Darüber hinaus zeigen sich deutliche Unterschiede zwischen integrierter, teilintegrierter und sequenzieller Planungsmethode. Zum einen werden bei weniger Diensten (in Abhängigkeit der eingesetzten Diensttypen) auch weniger Wechsel eingeplant. Zum anderen treten deutlich weniger Wechsel mit geringen Pufferzeiten auf. Die Unterschiede erklären den in Abbildung 11.9 angegebenen abstufenden Vernetzungseffekt zwischen den einzelnen Planungsmethoden - von (robust) integrierter Planung mit einem geringen Vernetzungseffekt hin zu einem größerem Vernetzungseffekt bei sequenzieller Planung. Dabei ist zu beachten, dass im Allgemeinen eine steigende Anzahl Dienststückwechsel nicht zwangsläufig zu einer größeren Verspätungspropagation führen muss, wenn beispielsweise ausreichend Pufferzeit zur Verknüpfung oder in den Dienststücken selbst vorhanden ist.

Kernaussagen In den berechneten Lösungen ist der Verdichtungspropagationseffekt im Vergleich zur gesamten gemessenen Verspätungspropagation generell stärker ausgeprägt als ein Vernetzungspropagationseffekt. In hochgradig kosteneffizienten und robusten Plänen ist ein Vernetzungseffekt von Verspätungspropagation kaum vorhanden, insbesondere in integriert erstellten robusten Plänen.

Robuste Pläne können folglich so erstellt werden, dass Verspätungspropagation in gemeinsam von Umlauf- und Dienst auszuführenden Aufgabensequenzen isoliert und dadurch eine Weitergabe über Dienststücke hinaus verhindert wird. Entscheidend dafür ist zum einen eine geschickte Vernetzung der Dienststücke, für die bei der Planerstellung ausreichend Freiheitsgrade vorhanden sind. Zum anderen trägt der Einbau von passenden Pufferzeiten zum Abbau von Verspätungen innerhalb eines Dienststückes bei. Beide Aspekte können kosteneffizient durch die integrierte Planung umgesetzt werden.

Als Indikator der Robustheit von Umlauf- und Dienstplänen können zum einen eingeplante Dienststückwechsel in Kombination mit dafür vorhandenen Pufferzeiträumen dienen, da an diesen Stellen im Plan potenzielle Verspätungen zwischen den einzelnen Ressourcen sowie den Ressourcenebenen weitergegeben werden können. Zum anderen liefert die Aufschlüsselung der zwischen Aufgaben eines Dienststückes vorhandenen Pufferzeiten beim Vergleich unterschiedlicher Ressourceneinsatzpläne einen Hinweis auf die möglicherweise robuster gestalteten Lösungen.

11.3.2 Planstruktur-Indikatoren zu Verdichtung und Vernetzung

In dem vorangegangenen Abschnitt 11.3.1 wurde der aus den Planstrukturen der errechneten Ressourceneinsatzpläne resultierende Einfluss auf die mögliche Verspätungspropagation des gesamten Ressourceneinsatzes untersucht. Dabei wurde die Weiter-

gabe von Verspätungen der *Verdichtung* von Aufgabensequenzen und der *Vernetzung* von Ressourcen bzw. Dienststücken zugeordnet. In diesem Abschnitt werden nun die planimmanenten Ursachen für die zuvor beschriebenen Effekte formalisiert und zu zwei Kennzahlen aggregiert. Dabei wird die Weitergabe von Verspätungen im Zusammenhang mit der jeweils zur Verfügung stehenden Pufferzeit betrachtet. Wie zuvor gezeigt ist dieser Zusammenhang auf zwei Konstruktionsebenen auszumachen. Zum einen innerhalb der einzelnen Aufgabensequenzen, die gemeinsam für Dienst und Umlauf zu einem Dienststück eingeplant werden. Zum anderen zwischen den einzelnen Aufgabensequenzen, d.h. beim Wechsel von Diensten auf einem Umlauf bzw. bei Dienstwechsel über eingeplante Pausen zu anderen Umläufen (siehe auch vorangegangene Abbildung 11.8 zur Aufschlüsselung möglicher Propagationseffekte).

Untersuchungsfragen Die Formalisierung planimmanenter Ursachen für die beobachteten Verspätungseffekte basiert auf den Fragestellungen:

KENN 1.) Welche Kennzahl ermöglicht allein anhand der Planstrukturen eine Bewertung des Ressourceneinsatzes hinsichtlich möglicher Verspätungspropagation durch Verdichtung?

KENN 2.) Welche Kennzahl ermöglicht allein anhand der Planstrukturen eine Bewertung des Ressourceneinsatzes hinsichtlich möglicher Verspätungspropagation durch Vernetzung?

Im Folgenden werden die Kennzahlen als Aufgabenwechsel-Pufferzeit-Indikator (*AP-Indikator*) für die Beschreibung der Verdichtung des Ressourceneinsatzes und als Dienststückwechsel-Pufferzeit-Indikator (*DP-Indikator*) für die Beschreibung der Vernetzung des Ressourceneinsatzes hergeleitet und definiert.

AP-Indikator für Verdichtung (ad KENN 1.) Betrachtet wird die Aufgabensequenz eines Dienststückes und die zwischen den Aufgaben zur Verfügung stehenden Pufferzeiten, also die Wartezeiten zwischen Ausrück-, Einrück-, Service- und Leerfahrten, in Bezug zu den auszuführenden Servicefahrten eines Dienststückes. So wird abgebildet, inwiefern ausreichend Pufferzeiten zur Verfügung stehen, um propagierte Verspätungen innerhalb eines Dienststückes für Servicefahrten zu absorbieren. Hierfür werden die einzelnen Pufferzeiträume betrachtet und Intervallgruppen von verschiedenen Pufferzeitengrößen gebildet. Da im Busverkehr in diskreten – in der Regel minutengenauen – Zeitintervallen geplant wird, wird im Folgenden eine minutengenaue Schrittgröße zur Intervallbildung genutzt. Als Resultat kann der Anteil der Pufferzeiten einer bestimmten Größe an allen Pufferzeiten zwischen Aufgaben, den

Aufgabenwechselpufferzeiten, errechnet und mit anderen Lösungen verglichen werden. Da sich jedoch Dienststücke in Zahl der auszuführenden Servicefahrten (und ihrer Länge) unterscheiden, wird zusätzlich für die Vergleichbarkeit unterschiedlicher Lösungen der Anteilswert anhand der durchschnittlichen Anzahl von Servicefahrten pro Dienststück normiert. Das Ergebnis ist *der durchschnittlich pro Servicefahrt zur Verfügung stehende Anteil an Pufferzeiten einer bestimmten Größe gemessen an der Anzahl aller Pufferzeiten in einem Dienststück.*

Für einen geeigneten Vergleich ist bei der Berechnung die Wahl des Schwellenwertes für die zu betrachtenden Pufferzeiten ausschlaggebend. Ist er zu gering gewählt, ähneln sich die Kennzahlenwerte unterschiedlich robuster Lösungen stark. Ist er zu groß gewählt, werden im Vergleich weniger entscheidende, zur Reduktion potenzieller Verspätungspropagation übermäßig große Pufferzeiten hervorgehoben. Der Schwellenwert zur Betrachtung ist im Idealfall in Abhängigkeit der auftretenden Verspätungen aus (historischen) Verspätungsdaten bestimmbar oder in Variation zu betrachten, so dass aussagekräftige Gruppen von Vergleichskennzahlen entstehen.

Beispiel: Haben im Durchschnitt in einem Dienststück 40% aller Pufferzeiten zwischen Aufgaben eine Länge von mindestens zwei Minuten und es werden im Schnitt vier Servicefahrten ausgeführt, dann stehen pro Servicefahrt und zugehörig eingeplanten Aufgaben wie z.B. Leerfahrten im Mittel 10% der Pufferzeiten zur Verfügung, die größer gleich zwei Minuten sind.

Mathematische Formulierung: Sei G die Menge aller Intervallgruppen und sei $P(k)$ eine Funktion, die für den Dienst $k \in K$ die Menge aller in ihm enthaltenen Dienststücke liefert. Weiterhin sei $T^S(p)$ eine Funktion, die für Dienststück p die Menge aller bedienten Servicefahrten liefert und sei $W^g(p)$ eine Funktion, die die Menge aller im Dienststück p enthaltenen Pufferzeiten bestimmt, die der durch $g \in G$ definierten Intervallgruppe entsprechen. Dann ist der *AP-Indikator* der Dienste K für Intervallgruppe $g \in G$, für den Fall, dass $\sum_{k\in K}\sum_{p\in P(k)}(|T^S(p)|)/\sum_{k\in K}(|P(k)|) > 0$ gilt, definiert als

$$v(K,g) \quad = \quad \frac{\sum_{k\in K}\sum_{p\in P(k)}(|W^g(p)|/|W^0(p)|)}{\sum_{k\in K}\sum_{p\in P(k)}(|T^S(p)|)/\sum_{k\in K}(|P(k)|)} \tag{11.1}$$

und ansonsten 0, wobei $\sum_{k\in K}\sum_{p\in P(k)}(|W^g(p)|/|W^0(p)|)$ über alle Dienststücke den durchschnittlichen Anteil aller Pufferzeiten, die in Intervallgruppe $g \in G$ liegen, an der Anzahl aller Pufferzeiten beschreibt, und $\sum_{k\in K}\sum_{p\in P(k)}(|T^S(p)|)/\sum_{k\in K}(|P(k)|)$ die durchschnittliche Anzahl Servicefahrten pro Dienststück.

Kennzahlbewertung und potenzielle Verspätungspropagation gegenübergestellt: Abbildung 11.12 zeigt die Bewertung verschiedener durch sequenzielle,

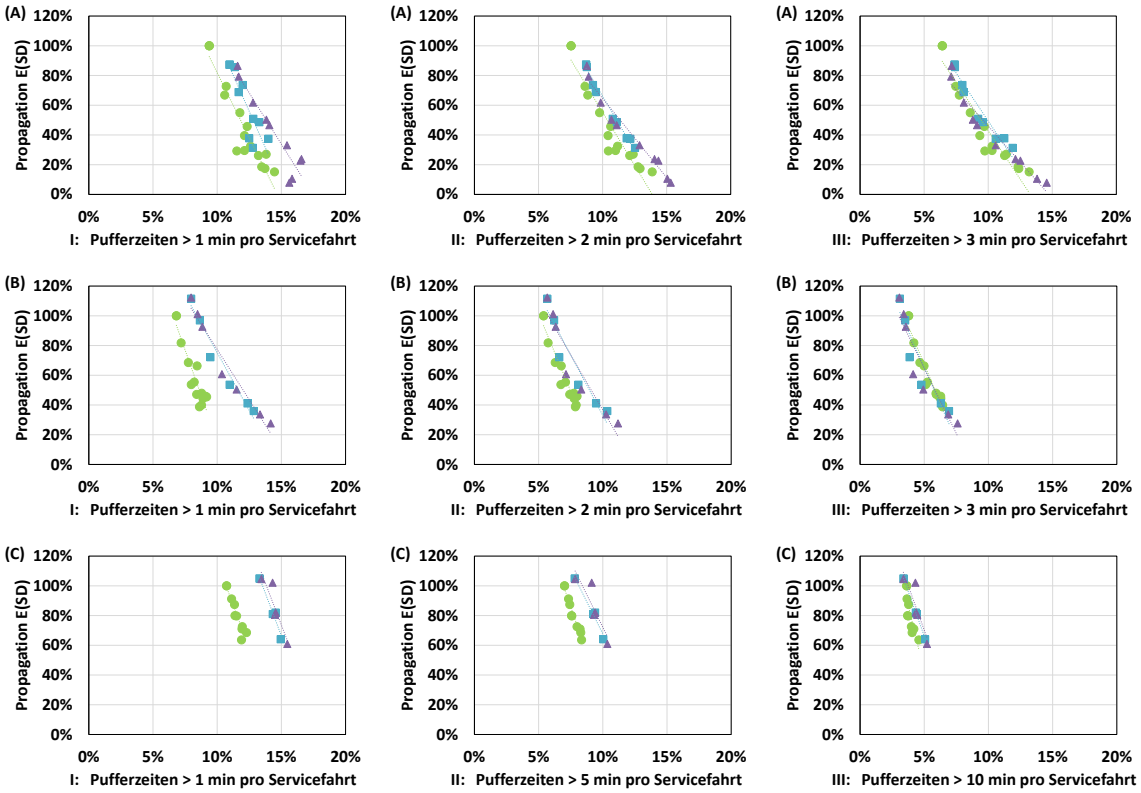

Abbildung 11.12: Kennzahl zur Verdichtung mit verschiedenen Schwellenwerten (Bilder I-III) gegenübergestellt der Verspätungspropagation auf Dienststückebene für sequenziell (*Dreiecke*), teilintegriert (*Vierecke*) und integriert (*Kreise*) erstellte Lösungen

teilintegrierte und integrierte Planung erstellter Lösungen mit dem *AP-Indikator* exemplarisch an den Testinstanzen A, B und C. Die aufgeführten Lösungen entsprechen denen im vorangegangenen Teilkapitel adressierten Lösungen (z.B. in Abbildung 11.6) und unterscheiden sich somit hinsichtlich Kosteneffizienz und Robustheit. Den Kennzahlenwerten ist jeweils gegenübergestellt die allein auf Dienststückebene gemessene Verspätungspropagation. Zur Vergleichbarkeit ist der Propagationswert in der Darstellung für jede Instanz auf die Verspätungspropagation innerhalb der Dienststücke des Basisfalls der integrierten Planung mit reiner Minimierung der Verspätungspropagation normiert.

In der Tendenz gilt, je größer der Anteil Puffer über einem konkreten Schwellenwert im Mittel pro Servicefahrt zur Verfügung steht, desto weniger werden Verspätungen im Dienststück weitergegeben. Dabei ist zu erkennen, dass für geringe Pufferzeiten nicht immmer eine klare Vergleichsaussage zwischen unterschiedlichen Lösungen getroffen werden kann. Mit Veränderung des Schwellenwertes zur Kennzahlerstellung wird deutlich, wie die erstellten Lösungen – auch bei unterschiedlichen Planungs-

methoden – gleichermaßen im Hinblick auf Verspätungspropagation innerhalb der Dienststückebene so bewertet werden können, dass im Vergleich der Lösungen die potenzielle Robustheit treffend gegeneinander abgewogen werden kann. Dabei werden überdies als Nebenaspekt der Kennzahlenbetrachtung Unterschiede im Umgang mit den auszuführenden Servicefahrten der einzelnen Instanzen und somit den zu Grunde liegenden Planungsproblemen selbst ersichtlich. Denn es werden erkennbar durch die im Fahrplan vorgegebenen Ankunfts- und Abfahrtszeiten der Servicefahrten unterschiedlich hohe Pufferzeiten pro Servicefahrt je Instanz erlaubt.

DP-Indikator für Vernetzung (ad KENN 2.) Für die Kennzahl werden zum einen die Dienststückwechsel auf Umlaufebene und die Dienstwechsel über eingeplante Pausen zu anderen Umläufen betrachtet. Zum anderen die jeweils für einen Wechsel zur Verfügung stehenden Pufferzeiten. So wird abgebildet, inwiefern Pufferzeiten für einzelne Dienststückwechsel bzw. Verknüpfungen von Aufgabensequenzen zur Verfügung stehen, um die zuvor innerhalb eines Dienststückes bzw. einer Aufgabensequenz entstandenen und bis zum Ende nicht abgebauten Verspätungen absorbieren zu können. Wie zuvor werden hierfür die einzelnen Pufferzeiten betrachtet und Intervallgruppen von verschiedenen Pufferzeitengrößen gebildet. Dabei beträgt die Schrittgröße zur Intervallbildung jeweils eine Minute. Als Ergebnis wird *der Anteil an Pufferzeiten über einer bestimmten Größe gemessen an der Anzahl aller Pufferzeiten für Dienststückwechsel* berechnet und kann mit anderen Lösungen verglichen werden. Die im Vergleich verschiedener Lösungen unterschiedlichen Ausprägungen der einem Wechsel vorangegangenen Aufgabensequenzen und somit auch auszuführenden Servicefahrten werden dabei durch den Bezug zu den insgesamt eingeplanten Dienststückwechseln berücksichtigt: Je weniger Servicefahrten im Durchschnitt gemeinsam von Umlauf und Dienst pro Dienststück auszuführen sind, desto mehr Dienststücke werden eingeplant, desto mehr Dienststückwechsel sind zu berücksichtigen.

Beispiel: Nach Ausführung einer gemeinsamen für einen Umlauf- und Dienst eingeplanten Aufgabensequenz bieten 80% aller Dienst(stück)wechsel einen Pufferzeitraum von mindestens 10 Minuten bevor andere Aufgabensequenzen auszuführen sind.

Mathematische Formulierung: Sei G die Menge aller Intervallgruppen, sei $D^g(k)$ eine Funktion, die für Dienst $k \in K$ die Menge aller Pufferzeiten zwischen allen aufeinanderfolgenden Dienststücken liefert, die der Intervallgruppe $g \in G$ entsprechen und auf unterschiedlichen Fahrzeugen durchzuführen sind (– für zwei aufeinanderfolgende Dienststücke p_1 und p_2 gilt $L(p_1) \cap L(p_2) = \emptyset$). Sei $D^g(l)$ eine entsprechende Funktion, die für Fahrzeugumlauf $l \in L$ die Menge aller Wartezeiten zwischen allen aufeinanderfolgenden Dienststücken liefert, die der Intervallgruppe $g \in G$ entspre-

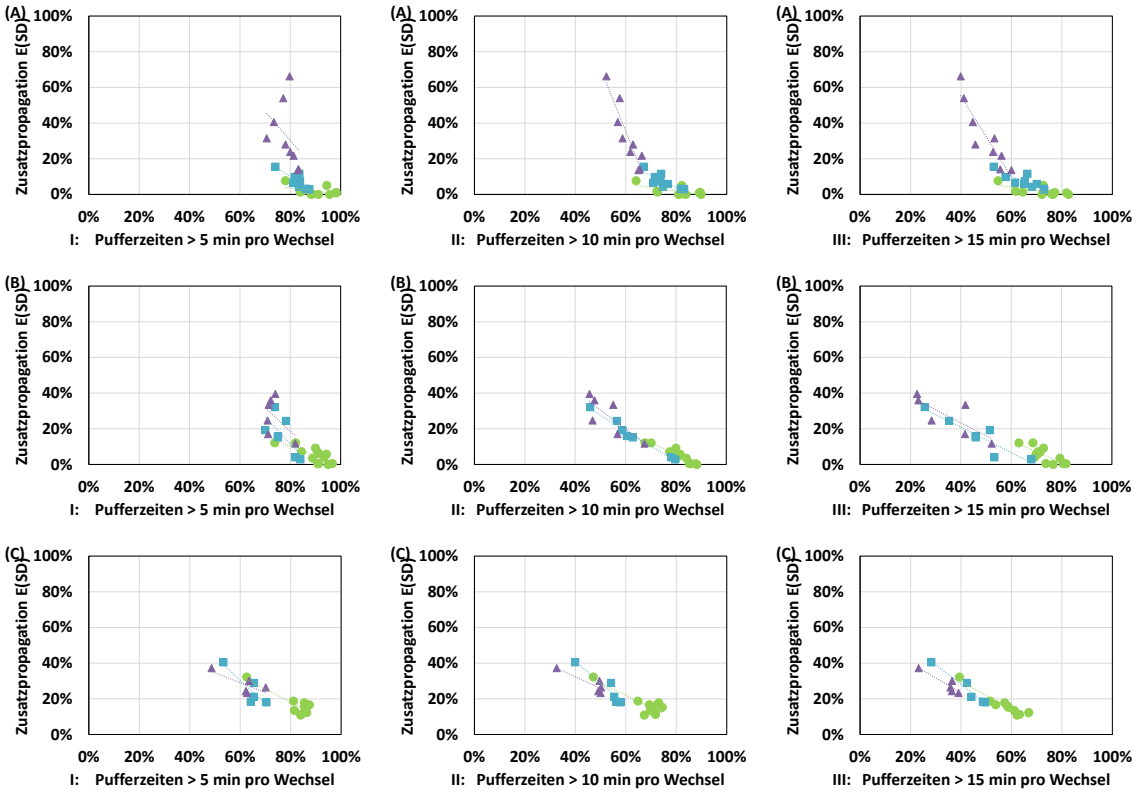

Abbildung 11.13: Kennzahl zur Vernetzung mit verschiedenen Schwellenwerten (Bilder I-III) gegenübergestellt der entstehenden zusätzlichen Verspätungspropagation des kombinierten Ressourceneinsatzes für sequenziell (*Dreiecke*), teilintegriert (*Vierecke*) und integriert (*Kreise*) erstellte Lösungen

chen. Dann ist der *DP-Indikator* für Dienste K, Fahrzeugumläufe L und Intervallgruppen G für den Fall, dass mindestens ein Dienststückwechsel stattfindet, definiert als

$$w(K, L, g) \quad = \quad \frac{\sum_{k \in K} |D^g(k)| + \sum_{l \in L} |D^g(l)|}{\sum_{k \in K} |D^0(k)| + \sum_{l \in L} |D^0(l)|} \tag{11.2}$$

und ansonsten 0, wobei $\sum_{k \in K} |D^g(k)| + \sum_{l \in L} |D^g(l)|$ die Summe aller Dienststückwechsel beschreibt, mit Pufferzeiten, die der Intervallgruppe $g \in G$ entsprechen und $\sum_{k \in K} |D^0(k)| + \sum_{l \in L} |D^0(l)|$ die Summe aller Dienststückwechsel angibt.

Kennzahlbewertung und potenzielle Verspätungspropagation gegenübergestellt: Abbildung 11.13 zeigt die Bewertung einzelner durch sequenzielle, teilintegrierte und integrierte Planung erstellter Lösungen mit dem *DP-Indikator* exemplarisch für Testinstanzen A, B und C. Die dargestellten Lösungen entsprechen denen im vorangegangenen Teilkapitel aufgeführten Lösungen (z.B. in Abbildung 11.6). Sie un-

terscheiden sich somit hinsichtlich Kosteneffizienz und Robustheit. Gegenübergestellt ist jeweils der bei kombinierter Betrachtung von Umläufen und Diensten auftretende Zusatzeffekt an Verspätungspropagation im Vergleich zu allein auf Dienststückebene gemessener Verspätungspropagation. Die Kennzahl ist aufgeführt für den Anteil an Pufferzeiträumen die größer als 5, 10 und 15 Minuten sind.

Erkennbar liefert die Betrachtung des Anteils an Pufferzeiten unter Einbeziehen auch geringer Zeiträume in den Beispielen keine klare Vergleichsaussage über potenziell auftretenden Verspätungungspropagation. Erst durch die Betrachtung größerer Pufferzeiträume werden eindeutigere Aussagen möglich. In der Tendenz können geringere Zusatzpropagationseffekte einem höheren Anteil größerer Pufferzeiten zugeschrieben werden.

11.3.3 Zusammenfassung der Untersuchungen

In diesem Teilkapitel wurden mit der *Verdichtung* und *Vernetzung* zwei wesentliche Eigenschaften zur Weitergabe von Verspätungen im kombinierten Ressourceneinsatz unterschiedlicher Ressourcentypen näher bestimmt und untersucht. Durch Variation des Umfanges in der Simulation des Fahrplanes in Kombination mit den auszuführenden Aufgabensequenzen auf Umlauf- und Dienstebene wurden dabei die potenziell entstehenden verschiedenen Effektgrößen von Verspätungspropagation im Zusammenhang mit den erstellten Ressourceneinsatzplänen aufgeschlüsselt. Die Weitergabe von Verspätungen wurde zum einen innerhalb von Dienststücken betrachtet. Zum anderen zwischen Dienststücken verschiedener Dienste auf Umlaufebene sowie über die Verknüpfung von Dienststücken durch eingeplante Fahrerpausen (auf andere Umläufe). Zur Analyse der planimmanenten Ursachen für die einzelnen Verspätungseffekte wurde die jeweils verfügbare Pufferzeit in Bezug zu den einzelnen Dienststückwechseln und Aufgabenwechseln gebracht.

(ad F1.) Es konnten unterschiedliche Ausprägungen der einzelnen Propagationseffekte bei Variation der Planungsmethoden festgestellt werden. Diese möglichen Unterschiede bei der Zusammensetzung der gesamten Verspätungspropagation des Ressourceneinsatzes erklären die im vorangegangenen Teilkapitel 11.2 herausgearbeitete Dominanz der verschiedenen Planungsmethoden untereinander. Im Vergleich zwischen (robuster) teilintegrierter mit sequenzieller Planung kann gemäß der vorgeschlagenen teilintegrierten Planungsmethodik besonders der potenzielle Vernetzungseffekt verringert werden. Im Vergleich zwischen (robuster) integrierter Planung mit teilintegrierter und sequenzieller Planung kann gemäß der vorgeschlagenen integrierten Planungsmethodik sowohl der Vernetzungs- als auch der Verdichtungspropagationseffekt verringert werden.

(ad F3.) Im Allgemeinen wiegt die Verdichtung von Aufgabensequenzen in Dienststücken stärker bei der Weitergabe von Verspätungen als die Vernetzung der Dienststücke. Wesentlich für die Robustheit des Ressourceneinsatzes ist daher eine passende Verteilung eingeplanter Pufferzeiten zwischen den einzelnen Aufgaben bzw. Fahrten innerhalb der Dienststücke. Eine solche passende Verteilung von Pufferzeiten wird im folgenden Teilkapitel 11.4 in Abschnitt 11.4.1 vertiefend untersucht. Für die gesamte Robustheit ist allerdings das Zusammenspiel mit einer geschickten Vernetzung der Ressourcen bzw. Dienststücke entscheidend. Auch in kosteneffizienten Plänen mit (deutlicher) Verspätungspropagation durch Verdichtung ist es durch die passende Vernetzung möglich, noch ausreichend Pufferzeiten zwischen Dienststücken bei Fahrer- bzw. Fahrzeugwechseln so einzupflegen, dass kein bzw. nur geringer zusätzlicher Vernetzungseffekt aufgebaut wird. Im Vergleich zu den Pufferzeitengrößen innerhalb der Dienststücke sind in der Regel deutlich längere Pufferzeiten zwischen den Dienststücken möglich, die somit bei passender Verteilung zur Verspätungsabsorption ausgenutzt werden können.

Mit dem *AP-Indikator* sowie dem *DP-Indikator* wurden darüber hinaus zwei Kennzahlen herausgearbeitet, die die Planstrukturen des kombinierten Ressourceneinsatzes vor dem Hintergrund von Verdichtungs- und Vernetzungseffekten abbilden. Anhand der Indikatoren lassen sich Ressourceneinsatzpläne (unabhängig der eingesetzten Planungsmethode) im Hinblick auf Robustheit bewerten. So kann unterschiedlich geplanter Ressourceneinsatz gemäß potenziell enstehender Verspätungspropagation auf Dienststückebene und potenziell zusätzlich entstehender Propagation durch Vernetzung der einzelnen Ressourcen miteinander verglichen und die Güte der Lösungen bezüglich Robuster Effizienz in ihrer Gesamtheit grob abgeschätzt werden.

11.4 Einfluss der Ausweitung der betrachteten Planungsaufgabe

In den vorangegangenen Teilkapiteln wurde ein bestehender Fahrplan als unveränderliche Eingangsgröße der Ressourceneinsatzplanung angesehen. In diesem Teilkapitel wird eine Anpassung der für die Ressourceneinsatzplanung zu Grunde liegenden Rahmenvorgaben des Fahrplans diskutiert. In Abstimmung der Fahr-, Umlauf- und Dienstplanungsebene werden durch ein eng auf den Ressourceneinsatz zugeschnittenes zeitliches Verschieben von Servicefahrten im Fahrplan Potenziale zur Steigerung Robuster Effizienz untersucht.

Mit der Ausweitung der betrachteten Planungsaufgaben und der damit einhergehenden Änderung der Problembeschaffenheit können insbesondere die Erkenntnisse zu den Forschungsfragen F1 und F3 weiter vertieft werden. Dabei wird vor allem der zusätzliche Spielraum zur Erhöhung der Robustheit (und Kosteneffizienz) durch geringfügige Fahrplananpassungen untersucht.

Die Untersuchung wird in zwei Experimentegruppen eingeteilt. In Abschnitt 11.4.1 wird basierend auf dem erstellten Ressourceneinsatz ein Verschieben der Fahrplanfahrten erlaubt, um durch Vorgaben aus dem Fahrplan zusätzliche Robustheit bei gleichzeitiger Beibehaltung der Kosten des Ressourceneinsatzes generieren zu können. Abschnitt 11.4.2 beschäftigt sich mit der gleichzeitigen Betrachtung von Fahrplananpassung und integrierter Planung zur Erhöhung von sowohl Robustheit als auch Effizienz des Ressourceneinsatzes. Abschnitt 11.4.3 fasst dieses Teilkapitel zusammen.

11.4.1 Fahrplananpassung basierend auf geplantem Ressourceneinsatz

In diesem Abschnitt werden die durch die unterschiedlichen Planungsmethoden errechneten Ressourceneinsatzpläne benutzt, um basierend auf den vorhandenen Aufgabensequenzen und Ressourcenverknüpfungen den zu Grunde liegenden Fahrplan so anzupassen, dass bei gleichem Ressourceneinsatz eine robustere Ausführung erreicht werden kann.

Untersuchungsfragen Die Experimente werden auf Grundlage der folgenden Fragen durchgeführt:

RET 1.) Welcher Spielraum zur Verringerung der Verspätungspropagation ist durch Retiming gegeben, ohne Robustheit durch zusätzliche Pufferzeiten zu erkaufen?

RET 2.) In welchem Ausmaß unterscheidet sich das Potenzial von Fahrtenverschiebungen in Abhängigkeit der ursprünglich eingesetzten Planungsmethoden?

RET 3.) Welche Eigenschaften im Hinblick auf die Verteilung von Pufferzeiten weisen die in Dienststücken verdichteten Aufgabensequenzen aus?

RET 4.) Welche Eigenschaften im Hinblick auf die Verteilung von Pufferzeiten weisen die angepassten Fahrpläne auf?

Genutzte Artefakte Für die Experimente wird das Retiming-Verfahren aus Abschnitt 9.2 genutzt. Im Vergleich der möglichen, unterschiedlichen Varianten – Aggregation von Verspätungsszenarien durch verschiedene Schätzer oder Nutzen einer

(aus den Szenarien abgeleiteten) Verspätungsverteilung – hat sich die Variante des Nutzens einer Verspätungsverteilung als am geeignetsten herausgestellt. Da durch diese am genauesten potenzielle Verspätungen abgebildet werden, können so die zur Verringerung von Verspätungspropagation bestmöglichen Planveränderungen dokumentiert werden. Eingabe für das Retiming-Verfahren bilden jeweils die sequenziell, teilintegriert und integriert erstellten Ressourceneinsatzpläne, die auch in Abschnitt 11.2.3 gemeinsam diskutiert wurden. Die Laufzeiten für das Retiming-Verfahren betragen jeweils nur wenige Sekunden. Die durchgeführte Simulation der resultierenden Ressourceneinsatzpläne fokussiert allein auf eine Messung der Stabilität der Pläne durch Weitergabe von Verspätungen ohne eingreifende Dispositionsmaßnahmen. Dabei wird die Fahrtenausführung der zu den angepassten Ressourceneinsatzplänen korrespondierend veränderten Fahrpläne simuliert.

Ergebnisdarstellung Tabelle 11.6 stellt exemplarische Ergebnisse für Retiming des Fahrplans im Vergleich mit ursprünglichen Lösungen dar. Dabei werden zum einen unter *approach* die unterschiedlichen Planungsmethoden angegeben, mit denen die ursprünglichen Lösungen berechnet wurden. Die Methoden werden nach dem eingesetzten Planungsschema (*seq*, *pInt*, *Int*) sowie dem Umfang der Berücksichtigung von (globaler) Verspätungspropagation unterschieden (keine, *GFD*, *R*; siehe auch Beschreibung von z.B. Tabelle 11.3). Zum anderen kennzeichnet *retiming* den Einsatz des Retiming-Verfahrens. Zusätzlich werden das erlaubte Verschiebeintervall (TW), die Anzahl der verschobenen Fahrten (*shifted # services*), sowie in Prozent im Vergleich zu den Gesamtfahrten, die verschobene Zeit in Minuten (*shifted time*) und die durchschnittliche verschobene Zeit pro verschobener Servicefahrt (*shifted t per serv*) aufgeführt. Darüber hinaus werden angegeben die Reduktion der Verspätungspropagation im Vergleich zur Ausgangslösung ($\Delta\%$ *E(SD)*), sowie die Kosten (*costs in % to int*) und Verspätungspropagation (*E(SD) in % to int*) im Vergleich zum Basisfall der integrierten, rein kostenminimierenden Ressourceneinsatzplanung. Die jeweils für Verschiebungen der Servicefahrten betrachtete Ausgangslösung wird durch den Eintrag von „x“ in den Spalten zu Fahrplananpassungen markiert.

Ergebnisinterpretation (ad RET 1.) Die Tabelle 11.6 verdeutlicht das vorhandene Potenzial der nachträglichen Fahrplanpassung zur Reduktion der potenziellen Verspätungspropagation ohne die Aufgaben bzw. Kostenstruktur des Ressourceneinsatzes zu verändern. Alle erstellten Lösungen können durch Verschieben von Fahrplanfahrten unabhängig von der genutzten Planungsmethodik deutlich robuster gestaltet werden. Die Verschiebeintervalle sind dabei gering, das Ausmaß der Verschiebung

Tabelle 11.6: Einfluss von nachträglichem Retiming auf die Robustheit von mit unterschiedlichen Planungsmethoden erstellten Ressourceneinsatzplänen

approach	TW±	#services	shifted time	t per serv	Δ% E(SD)	costs in % to int	E(SD) in % to int
				Testinstanz A			
int	x	x	x	x	x	100,0	100,0
retiming	2	109 (49,3%)	201	1,84	-24,7	100,0	75,3
	4	109 (49,3%)	339	3,11	-34,1	100,0	65,9
	6	99 (64,0%)	389	3,93	-36,0	100,0	64,0
intGFD	x	x	x	x	x	100,0	94,0
retiming	2	106 (48,0%)	194	1,83	-24,8	100,0	70,7
	4	106 (48,0%)	329	3,10	-34,0	100,0	62,0
	6	107 (48,4%)	440	4,11	-35,9	100,0	60,2
intR	x	x	x	x	x	104,1	26,2
retiming	2	97 (43,9%)	179	1,85	-27,9	104,1	18,9
	4	97 (43,9%)	307	3,16	-39,8	104,1	15,8
	6	101 (45,7%)	385	3,81	-45,5	104,1	14,3
				Testinstanz B			
int	x	x	x	x	x	100,0	100,0
retiming	2	183 (54,0%)	301	1,64	-15,6	100,0	84,4
	4	200 (59,0%)	475	2,38	-20,9	100,0	79,1
	6	197 (58,1%)	581	2,95	-22,0	100,0	78,0
intGFD	x	x	x	x	x	100,0	94,1
retiming	2	184 (54,3%)	301	1,64	-12,6	100,0	82,2
	4	201 (59,3%)	478	2,38	-18,2	100,0	76,9
	6	196 (57,8%)	579	2,95	-19,6	100,0	75,7
intR	x	x	x	x	x	104,8	41,2
retiming	2	210 (61,9%)	377	1,80	-15,3	104,8	34,9
	4	205 (60,5%)	556	2,71	-23,4	104,8	31,6
	6	206 (60,8%)	726	3,52	-27,0	104,8	30,1
pInt	x	x	x	x	x	104,1	128,0
retiming	2	160 (47,2%)	251	1,57	-6,4	104,1	119,8
	4	162 (47,8%)	346	2,14	-10,4	104,1	114,6
	6	164 (48,4%)	419	2,55	-12,9	104,1	111,4
pIntGFD	x	x	x	x	x	104,1	106,5
retiming	2	158 (46,6%)	246	1,56	-8,3	104,1	97,6
	4	159 (46,9%)	341	2,14	-13,1	104,1	92,5
	6	161 (47,5%)	416	2,58	-15,3	104,1	90,1
pIntR	x	x	x	x	x	116,8	59,2
retiming	2	103 (30,4%)	174	1,69	-4,3	116,8	56,7
	4	103 (30,4%)	246	2,39	-4,2	116,8	56,7
	6	102 (30,1%)	281	2,75	-3,8	116,8	57,0
seq	x	x	x	x	x	111,9	141,0
retiming	2	154 (45,4%)	243	1,58	-8,0	111,9	129,6
	4	168 (49,6%)	351	2,09	-10,0	111,9	126,8
	6	168 (49,6%)	412	2,45	-11,0	111,9	125,4
seqGFD	x	x	x	x	x	111,9	128,9
retiming	2	159 (46,9%)	243	1,58	-8,2	111,9	118,3
	4	162 (47,8%)	343	2,12	-12,9	111,9	112,2
	6	163 (48,1%)	414	2,54	-15,5	111,9	108,9
seqR	x	x	x	x	x	123,0	62,1
retiming	2	105 (31,0%)	172	1,64	-4,5	123,0	59,3
	4	104 (30,7%)	247	2,38	-4,1	123,0	59,6
	6	105 (31,0%)	285	2,71	-3,8	123,0	59,8
				Testinstanz C			
int	x	x	x	x	x	100,0	100,0
retiming	2	389 (58,9%)	713	1,83	-18,9	100,0	81,1
	4	413 (62,5%)	1254	3,04	-24,0	100,0	76,0
	6	417 (63,1%)	1621	3,89	-29,1	100,0	70,9
intGFD	x	x	x	x	x	100,0	77,4
retiming	2	385 (58,2%)	708	1,84	-18,3	100,0	63,2
	4	394 (59,6%)	1209	3,07	-25,3	100,0	57,8
	6	405 (61,3%)	1571	3,88	-29,6	100,0	54,5
intR	x	x	x	x	x	103,4	43,5
retiming	2	387 (58,5%)	698	1,80	-19,5	103,4	35,0
	4	384 (58,1%)	1174	3,06	-27,8	103,4	31,4
	6	386 (58,4%)	1541	3,99	-27,6	103,4	31,5

von Fahrplanfahrten ist jedoch stark ausgeprägt. So werden häufig gut die Hälfte der vorhandenen Fahrten verschoben. Allerdings findet die einzelne Verschiebung zum einen nur in geringem Ausmaß statt (bis maximal sechs Minuten zum Aufzeigen von Potenzialen). Zum anderen sollen die Experimente – auch als Vergleichswert für die in Abschnitt 11.4.2 folgenden Untersuchungen – die größtmöglichen Potenziale aufzeigen ohne z.B. die starre Beibehaltung möglicher Taktungen im Fahrplan berücksichtigen zu müssen.

Je größer das Verschiebeintervall gewählt wird, desto größer ist erwartungsgemäß die Verringerung der potenziellen Verspätungspropagation bzw. die zusätzlich erreichbare Robustheit. Es zeigt sich allerdings auch, dass aufgrund der Beschränkung zur Beibehaltung der Kostenstruktur und somit zur Beibehaltung der Aufgabendauern bereits ein Verschiebeintervall von sechs Minuten zu keiner bzw. kaum zur weiteren Verbesserung der Robustheit führt. Dies wird beim Vergleich mit dem Verschiebeintervall von vier Minuten beispielsweise für Testinstanz B für robust integriert, teilintegriert und sequenziell erstellte Pläne deutlich, die bereits über ausreichend eingeplante Pufferzeiten verfügen.

(ad RET 2.) Aufgrund der zuvor ausgewiesenen längeren Aufgabensequenzen in den Planstrukturen integriert erstellter Plänen (vgl. Abschnitt 11.2.3) besitzen diese Ressourceneinsatzpläne im Vergleich zu sequenziell oder teilintegriert erstellten Pläne größeres Potenzial für eine Reduktion der Verspätungspropagation (siehe Spalte $\Delta\%$ *E(SD)* der Tabelle 11.6). Es können mehr Servicefahrten angepasst werden ohne die Kostenstruktur durch Verlängerung von Dienststücken bzw. der gemeinsam von einem Umlauf- und Dienst auszuführenden Aufgaben zu ändern. Dafür ist deutlich zu erkennen, dass bei integriert erstellten Plänen innerhalb der Dienststücke tatsächlich auch mehr Servicefahrten verschoben werden. Bei robust erstellten teilintegrierten bzw. sequenziellen Plänen sind weniger Verschiebungen möglich, so dass hier das Potenzial zur Verbesserung bereits robust erstellter Pläne geringer ausgeprägt ist.

Planstrukturen Eine Untersuchung der entstehenden Planstrukturen erklärt das gezeigte Potenzial zur Verspätungsreduktion. Dabei werden im Folgenden sowohl die Ebene des Ressourceneinsatzes als auch die Ebene des zugrunde liegenden Fahrplans adressiert.

(ad RET 3.) Abbildung 11.14 verdeutlicht exemplarisch für Lösungen der Testinstanz B die Veränderung von Pufferzeiten in den Dienststücken sowie die daraus resultierende Entzerrung der Aufgabensequenzen zur Ausführung von Servicefahrten. Betrachtet werden mit (robust) integrierter Planungsverfahren erstellte Lösungen (Bilder I) und die darauf basierenden, resultierenden Ressourceneinsatz- und Fahr-

pläne bei Berücksichtigen von Verschiebeintervallen von jeweils zwei, vier und sechs Minuten (Bilder II-IV). Die Bilder V aggregieren die Betrachtung indem als Vergleich für die jeweilige Ausgangslösung und den sich durch Retiming ergebenden Lösungen die in Abschnitt 11.3.2 vorgeschlagene Kennzahl AP-Indikator zur Verdichtung der Aufgaben in Dienststücken aufgeführt wird.

Der AP-Indikator mit verschiedenen Schwellenwerten (hier von eins bis vier Minuten) zeigt, dass sich bei gleichbleibender Gesamtpufferzeit innerhalb der Dienststücke der Anteil an Pufferzeiten über einer bestimmten Größe für die Ausführung pro Servicefahrt verschiebt (insbesondere der Anteil bei über eins bis drei Minuten). Beeinflusst wird hierbei allein die Verteilung der Pufferzeiten innerhalb von Dienststücken. Diese Veränderung ist an den Bildern I bis IV zu erkennen. Da sich die Lösungen der ersten beiden Spalten allein in der Verknüpfung der Dienststücke unterscheiden, verändert sich die Verteilung ähnlich verglichen mit der Ausgangslösung. Die robustere Ausgangslösung in der dritten Spalte mit geringeren Aufgaben pro Dienststück weist größere Pufferzeiten auf. Die Veränderungen über die jeweiligen Stufen des Retimings spiegeln überdies eine gemäß der vorhandenen Freiheitsgrade leichte Anpassung der Pufferzeiträume an die zugrunde liegende potenzielle Verteilungsfunktion von Verspätungen der betrachteten Servicefahrten wieder. Darüber hinaus werden beim Retiming mehr Pufferzeiträume im Vergleich zu den Ausgangslösungen geschaffen. Die Anzahlen bewegen sich je nach Größe des Verschiebeintervalls zwischen ca. 5% bis 7%. Als Konsequenz werden die Anzahlen vormals großer Pufferzeiträume verringert.

(ad RET 4.) Zusätzlich zur Sicht auf den Ressourceneinsatzplan stellt Abbildung 11.15 die Auswirkungen der Veränderungen in der Ganglinie der Servicefahrten dar. Aufgeführt werden exemplarisch die zur reinen kostenminimalen integrierten Planung aufgabengleiche Lösung der Testinstanz B, die durch Berücksichtigen von Verspätungspropagation ermittelt wurde, sowie die durch Retiming erfolgten Anpassungen der Ganglinie und der Verteilung von Pufferzeiten bei Verschiebeintervallen von zwei, vier und sechs Minuten. Für jede dieser Lösungen ist die Auslastung mit Servicefahrten inklusive eingeplanter gemeinsamer Pufferzeit von korrespondierenden Umlauf(-aufgaben) und Dienst(-aufgaben) pro Uhrzeit dargestellt. Da bei der nachträglichen Fahrplananpassung als Zielfunktion die Gesamtverspätungspropagation minimiert wird, weisen die entstehenden Veränderungen auf für die Gesamtheit des einzuplanenden Ressourceneinsatz in Kombination mit den auszuführenden Servicefahrten dienliche Planungsempfehlungen hin. Im Einzelnen wird eine Umverteilung von Pufferzeit durch Anpassen der Servicefahrtzeiten in unterschiedlichem Umfang ersichtlich. So zeigen sich insbesondere um die mit höherem Ressourceneinsatz aus-

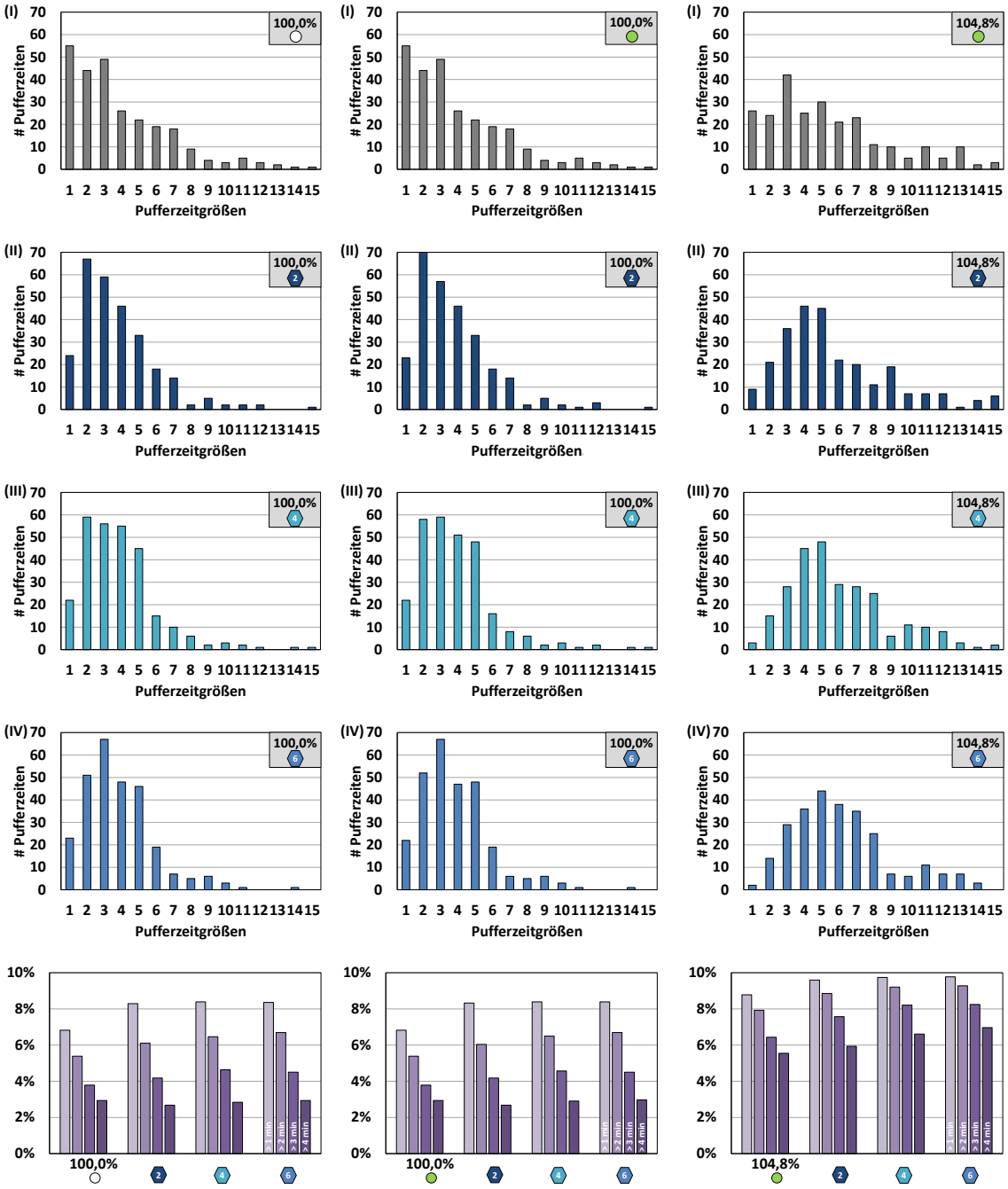

Abbildung 11.14: Häufigkeiten der Pufferzeiten zwischen Aufgabenwechseln in allen Dienststücken des geplanten Ressourceneinsatzes am Beispiel (robust) integriert erstellter Lösungen für Testinstanz B (Bilder I-IV), sowie zugehöriger AP-Indikator

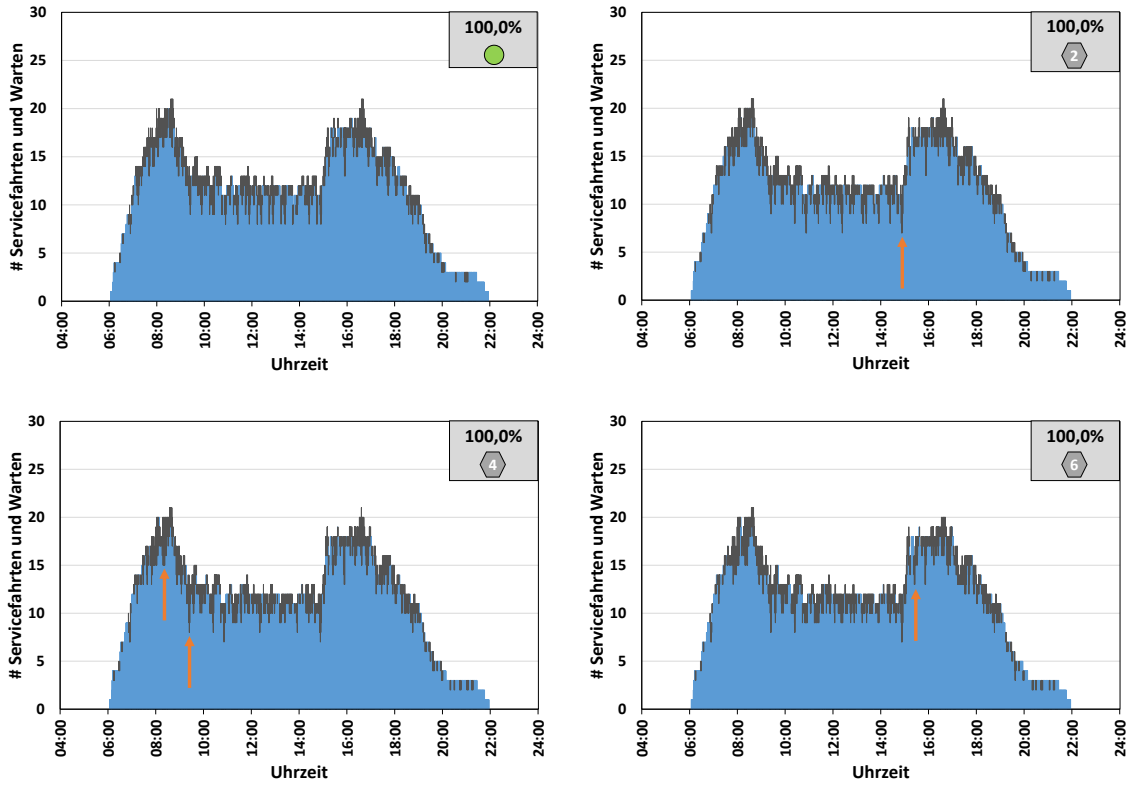

Abbildung 11.15: Anpassungen der Ganglinie (*hellblau*) und Verteilung von Pufferzeiten (*grau*) durch Retiming am Beispiel einer kostenminimalen (robust dekomponierten) Umlauf- und Dienstplanlösung für Testinstanz B, inklusive markierter wesentlicher Änderungen (*orangefarbene Pfeile*)

geführten Auslastungsspitzen deutliche Anpassungen, die in den Bildern mit Pfeilen markiert sind. Als Effekte werden zum einen tiefere, etwas breiter als usprünglich mit Pufferzeit gefüllte Lücken nach der ersten Auslastungsspitze (um ca. 9.30 Uhr) und vor der zweiten Auslastungsspitze (um ca. 15.30 Uhr) generiert. Pufferzeiten werden also anschaulich so verschoben, dass in größerem Umfang Propagation aus der ersten Auslastungsspitze absorbiert bzw. möglichst wenig Verspätung in die zweite Spitze hinein genommen wird. Zum anderen zeigt sich ein leichtes Abflachen der beiden Spitzen. Dabei werden die einzelnen Spitzen so geteilt, dass zur Reduktion von Verspätungspropagation (mehrere) gleich hohe Spitzen mit deutlicher Pufferzeit dazwischen geschaffen werden.

Kernaussagen Durch das Anpassen der Lage der Servicefahrten im Fahrplan mit der Hilfe von Verschiebeintervallen ist eine Reduktion potenzieller Verspätungspropagation bei exakt gleichbleibenden Ressourceneinsatzkosten möglich. Dabei werden

keine Einsatzdauern verändert, sondern lediglich vorhandene Pufferzeiten umverteilt, wodurch das Potenzial einer solchen Vorgehensweise begrenzt ist. Es zeigt sich dennoch, dass die Verschiebungen der Servicefahrten so gestaltet werden können, dass in der Tendenz mehr Pufferzeiträume zwischen der Ausführung von Servicefahrten vorhanden sind. Durch die in Bezug zum Ressourceneinsatz und möglichen Verspätungen passgenauere Ausgestaltung der Pufferzeiträume werden einzelne Verspätungen nicht weiter aufgebaut und können z.T. absorbiert werden. Die Ergebnisse bieten eine obere Schranke für das erreichbare Robustheit-/Stabilitätspotenzial des jeweils geplanten Ressourceneinsatzes.

11.4.2 Integrierte Planung mit Zeitfenstern im Vergleich

Der vorangegangene Abschnitt 11.4.1 hat das Potenzial des nachträglichen Verschiebens von Servicefahrten zur Erhöhung der Robustheit des Ressourceneinsatzes aufgezeigt. Dabei erfolgte die Veränderung des Fahrplans unter strikter Beibehaltung der Kosten und vorhandenen Aufgabensequenzen. Daran anknüpfend wird in diesem Abschnitt herausgearbeitet, welchen Einfluss auf Robustheit und Effizienz eine simultan durchgeführte Ressourceneinsatzplanung und Veränderung der Fahrplanstruktur haben können. Durch eine solche Planung sind sowohl mehr Freiheitsgrade auf Ebene der Fahrplananpassung und als auch gleichzeitig mehr Freiheitsgrade bei der Ressourceneinsatzplanung vorhanden.

Untersuchungsfragen Folgende Fragestellungen bilden die Grundlage für die durchgeführten Untersuchungen:

ITW 1.) Welches Potenzial bezüglich Robuster Effizienz besteht durch integrierte Planung mit Zeitfenstern gegenüber sequenzieller, teilintegrierter und integrierter Ressourceneinsatzplanung?

ITW 2.) Welches Potenzial besteht darüber hinaus gegenüber dem Retiming von sequenziell, teilintegriert und integriert erstellten Plänen?

ITW 3.) Inwiefern sind Mengen von Servicefahrten im Fahrplan identifizierbar, deren geringe Verschiebung sowohl Robustheit als auch Kosteneffizienz des Ressourceneinsatzes erhöhen können?

ITW 4.) Welche Ausprägungen weist der durch (robust) integrierte Planung mit Zeitfenstern erstellte Ressourceneinsatz im Hinblick auf die Verdichtung von Aufgabensequenzen (in Dienststücken) und die Vernetzung von Dienststücken auf?

Genutzte Artefakte Zur Planung wird der in Abschnitt 9.4 vorgeschlagene Lösungsansatz zur integrierten robusten Planung mit Zeitfenstern eingesetzt. Die Lösungen werden insbesondere verglichen mit Lösungen, die durch (robuste) integrierte Planung erstellt wurden, sowie mit Lösungen, die durch zusätzliches Retiming auf robust integriert erstellten Pläne generiert wurden. Denn im Abschnitt zuvor wurde gezeigt, dass aufgrund der Aufgabenverkettungen im Vergleich zu teilintegrierter bzw. sequenzieller Planung bei integriert erstellten Plänen das größte Potenzial für Retiming vorhanden ist. Aufgrund der Problemkomplexität werden die für integrierte Planung mit Zeitfenstern berechneten Ergebnisse hauptsächlich durch Einsatz von Kantengenerierung im Lösungsframework gemäß Abschnitt 9.4.3 ermittelt. Mögliche Unterschiede der Lösungsqualität im Vergleich zur Betrachtung des kompletten Lösungsraumes werden in Tabelle 9.2 aufgezeigt und im genannten Abschnitt diskutiert.

Ergebnisdarstellung Die Abbildungen 11.16 stellen die Ergebnisse der mit unterschiedlichen Planungsmethoden berechneten Umlauf- und Dienstpläne gegenüber. Ergänzend zu den zuvor aufgeführten Abbildungen 11.6 repräsentieren die durch Rauten dargestellten Ergebnisse die durch (robuste) integrierte Planung mit Zeitfenstern erstellten Lösungen. Bezugswert aller Lösungen ist auch hier jeweils eine Lösung, die durch rein kostenminimierende integrierte Planung berechnet wurde.

Zusätzlich schlüsselt Tabelle 11.7 exemplarische Lösungen für robuste integrierte Planung mit Zeitfenstern auf und vergleicht sie mit Lösungen, die durch robuste integrierte Planung mit nachträglichem Retiming erstellt wurden. Dabei werden die genutzte Planungsmethode (*approach*), das erlaubte Verschiebeintervall (TW), sowie die Änderungen im Fahrplan in Bezug auf die Anzahl verschobener Fahrten (*shifted # services*), die verschobene Gesamtzeit in Minuten (*shifted time*) und die durchschnittliche zeitliche Verschiebung einer angepassten Servicefahrt (*shifted t per serv*) angegeben. Außerdem wird der benötigte Ressourceneinsatz in geplanter Umlauf- (#v) und Dienstanzahl (#d), sowie die geplanten Kosten (*costs in % to int*) und die potenzielle Verspätungspropagation (*E(SD) in % to int*) in Bezug zum Basisfall der integrierten Planung mit reiner Kostenminimierung aufgeführt.

Ergebnisinterpretation (ad ITW 1.) Die Abbildungen 11.16 verdeutlichen die aus den größten Freiheitsgraden für die Planung des Ressourceneinsatzs resultierende Dominanz der integrierten Planung mit Zeitfenstern gegenüber den anderen Planungsmethoden. Erkauft wird diese Dominanz durch weniger Ressourcen mit mehr auszuführenden Aufgaben, oder bei gleicher Ressourcenanzahl durch einen angepassten Fahrplan, so dass trotz mehr pro Ressource eingeplanter Aufgaben weniger Ver-

Tabelle 11.7: Vergleich robust integriert erstellter Pläne (intR) mit nachträglichem Retiming (retiming) und robust integriert erstellter Pläne mit Zeitfenstern (intTW) mit ähnlicher Robustheit

approach	TW±	shifted #services	shifted time	t per serv	#v	#d	costs in % to int	E(SD) in % to int
				Testinstanz A				
intR	x	x	x	x	17	25	100,0	62,3
retiming	2	103 (48,8%)	197	1,91	17	25	100,0	47,5
	4	105 (49,8%)	346	3,30	17	25	100,0	43,5
intTW	2	118 (55,9%)	204	1,73	15	25	96,5	42,7
intR	x	x	x	x	17	25	100,2	51,1
retiming	2	98 (46,4%)	184	1,88	17	25	100,2	41,0
	4	100 (47,4%)	320	3,20	17	25	100,2	37,3
intTW	2	30 (14,2%)	54	1,80	16	24	96,5	38,6
intR	x	x	x	x	17	25	101,0	37,2
retiming	2	110 (52,1%)	201	1,83	17	25	101,0	26,7
	4	108 (51,2%)	326	3,02	17	25	101,0	22,2
intTW	2	59 (28,0%)	101	1,71	16	25	99,1	24,9
				Testinstanz B				
intR	x	x	x	x	21	29	100,0	76,3
retiming	4	198 (58,4%)	484	2,44	21	29	100,0	65,1
	6	204 (60,2%)	600	2,94	21	29	100,0	63,7
intTW	4	139 (41,0%)	364	2,62	20	28	97,6	64,5
intR	x	x	x	x	21	29	100,2	66,0
retiming	4	198 (58,4%)	529	2,67	21	29	100,2	54,0
	6	206 (60,8%)	663	3,22	21	29	100,2	52,6
intTW	4	114 (33,6%)	296	2,60	20	29	98,2	52,7
intR	x	x	x	x	21	31	105,8	40,4
retiming	4	202 (59,6%)	571	2,83	21	31	105,8	29,6
	6	212 (62,5%)	720	3,40	21	31	105,8	28,2
intTW	4	20 (5,9%)	49	2,45	21	29	101,9	28,8
				Testinstanz C				
intR	x	x	x	x	33	65	100,0	77,4
retiming	4	394 (59,6%)	1209	3,07	33	65	100,0	57,8
	6	405 (61,3%)	1571	3,88	33	65	100,0	54,5
intTW	4	90 (13,6%)	210	2,51	32	62	96,9	53,5
intR	x	x	x	x	33	65	100,0	68,5
retiming	4	386 (58,4%)	1152	2,98	33	65	100,0	49,9
	6	395 (59,8%)	1541	3,90	33	65	100,0	48,7
intTW	4	77 (11,6%)	197	2,56	32	63	97,8	48,3
intR	x	x	x	x	33	65	100,5	53,8
retiming	4	357 (54,0%)	1102	3,98	33	65	100,5	43,4
	6	366 (55,4%)	1476	4,03	33	65	100,5	41,3
intTW	4	52 (7,9%)	124	2,38	33	64	99,0	43,1

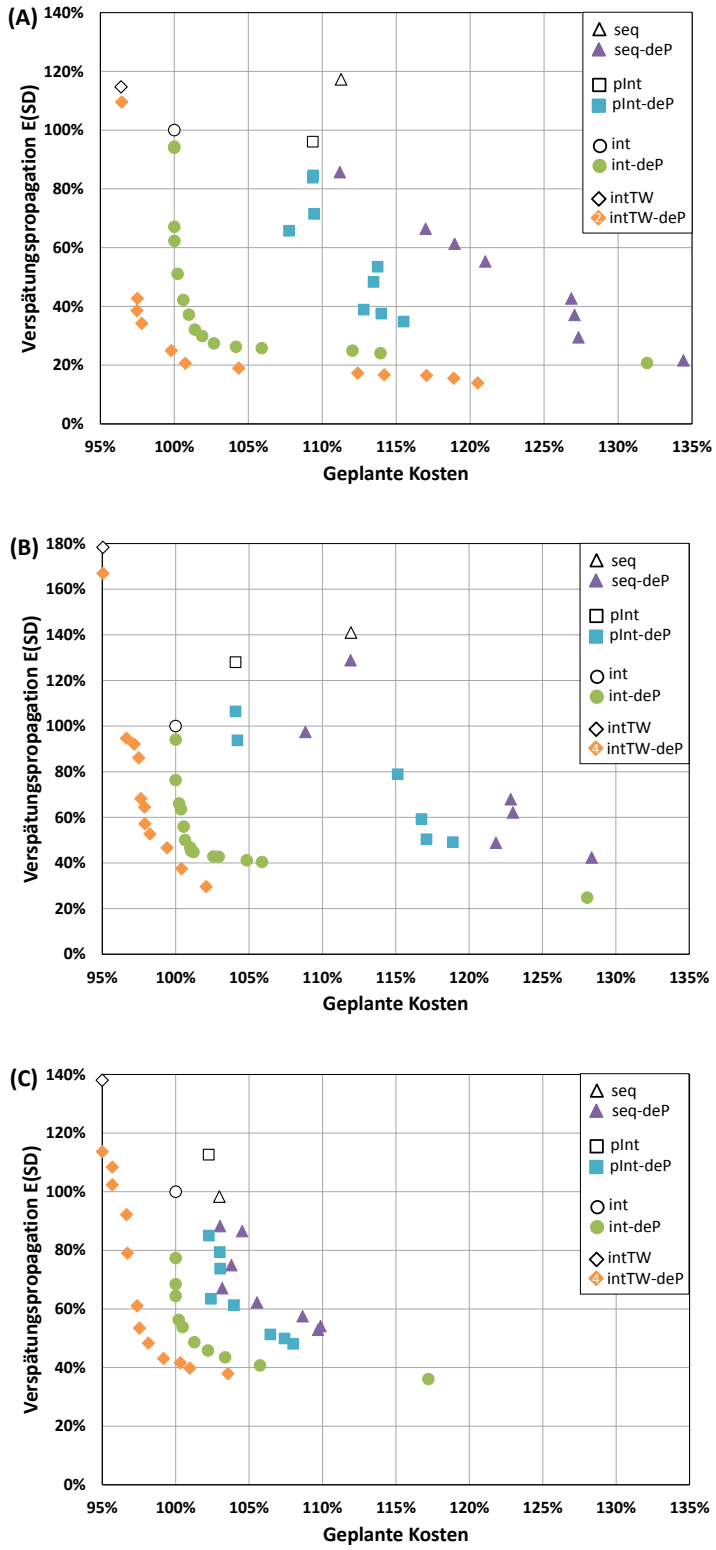

Abbildung 11.16: Gegenüberstellung geplanter Kosten und erwarteter Verspätungspropagation bei sequenzieller (seq), teilintegrierter (pInt), integrierter (int) und integrierter Planung mit Zeitfenstern (intTW) mit Berücksichtigen von Verspätungspropagation (deP)

spätungspropagation auftreten kann. Es zeigt sich, dass in Abhängigkeit der Instanz und der gewählten Verschiebeintervallgröße Lösungen mit z.T. deutlich weniger Ressourcen erstellt werden können als bei integrierter Planung. Ohne Berücksichtigung von Robustheitsaspekten führt dies zu wesentlich unrobusteren Plänen (mit weniger Warte- und Standzeiten bei der Aufgabenausführung). Sobald Robustheit als Teilziel verfolgt wird, können robustere Pläne zu gleichen Kosten bzw. sogar robustere Pläne zu geringeren Kosten erstellt werden. Die erkennbaren vereinzelten Sprünge in den Lösungsfronten auf der Kostenseite weisen jeweils auf einen gesteigerten Ressourceneinsatz bzw. auf das Gleichziehen mit der ursprünglich benötigten Ressourcenanzahl bei integriert erstellten Plänen hin.

(ad ITW 2.) Tabelle 11.7 zeigt ergänzend, dass durch integrierte Planung mit Zeitfenstern aufgrund der verfügbaren Freiheitsgrade die Lösungsqualität im Vergleich zu einer robusten integrierten Planung mit nachträglichem Retiming noch gesteigert werden kann. Dabei müssen im Vergleich zu nachträglichen Fahrplananpassungen weniger Fahrten verschoben werden. Die einzelnen Verschiebungen pro Fahrt sind zusätzlich weniger stark ausgeprägt. Überdies lassen sich sogar bei kleineren Verschiebeintervallen robustere Pläne erstellen. Die potenziell höhere Robustheit lässt sich darüber hinaus zu geringeren Kosten des Ressourceneinsatzes erreichen.

Planstrukturen Die enstehenden Planstrukturen verdeutlichen Veränderungen auf der Ebene des Ressourceneinsatzes sowie des Fahrplanes und erklären so das gezeigte Potenzial zur Verringerung der Verspätungspropagation bei gleichzeitiger Verringerung der Umlauf- und Dienstanzahl. Exemplarisch für Testinstanz C zeigt Abbildung 11.17 die Auswirkungen des geplanten Ressourceneinsatzes auf Verspätungspropagation im Zusammenhang mit der (durch Fahrtenverschiebungen angepassten) Ganglinie der auszuführenden Servicefahrten. Verglichen werden rein kostenminimierende integrierte Planung, robuste integrierte Planung und (robuste) integrierte Planung mit Zeitfenstern. Die aufgeführten Propagationswerte stellen die auftretende propagierte Verspätung in Minuten für die jeweils zum betrachteten Zeitpunkt im Plan beginnenden Servicefahrten dar. Die Werte basieren auf der für die Planerstellung genutzten und während der Optimierung für alle Lösungen gleichermaßen eingesetzten Menge von Verspätungsszenarien.

(ad ITW 3.) Deutlich erkennbar unterscheiden sich die aus den verschiedenen Lösungen resultierenden Propagationseffekte. Bemerkenswert ist dabei, dass trotz geringeren Ressourceneinsatzes – es werden ein Umlauf und drei Dienste weniger als bei reiner integrierter Planung benötigt – und insgesamt weniger gemeinsam auf Umläufen und Diensten eingeplanter Pufferzeit, geringere Propagationseffekte in der

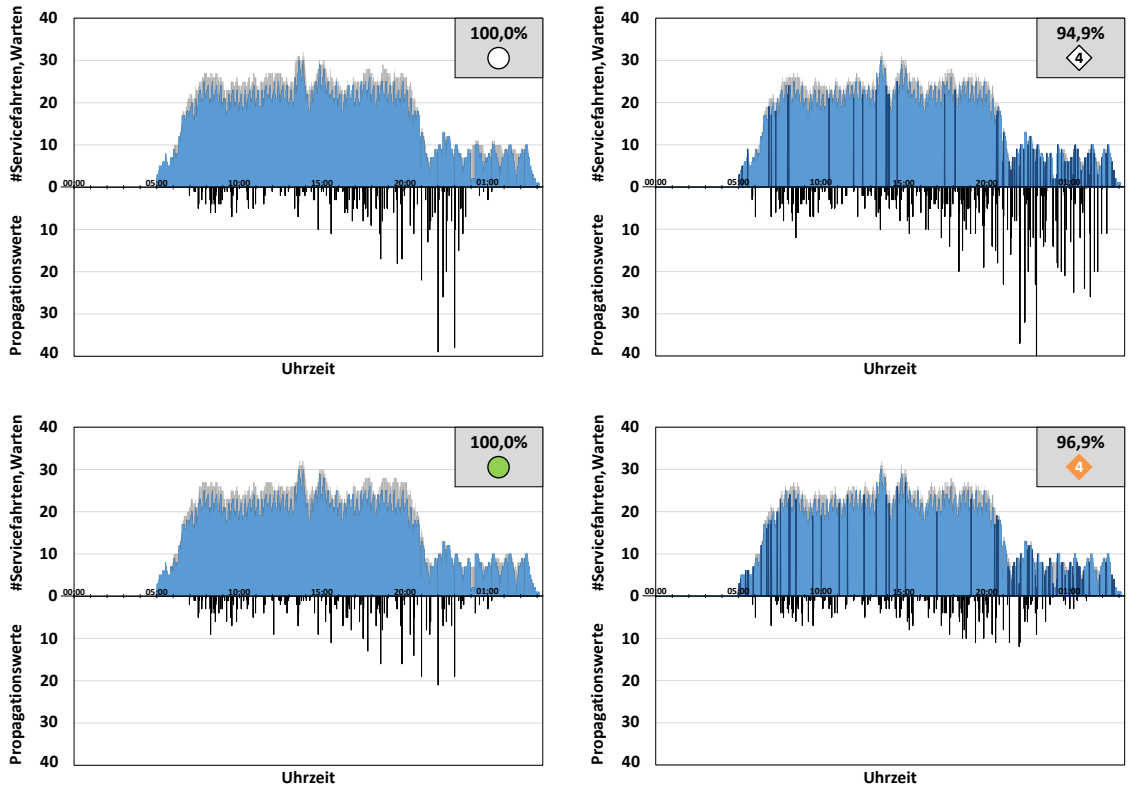

Abbildung 11.17: Gegenüberstellung von Propagationseffekten (*schwarz*) im eingeplanten Ressourceneinsatz, sowie Fahrtenverschiebungen (*dunkelblau*), Anpassungen der Ganglinie (*hellblau*) und Verteilung von Pufferzeiten (*grau*) durch integrierte (robuste) Planung mit Zeitfenstern im Vergleich zu integrierter Planung für Testinstanz C

durch integrierte robuste Planung mit Zeitfenstern erstellten Lösung auftreten. Im Gegensatz dazu führen Fahrtenverschiebungen ohne Berücksichtigen von Robustheitsaspekten zu deutlich unrobusteren Plänen, bei zusätzlicher Reduktion des Ressourceneinsatz um zwei weitere Dienste. Wie in den Bildern dargestellt, überschneiden bzw. unterscheiden sich die verschobenen Fahrtenmengen. In den Ganglinien sind jeweils die Zeitpunkte (dunkelblau) markiert, an denen mindestens eine zeitlich verschobene Servicefahrt startet. Im Vergleich der mit erlaubten Fahrplananpassungen erstellten Lösungen sind bei robuster Planung 90 Fahrten, sowie 91 Fahrten bei rein kostenminimierender Planung verschoben. Von den verschobenen Fahrten sind in 40 (44,4%) Fällen die gleichen Fahrten betroffen. Von diesen sind wiederum 13 (14,4%) Fahrten in genau gleichem Umfang und in gleicher Richtung verschoben. Das geringe Ausmaß an identischen Fahrtenverschiebungen lässt sich auch beim Vergleich anderer Lösungen und Instanzen beobachten. So zeigt beispielsweise der Vergleich äquivalenter Lösungen für Testinstanz B Fahrtenverschiebungen, die zu ca. 21% die gleichen

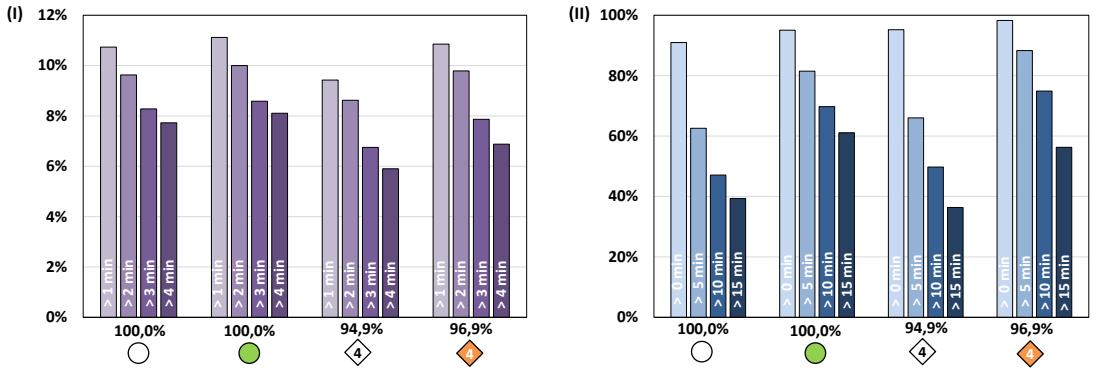

Abbildung 11.18: Gegenüberstellung der Kennzahlbewertung zur Verdichtung (Bild I) und Vernetzung (Bild II) des (robust) integriert (mit Zeitfenstern) eingeplanten Ressourceneinsatzes für Testinstanz C

Fahrten betreffen, allerdings insgesamt nur 7% identisch verschobenen Fahrten aufweisen.

(ad ITW 4.) Ergänzend zu den Ganglinien und eingeplanten Pufferzeiten der verschiedenen Lösungen stellt Abbildung 11.18 die in Abschnitt 11.3.2 vorgeschlagenen Plankennzahlen zu Verdichtung und Vernetzung gegenüber. Der Vergleich verdeutlicht die möglichen Ursachen dafür, dass durch robuste integrierte Planung mit Zeitfenstern erstellte Pläne trotz geringeren Ressourceneinsatzes und in Summe geringeren Pufferzeiten weniger Verspätungen propagieren können als (robust) integriert erstellte Pläne. Es stehen zum einen innerhalb der Dienststücke ausreichend Pufferzeiten zwischen der Ausführung der Servicefahrten zur Verfügung. Im Vergleich der Anteilswerte von Pufferzeiten in den Dienststücken bis zu drei Minuten sind die robusten Lösungen fast identisch ausgeprägt. Darüber hinaus steht bei der Verknüpfung der Aufgabensequenzen bis zu der Betrachtung der Anteile von über 15 Minuten Pufferzeit verhältnismäßig mehr Pufferzeit zur Verfügung. Sowohl innerhalb der Dienststücke als auch bei der Verknüpfung von Aufgabensequenzen ist erst bei der Betrachtung größerer Pufferzeiträume der robust integriert geplante Ressourceneinsatz dem unter Einbeziehen von Zeitfenstern geplanten Ressourceneinsatz überlegen. Die geringere Gesamtverspätungspropagation bei robuster integrierter Planung mit Zeitfenstern kann somit durch die durch Fahrtenverschiebungen besser auf die auszuführenden Fahrtensequenzen angepasste Verteilung von Pufferzeiten erreicht werden. Im Gegensatz dazu weist rein kostenminimierende integrierte Planung mit Zeitfenstern ohne Robustheitsbetrachtung deutlich geringere Pufferzeiten bei Verknüpfung der Aufgaben bzw. Aufgabensequenzen auf.

Kernaussagen Die Untersuchungen zeigen, dass die äußerst umfangreiche und rechenintensive Problemstellung der simultan ausgeführten Umlauf- und Dienstplanung mit gleichzeitiger Anpassung des Fahrplans unter Einbeziehung von sowohl Kosten als auch Robustheitsaspekten geeignet lösbar ist. Ein solches Verfahren ist aufgrund der Freiheitsgrade überlegen gegenüber einer Ressourceneinsatzplanung mit nachträglichem Retiming.

Es können durch passende Fahrtenverschiebungen gleichzeitig der Ressourceneinsatz und die geplanten Kosten verringert sowie die Robustheit des Ressourceneinsatzes erhöht werden. Dabei ist (im Vergleich zu rein kosteniminierender Planung mit Zeitfenstern) die Menge von Servicefahrten identifizierbar, deren geringfügige Anpassung die Kosten für den Ressourceneinsatz verringern kann ohne die Robustheit zu gefährden bzw. die Menge von Servicefahrten, deren Anpassung in Abstimmung mit dem Ressourceneinsatz zu robusteren Pläne führen kann.

11.4.3 Zusammenfassung der Untersuchungen

In diesem Teilkapitel wurden mit dem Einbeziehen von Zeitfenstern zum Verschieben von Servicefahrten der Einfluss zusätzlicher Freiheitsgrade in der Planung untersucht, um Robustheit und Kosteneffizienz des Ressourceneinsatzes zu erreichen. Dabei wurden Lösungen durch integrierte Planung mit Zeitfenstern, sowie durch nachträgliches Retiming des Fahrplans basierend auf sequenziell, teilintegriert und integriert erstellten Ressourceneinsatzplänen analysiert und miteinander bzw. mit den in den vorangeganenen Teilkapiteln ausgearbeiteten Ergebnissen verglichen. Die unterschiedlichen Zeitfenstergrößen zum Anpassen des Fahrplans wurden so gewählt, dass nur geringe Verschiebungen einzelner Fahrten möglich sind, um die ursprüngliche Fahrplanstruktur weitestgehend beibehalten zu können. Größere Verschiebeintervalle wurden beim Retiming eingesetzt, um eine Konvergenz der erreichbaren Auswirkungen in den Ergebnissen zu verdeutlichen. Mit Retiming wurde Einfluss auf die Verdichtung von gemeinsam auf Umlauf- und Dienst auszuführenden Aufgabensequenzen genommen. Die integrierte Planung mit Zeitfenstern beeinflusst zusätzlich zur Verdichtung auch die mögliche Vernetzung von Aufgabensequenzen unterschiedlicher Ressourcen.[3]

Im Vergleich zur reinen Umlauf- und Dienstplanung resultieren aus der zusätzlichen Anpassung des Fahrplans Verbesserungen der Ergebnisse, so dass eine weitere Verschiebung der Lösungsfronten hin zu geringeren Kosten und geringerer Verspätungspropagation möglich ist.

[3] Für die grundlegende Diskussion über den Einfluss von Verdichtung und Vernetzung auf Verspätungspropagation sei auf Abschnitt 11.3.1 verwiesen.

(ad F2.) Dabei können zum einen durch Retiming-Verfahren Pläne kostenneutral robuster gestaltet werden. (ad F3.) Innerhalb von Dienststücken werden dazu in der Tendenz mehr Pufferzeiträume zwischen einzelnen Aufgaben geschaffen und vormals sehr kleine bzw. übermäßig große Pufferzeiträume der anhand der eingeplanten Aufgabensequenzen erwartbaren Verspätungspropagation angepasst. Aufgrund der höheren Aufgabendichte in integriert erstellten Plänen, kann dort im Vergleich zu sequenzieller oder teilintegrierter Planung ein größerer Effekt zur Erhöhung der Robustheit erzielt werden.

(ad F1.) Zum anderen zeigt sich, dass eine integrierte Planung mit Zeitfenstern den Einfluss des nachträglichen Retimings von (robust) integriert erstellten Ressourceneinsatzplänen noch übertreffen kann, da ein gleichzeitiges Verändern von sowohl Kosten als auch Robustheit möglich ist, statt durch Fahrplananpassungen allein die Robustheit des Ressourceneinsatzes zu beeinflussen. Zusammenfassend lässt sich feststellen, je mehr Planungsstufen integriert betrachtet werden, desto größer sind sowohl die Freiheitsgrade für Robustheit als auch Effizienz in der Ressourceneinsatzplanung. Darüber hinaus ist festzuhalten, dass die sehr komplexe Problemstellungen der Umlauf- und Dienstplanung in Kombination mit Fahrplananpassungen und Berücksichtigung zweifacher Zielsetzungen aufgrund der explodierenden Problemgrößen nur durch die vorgeschlagenen Methoden aus Abschnitt 9.4.3 bzw. 9.4.2 in dem dargestellten Umfang handhabbar ist.

(ad F3) Es ist anzumerken, dass die sich für den Fahrplan ergebenden Änderungen in höchstem Maße abhängig von den verschiedenenen bei der Ressourceneinsatzplanung zu berücksichtigenden Dienstregeln und Restriktionen sind. So bewirken beispielsweise andere Vorgaben für Pausenzeiten bei Einsatz der integrierten Planung mit Zeitfenstern das Verschieben anderer Servicefahrten. Die erfolgten Anpassungen sind deshalb schwer verallgemeinerbar. Sie ermöglichen jedoch die Identifikation von Servicefahrten, deren Verschiebung sowohl Robustheit als auch Kosteneffizienz des Ressourceneinsatzes erhöhen kann. Dabei liefern sie bei Einsatz größerer Zeitfenster einen Hinweis auf anzupassende Frequenzen und Umsteigebeziehungen im Fahrplan.

11.5 Einfluss der Ausweitung des betrachteten Robustheitsbegriffs

In den vorangegangenen Teilkapiteln wurde als Robustheit die reine Stabilität bei der Ausführung des berechneten Ressourceneinsatzes evaluiert. Somit wurde zur Bewertung der aus den verschiedenen Lösungsmethoden resultierenden Ressourceneinsatz-

pläne bisher eine untere Schranke der insgesamt möglichen Robustheit betrachtet. In diesem Teilkapitel wird die Sicht auf die Ausführungsebene erweitert, indem bei der Evaluation zusätzlich die Flexibilität des Ressourceneinsatzes berücksichtigt wird (vgl. Abschnitt 3.3). Dafür werden dispositive Maßnahmen während der Aufführung simuliert (vgl. Steuerung der Planausführung in Abschnitt 2.4). Die Maßnahmen beinhalten einen kostenfreien Ressourcentausch zur Erfüllung von Aufgaben und zum Auflösen von potenziellen Verspätungskonflikten. Um eine Vergleichbarkeit aller Ergebnisse dieser Arbeit im Hinblick auf die gemessene Verspätungspropagation und den ausgeführten Ressourceneinsatz zu gewährleisten, ist dabei das Absagen von Fahrten als dispositive Maßnahme nicht erlaubt.

Mit der Ausweitung des betrachteten Robustheitsbegriffs sollen die bisher herausgearbeiteten Erkenntnisse zu den drei in Kapitel 4 formulierten Forschungsleitfragen vertieft werden. Insbesondere Forschungsleitfrage F2 nach realisierbarer Robustheit in kosteneffizienten Plänen wird im Folgenden verstärkt untersucht. Dabei wird ein mögliches Zusammenspiel zwischen Stabilitäts- und Flexibilitätsaspekten des Ressourceneinsatzes thematisiert.

Im Folgenden diskutiert Abschnitt 11.5.1 den Einsatz kostenneutraler dispositiver Maßnahmen während der Ausführung. Abschnitt 11.5.2 fasst die Untersuchungen zusammen.

11.5.1 Berücksichtigen von Ressourcentausch in der Ausführung

In diesem Abschnitt werden die zuvor bereits betrachteten Ressourceneinsatzpläne durch eine erweiterte Simulation evaluiert. Diese bildet neben der Weitergabe von Verspätungspropagation auch kostenneutrale dispositive Maßnahmen ab. So wird der Einfluss von Flexibilität in den erstellten Ressourceneinsatzplänen auf Robustheit während der Ausführung herausgearbeitet und kann mit dem reinen Einfluss von Stabilität verglichen werden.

Untersuchungsfragen Die durchgeführten Experimente basieren auf folgenden Fragestellungen:

FLEX 1.) Ist in den erstellten Plänen Potenzial zur Reduktion von Verspätungspropagation durch Tauschoperationen während der Ausführung vorhanden?

FLEX 2.) In welchem Ausmaß unterscheidet sich das Potenzial in Abhängigkeit der ursprünglich eingesetzten Planungsmethoden?

FLEX 3.) Welche Planstrukturen begünstigen den Einsatz von Tauschoperationen?

Genutzte Artefakte Es werden durch sequenzielle, teilintegrierte, integrierte und integrierte Planung mit Zeitfenstern erstellte Ressourceneinsatzpläne und zugehörige Fahrpläne betrachtet, die zuvor in Abschnitt 11.4.2 gemeinsam diskutiert wurden. Die generierten Pläne werden nun durch Simulation mit zusätzlichen Dispositionsregeln evaluiert (vgl. Abschnitt 5.3.1). Um einen größtmöglichen Effekt zu zeigen, werden Tauschoperationen bereits dann durchgeführt, wenn durch eine primäre Verspätung bereits ein Schwellenwert von sehr geringer Verspätungspropagation überschritten würde (konkret liegt der Parameter der Verspätungspropagationstoleranz θ bei 10 Sekunden). Dieser Schwellenwert wird im Rahmen einer Sensitivitätsanalyse variiert.

Ergebnisdarstellung Abbildung 11.19 stellt die geplanten Kosten der einzelnen Lösungen ihrer potenziellen Verspätungspropagation gegenüber. Dabei repräsentieren die schwarz eingefärbten Punkte die in den vorangegangenen Teilkapiteln im Hinblick auf reine Stabilität des Ressourceneinsatzes evaluierten Lösungen. Die farbigen Punkte zeigen die Evaluation der potenziellen Verspätungspropagation durch Ausnutzen von Stabilitäts- und Flexibilitäts-Eigenschaften der errechneten Pläne.

Ergänzend schlüsselt Tabelle 11.8 die Anzahl der während der Simulationsläufe durchgeführten Tauschoperationen für unterschiedlich ausgeprägte Lösungen auf. Am Beispiel der Testinstanz C werden als Vergleich der verschiedenen Planungsvarianten und Lösungsmethoden jeweils Ergebnisse zu rein kostenminimierender Planung und robuster kostenminimierender Planung mit Berücksichtigen von Verspätungspropagation durch globale Flussdekomposition angegeben. Dabei werden neben der genutzten Planungsvariante (*scheme*), der Planungsmethode (*approach*), den geplanten Kosten (*planned costs in %*), der gemessenen Verspätungspropagation (*E(SD) in %*) und dem Unterschied zu reiner Stabilitätsmessung (Δ *to STAB*), Statistiken zu den Tauschoperationen (*swaps*) aufgeführt. Angegeben werden die durchschnittliche Anzahl (*avg*) Tauschoperationen über alle Simulationsläufe inklusive Standardardabweichung (*std*), sowie die geringste (*min*) und die höchste (*max*) gemessene Anzahl ausgeführter Tauschvorgänge.

Ergebnisinterpretation (ad FLEX 1.) Die Abbildung 11.19 illustriert das Potenzial von Tauschoperationen während der Ausführung zur Reduktion von Verspätungspropagation. Gleichzeitig verdeutlicht sie einen Unterschied in der Bewertung potenzieller Robustheit der erstellten Ressourceneinsatzpläne. Im Vergleich zur Messung von Stabilität durch das reine Abbilden von Verspätungspropagation (schwarze Punkte) sind durch das Einbeziehen von Flexibilitätsaspekten (farbige Punkte) die erstellten Pläne in der Regel robuster ausgewiesen. So können als stabil gemessene

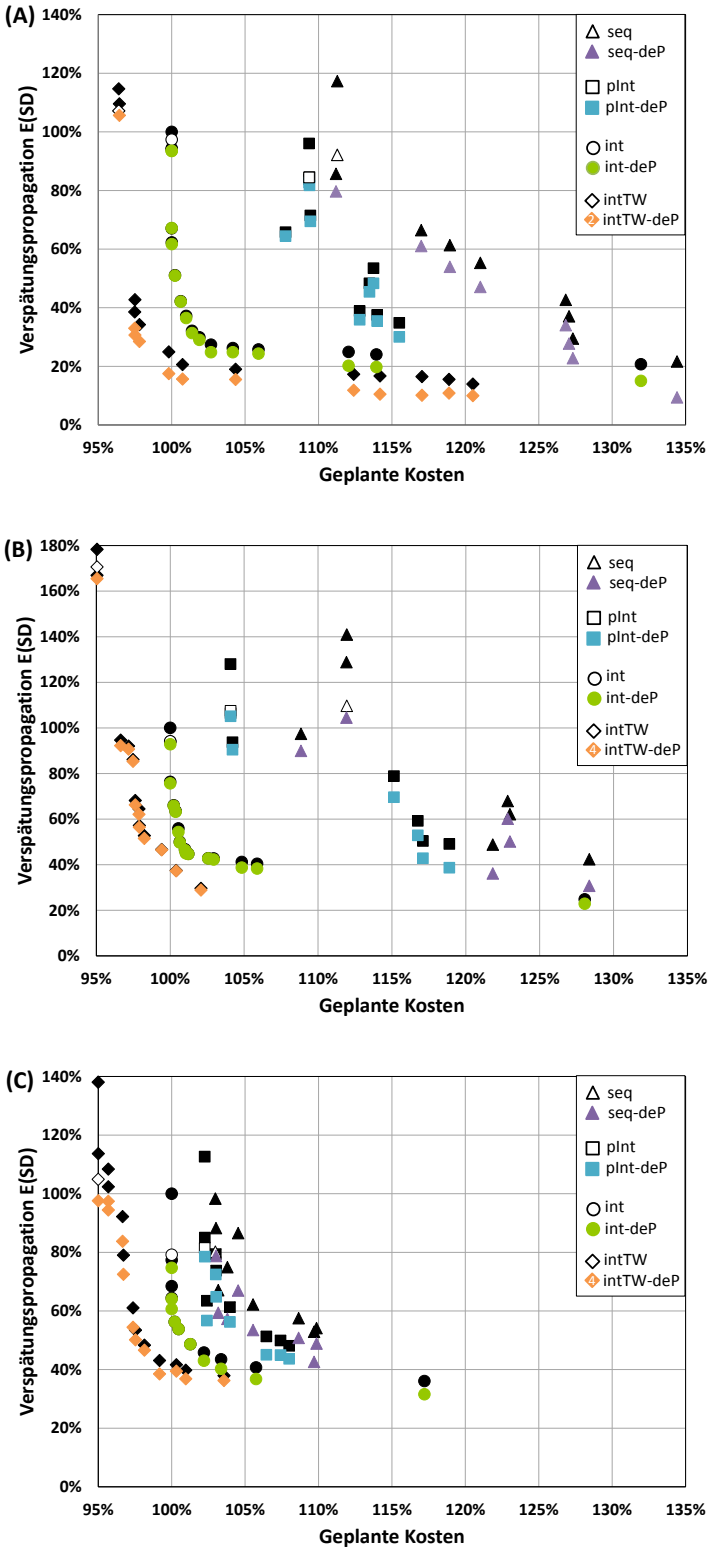

Abbildung 11.19: Gegenüberstellung geplanter Kosten und erwarteter Verspätungspropagation bei sequenzieller (seq), teilintegrierter (pInt), integrierter (int) und integrierter Planung mit Zeitfenstern (intTW) mit Berücksichtigen von Verspätungspropagation (deP) und Variation der Robustheitsevaluation

Tabelle 11.8: Gemessene Verspätungspropagation und Anzahl Tauschoperationen während der Ausführung als arithmetischer Mittelwert über die Simulationsläufe am Beispiel von Testinstanz C

scheme	approach	E(SD) in %	Δ to STAB	planned costs in %	swaps avg	swaps std	swaps min	swaps max
SEQ	VSP+CSP	80,2	-18,1	103,0	17,2	4,3	5	28
	VSP-GFD+CSP	78,8	-9,5	103,0	2,5	1,9	0	8
PINT	VSP CSP	81,5	-31,1	102,3	19,0	4,5	4	26
	VSP CSP GFD	78,5	-6,6	102,3	1,5	0,8	0	7
INT	IVCSP	79,2	-20,8	100,0	20,4	4,8	5	31
	IVCSP GFD	74,7	-3,3	100,0	0,9	1,1	0	5
INT-TW	IVCSP-TW	104,8	-33,2	94,9	26,1	5,4	7	36
	IVCSP-TW GFD	97,6	-16,1	94,9	2,9	1,6	0	10

Pläne noch ein wenig robuster bewertet werden und eher unstabile Pläne z.T. als signifikant robuster. Dabei ist zu beachten, dass bei den Experimenten die Schwelle zur Durchführung eines Ressourcen- bzw. Aufgabentausches sehr niedrig angesetzt ist. So werden beispielsweise ab einem Schwellenwert von bereits drei Minuten etwas weniger Tausche durchgeführt. Die dargestellten Ergebnisse sind daher als obere Schranke der potenziellen Robustheit zu sehen. Die mögliche Robustheit der Pläne liegt je nach Ausprägung der Steuerung des Ressourceneinsatzes in der Ausführungsphase zwischen den jeweils ausgewiesenen Werten zur Messung reiner Stabilität und der kombinierten Evaluation von Stabilität und Flexibilität.

(ad FLEX 2.) Auffällig ist die Verbesserung der Robustheit in der Evaluation der rein kostenminimierend erstellten Pläne. Wie in Tabelle 11.8 gezeigt wird, kann durch eine hohe Anzahl von Tauschoperationen ein zuvor unstabil eingeplanter Ressourceneinsatz z.T. kompensiert werden. Im Vergleich zwischen rein kostenminimierenden Lösungen und den korrespondierenden aufgabengleichen Lösungen unter Einbeziehen von Verspätungspropagation verringert sich so der Abstand bei der gemessenen Propagation von Verspätungen. Das Vermeiden von Verspätungspropagation durch geschickte Ressourcenverknüpfung wird also von der Planungs- in die Ausführungsphase verlagert. Der Qualitätsabstand zwischen den eher unstabil und stabil eingeplanten aufgabengleichen Lösungen kann dadurch aber nicht vollständig geschlossen werden.

In der Robustheitsbewertung hochgradig stabiler und kosteneffizienter, integriert erstellter Pläne zeigen sich dahingegen kaum Verbesserungen. In den zugehörigen Simulationen werden nur wenige Tauschoperationen durchgeführt. Wie in der Abbildung 11.19 für die einzelnen Instanzen zu erkennen ist, werden in der Regel erst

mit deutlich höheren Plankosten – also zunehmender Zahl eingesetzter Ressourcen – Tauschoperationen zur weiteren Reduktion von Verspätungspropagation realisiert.

Im Vergleich zu integriert erstellten Plänen wird überdies umlaufbasiert erstellter Ressourceneinsatz in den Experimenten deutlicher durch Tauschoperationen in der Ausführung beeinflusst. Ebenfalls offenbaren Lösungen, die durch integrierte Planung mit Zeitfenstern auch den zugrunde liegenden Fahrplan anpassen, z.T. deutliches Potenzial zur Reduktion von Verspätungspropagation durch Tauschoperationen. Im Folgenden wird mit Blick auf die Planstrukturen untersucht, ob die vorhandenen Möglichkeiten zum Ressourcentausch oder eher der Bedarf an Tauschoperationen für die unterschiedlichen Robustheitsbewertungen der erstellten Ressourceneinsatzpläne ausschlaggebend sind.

Planstrukturen (ad FLEX 3.) Die Möglichkeiten zum Aufgaben- bzw. Ressourcentausch sind in den erstellten Lösungen verschieden ausgeprägt. Abbildung 11.20 stellt in den Bildern I die Anzahl möglicher Tauschpunkte von Lösungen der unterschiedlichen Planungsvarianten gegenüber. Dabei können für den Tausch von Ressourcen maximal so viele Tauschpunkte vorhanden sein wie Dienststücke geplant sind (siehe Definition in Abschnitt 5.2.3). Die Anzahl der sich daraus ergebenden Möglichkeiten zum Tausch von (in sich abgeschlossenen) Aufgabensequenzen zwischen unterschiedlichen Ressourcen ist deutlich größer. Insbesondere durch den Tausch an Depot- bzw. Pausenstandorten und damit zur Verfügung stehender Auswahl an Ressourcen zum Tausch, liegen diese in den betrachteten Instanzen insgesamt um das ca. Vier- bis Neunfache höher als die Anzahl der Tauschpunkte vorgibt.

Die Abbildung 11.20 zeigt in den Bildern I der linken Spalte exemplarisch für die Testinstanzen A, B und C, dass in den betrachteten Lösungen eine Vielzahl an Tauschpunkten generiert werden. In vielen Lösungen liegt deren Anzahl bei mindestens der Hälfte aller maximal möglichen Tauschpunkte, bei einigen deutlich höher. So können in nahezu allen Ressourceneinsatzplänen mindestens die Aufgabensequenzen jedes zweiten Dienststücks kostenneutral zwischen Umläufen getauscht werden. Dabei weisen ein Großteil der Tauschpunkte mindestens dreißig Minuten Pufferzeit zwischen beendeten Aufgaben und zu tauschenden Aufgabensequenzen aus. Viele Tauschpunkte besitzen sogar über sechzig Minuten Pufferzeit und erlauben so auch sehr große kumulierte Verspätungen zu absorbieren. Nahezu alle Tauschpunkte werden durch die Pausenbildung bei der Diensterstellung zentral an Depots bzw. Pausenorten eingeplant, so dass aureichend Ressourcen zum Tausch verfügbar sind.

Die in der rechten Spalte der Abbildung 11.20 dargestellten Bilder II zeigen darüber hinaus den Einsatz von Tauschvorgängen in Abhängigkeit eines vorhandenen Ver-

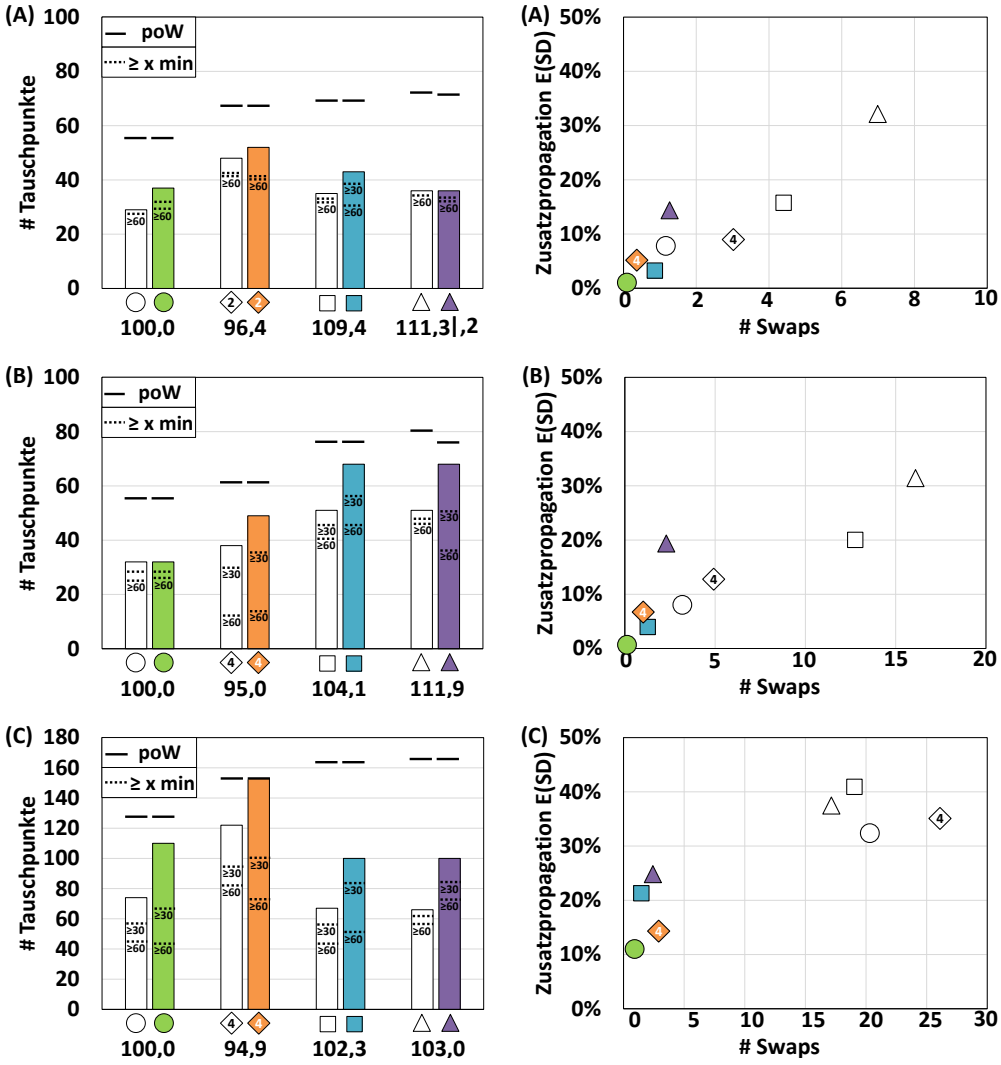

Abbildung 11.20: Anzahl Tauschpunkte mit Bezug zur Anzahl eingeplanter Dienststücke (poW) im durch sequenzielle, teilintegrierte, integrierte und integrierte Planung mit Zeitfenstern erstellten Ressourceneinsatz, sowie tatsächlich durchgeführte Tausche (Swaps) in Bezug zur auftretenden Verspätungspropagation durch Vernetzung der Ressourcen

netzungseffektes des kombinierten Ressourceneinsatzes. Erkennbar werden bei den kosteneffizienten, robust integriert erstellten Lösungen kaum Tauschvorgänge durchgeführt, obwohl Tauschmöglichkeiten vorhanden sind. Überdies implizieren weniger vorhandene Tauschpunkte nicht notwendigerweise weniger durchgeführte Tauschvorgänge. Eine Erklärung bietet der nicht vorhandene Bedarf nach Tauschvorgängen aufgrund der in Abschnitt 11.3 gezeigten Isolation von Verspätungspropagation in Dienststücken bei eben solchen Plänen. Im Gegensatz dazu existiert bei anderen – speziell auch umlaufbasiert erstellten – Lösungen die Notwendigkeit bzw. Gelegenheit den dort stärker ausgeprägten Vernetzungseffekt von Verspätungspropagation durch Tauschoperationen verringern zu können.

(ad FLEX 2.) Im Vergleich zwischen integrierter Planung und integrierter Planung mit Zeitfenstern offenbaren durch letztere Planungsvariante erstellte Pläne mehr Tauschpunkte. Dies liegt in einer Veränderung der Dienstplanstruktur begründet. In den betrachteten Lösungen führen die Fahrplananpassungen zu einem geringeren Ressourceneinsatz. Gleichzeitig werden aufgrund der erfolgten Anpassungen in der Tendenz Diensttypen mit kürzeren Mindestpausenzeiten und einer höheren Anzahl von Dienststücken bevorzugt erstellt (siehe Anzahl Dienststücke in Abbildung 11.20). Durch die erhöhte Anzahl an Dienststücken werden mehr Möglichkeiten für Tauschpunkte von kompletten auf einem Umlauf und Dienst gemeinsam eingeplanten Aufgabensequenzen geschaffen. Diese werden aufgrund der in der Tendenz höheren Vernetzungseffekte von Verspätungspropagation eher genutzt.

11.5.2 Zusammenfassung der Untersuchungen

Dieses Teilkapitel erweitert die Bewertung der Robustheit des Ressourceneinsatzes um potenzielle Flexibilität in der Ausführung. Dazu wurden die zuvor bezüglich reiner Stabilität evaluierten Ressourceneinsatzpläne zusätzlich mit kostenneutralen dispositiven Maßnahmen während der Ausführung simuliert.

(ad F2.) Es zeigt sich, dass die errechneten Ressourceneinsatzpläne bereits ohne explizites Berücksichtigen von Flexibilitätsaspekten in der Planung während der Ausführung so modifizierbar sind, dass die Weitergabe von Verspätungen zusätzlich reduziert werden kann.

(ad F1.) Im Vergleich der Lösungen profitieren insbesondere unrobuste Pläne von einem Ressourcentausch während der Ausführung bzw. Lösungen, in denen potenziell im größerem Umfang Verspätungen durch Vernetzung der verschiedenen Ressourcen weitergeben wird. Kosteneffizient stabil erstellte Pläne weisen dagegen zwar nahezu die gleichen Möglichkeiten zum Tausch auf, allerdings ohne massiv Tauschoperationen in der Ausführung zu benötigen. Aus lösungsmethodischer Sicht betrachtet werden

also speziell bei unrobusten Plänen die Planungsentscheidungen der globalen Flussdekomposition zur Verringerung von Verspätungspropagation in die Ausführungsphase verlegt.

(ad F3.) Dies kann als Hinweis auf eine mögliche Wechselbeziehung zwischen Stabilitäts- und Flexibilitätsaspekten aufgefasst werden. Durch geschickte stabile Planung im Vorfeld brauchen für eine robuste Ausführung des Ressourceneinsatzes deutlich weniger kostenneutrale Anpassungen durchgeführt werden. Im Allgemeinen sind die Möglichkeiten für solche Tauschoperationen aufgrund von frequentierten gemeinsamen Depot bzw. Pausenorten, durch die kostengünstigere und somit bei der Planung forcierte Rückkehr von Fahrzeugen ins Depot häufig in großem Umfang für den Ressourceneinsatz bereits eingeplant.

11.6 Konsequenzen zum Erreichen Robuster Effizienz in der Planung

In diesem Teilkapitel werden die in den einzelnen Teilkapiteln 11.2 bis 11.5 vor dem Hintergrund der in Kapitel 4 formulierten Forschungsleitfragen herausgearbeiteten Erkenntnisse zum Erreichen eines robusten und kosteneffizienten Ressourceneinsatz übergreifend zusammmengefasst. Für die detaillierten Erkenntnisse der im Einzelnen betrachteten Untersuchungsfragen (siehe Abbildungen 11.1 und 11.2) sei auf die Untersuchungen bzw. Zusammenfassungen der jeweiligen Abschnitte der vorangegangenen Teilkapitel verwiesen.

Mit den durchgeführten Untersuchungen wurden drei wesentliche Dimensionen für das Erreichen eines robusten und kosteneffizienten Ressourceneinsatzes betrachtet: *Planungsmethodik*, *Planstrukur* und *Robustheitsaspekt*. Die Ausprägungen in den einzelnen Dimensionen bedingen sich zum einen gegenseitig. So wird ausgehend von der Wahl der Planungsmethode die Ausgestaltung der Planstrukturen und die dadurch ermöglichte Ausprägung der Robustheitsaspekte Stabilität und Flexibilität beeinflusst. Zum anderen kann in jeder Dimension unabhängig von den anderen eine für das Erreichen Robuster Effizienz zu bevorzugende Auswahl bzw. wünschenswerte Ausprägung identifiziert werden. Abbildung 11.21 vereinigt die zentralen Untersuchungsergebnisse dazu, die im Folgenden kurz erläutert werden.

Erstens wurde hinsichtlich einer geeigneten Wahl der *Planungsmethodik* gezeigt, dass das Berücksichtigen von Verspätungspropagation zur methodengeführten Pufferverteilung einer festen Pufferzeitvorgabe als Lösungsmethode vorzuziehen ist. Auch wurde gezeigt, dass die globale, den gesamten Ressourceneinsatz einbeziehende Be-

trachtung von Robustheit deutlich gegenüber lokalen Entscheidungen bei der Planerstellung zu präferieren ist. Ebenso deutlich wurde gezeigt, dass eine integrierte Ressourceneinsatzplanung aufgrund der bei der Planung zur Verfügung stehenden Freiheitsgrade zu favorisieren ist (vgl. Abschnitt 11.2). Darüber hinausgehende Freiheitsgrade in Bezug auf die Zielsetzung der Ressourceneinsatzplanung können durch (integrierte) Eingriffe in die Fahrplanung geschaffen werden (vgl. Abschnitt 11.4).

Zweitens wurde im Hinblick auf erstrebenswerte *Planstrukuren* herausgearbeitet, dass für die Ausgestaltung des robusten und kosteneffizienten Ressourceneinsatzes die Verdichtung von Aufgaben in Dienststücken als wesentlicher Faktor für Propagation vor einer Vernetzung der unterschiedlichen Ressourcen zu beachten ist (vgl. insbesondere Abschnitt 11.3). Im Idealfall kann bereits die Zusammensetzung von Aufgabensequenzen zu Dienststücken so erfolgen, dass eine rasche Absorption von Verspätungspropagation erzielt wird, wie es insbesondere durch nachträgliche auf den Ressourceneinsatz abgestimmte Fahrplananpassung gezeigt wurde (vgl. Abschnitt 11.4). Für die gesamte Minimierung von Verspätungspropagation ist jedoch das Zusammenspiel mit einer passenden Vernetzung entscheidend, da diese helfen kann, mögliche Propagation durch größere Pufferzeiten aufzufangen (siehe ebenfalls Abschnitte 11.3 und 11.4).

Drittens wurde zusätzlich gezeigt, dass Stabilität als *Aspekt der Robustheitsbetrachtung* eine flexible Plangestaltung nicht ausschließt. Der Vorteil besteht darin, dass stabile Pläne auch Tauschoperationen beinhalten und gleichzeitig wenige Tauschoperationen während der Ausführung durchzuführen sind. So ist für den Umgang mit alltäglichen Störungen die Stabilität des Ressourceneinsatzes einer rein flexiblen Plangestaltung vorzuziehen (vgl. Abschnitt 11.5).

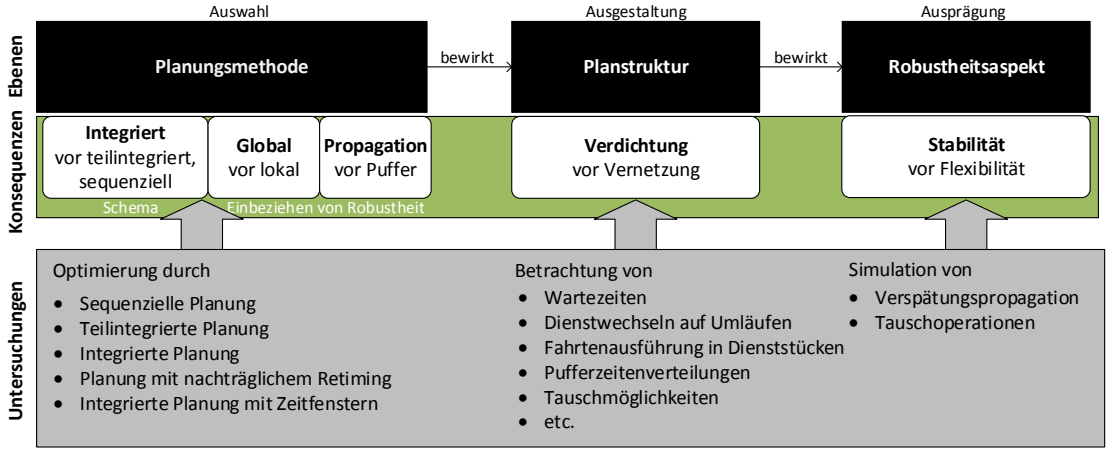

Abbildung 11.21: Konsequenzen zum Erreichen Robuster Effzienz basierend auf den durchgeführten Untersuchungen im Hinblick auf die Dimensionen Planungsmethode, Planstruktur und Robustheitsaspekt

Kapitel 12

Zusammenfassung und Ausblick

In dieser Arbeit wurde untersucht, wie Aufgaben eines mehrschichtigen, voneinander abhängigen Ressourceneinsatzes so eingeplant werden können, dass die entstehenden Ressourceneinsatzpläne gleichzeitig sowohl kosteneffizient als auch robust in ihrer Ausführung sind. Als Anwendungsfall für den proaktiven Umgang mit Unsicherheit bei der Zusammenstellung von Aufgabensequenzen und für den Umgang mit den beiden Zielkriterien Kosteneffizienz und Robustheit wurde die Umlauf- und Dienstplanung im öffentlichen Personenverkehr mit möglichen Verspätungen von Fahrten betrachtet. Ausgehend von Forschungsfragen nach quantitativem Einfluss verschiedener Planungsmethoden, sowie qualitativ abzuleitenden Empfehlungen für den Ressourceneinsatz wurden Optimierungsmodelle und -verfahren entwickelt, die durch die explizite Betrachtung von potenzieller Verspätungspropagation bei gleichzeitiger Berücksichtigung mehrerer Planungsstufen eine Lücke im Stand der Forschung schließen. Sie erlauben es, robust und kosteneffiziente Ressourceneinsatzpläne zu erstellen, und vermeiden so den Nachteil zum Teil bereits im Praxiseinsatz befindlicher traditioneller Optimierungskomponenten, deren Anwendung zu Plänen mit hoher Ressourcenauslastung aber auch Störanfälligkeit führen kann. Die entstandenen Lösungskomponenten wurden zusätzlich genutzt, um Erkenntnisse über das Zusammenspiel von robust effizienten Planstrukturen auf Umlauf- und Dienstebene zu erlangen. Darüber hinaus wurden mit Stabilität und Flexibilität verschiedene Robustheitsaspekte in die Bewertung des Ressourceneinsatzes einbezogen. Insgesamt wurden so die Untersuchungen im Rahmen dieser Arbeit in den drei Dimensionen *Planungsmethodik*, *Planstruktur* und *Robustheitsaspekt* vertieft. Im Folgenden werden zunächst die Beiträge der einzelnen Kapitel mit Blick auf diese drei Dimensionen herausgearbeitet, bevor auf die wesentlichen Erkenntnisse der untersuchten Forschungsfragen, sowie Implikationen für die Praxis und weiterführende Forschungsmöglichkeiten eingegangen wird.

Nach einer Einführung in die Problemstellung in Kapitel 1, wurde in Kapitel 2 die Ressourceneinsatzplanung vor dem Hintergrund der Verknüpfung von Planungs- und Steuerungsaufgaben eines ÖPNV-Betriebes betrachtet. Dabei wurden die Rahmenbedingungen für das Einbeziehen von Robustheitsaspekten in die Ressourcenein-

satzplanung herausgestellt. In Kapitel 3 wurden bestehende Ansätze zur Ressourceneinsatzplanung in Bezug zur Zielsetzung der Arbeit diskutiert. Neben methodischen Grundlagen zum Einbeziehen von Unsicherheit und mehreren Zielkriterien in die Planung wurden sowohl Ansätze zur rein kosteneffizienten Ressourceneinsatzplanung betrachtet, als auch Ansätze aus verschiedenen Problemdomänen des öffentlichen Personenverkehrs, die Robustheitsaspekte in der Planung berücksichtigen. Basierend auf den diskutierten Ansätzen wurde der Forschungsbedarf in Kapitel 4 hergeleitet und das Konzept der Robusten Effizienz als Leitbild für die weitere Arbeit dargestellt. Dabei wurden drei grundlegende Forschungsfragen formuliert.

F1: *Welchen Einfluss haben unterschiedliche Planungsmethoden auf Robustheit und Effizienz der erstellten Pläne?*

F2: *Wieviel Robustheit kann in kostenoptimalen Plänen realisiert werden?*

F3: *Welche Planungsempfehlungen können für den robusten und kosteneffizienten Fahrzeug- und Personaleinsatz hergeleitet werden?*

Gestaltungsorientierte Sicht auf diese Arbeit

Zur Beantwortung der Forschungsfragen wurden die folgenden Schritte unternommen.

Simulationsframework zur Evaluation von Robustheit in Bezug auf Stabilität und Flexibilität: In Kapitel 5 wurden als Grundlage zur Bewertung der Robustheit von Ressourceneinsatzplänen Kennzahlen herausgearbeitet und ein Simulationsframework entwickelt, mit dem geplanter Ressourceneinsatz hinsichtlich seiner Robustheit in der Ausführung evaluiert werden kann. Als wesentliche Kennzahlen werden die Propagationen von Verspätungen abgebildet sowie auf unterschiedlichen Planebenen dimensionierte Tauschmöglichkeiten von Aufgaben und Ressourcen. Während sowohl exakte Verspätungspropagation als auch Tauschmöglichkeiten als potenzielle Dispositionsmaßnahmen in das Simulationsframework eingebettet wurden, wurde zur Bewertung der Robustheit während der Umlauf- und Dienstplanoptimierung eine Propagationsformel vorgeschlagen, die die potenzielle Verspätungspropagation einzelner Ressourceneinsätze bei der Planerstellung erfassen kann.

Stufenweises Einbeziehen der Betrachtung potenzieller Verspätungspropagation in Varianten der Umlauf- und Dienstplanung: Aufbauend auf den herausgearbeiteten Kennzahlen zur Bewertung der Robustheit wurde speziell das direkte Abbilden potenzieller Verspätungspropagation als elementarer Bestandteil der entwickelten

Lösungsansätze zur Umlauf- und Dienstplanung umgesetzt. So wurden unter Einbeziehen der Betrachtung von Verspätungspropagation durch sukzessives Vergrößern der bei der Planung verfügbaren Freiheitsgrade aufeinander aufbauende Lösungsansätze entwickelt. Diese kombinieren in unterschiedlichem Ausmaß Umlauf- und Dienstplanung und erlauben jeweils eine Steuerung des Zielkonflikts zwischen Robustheit und Kosteneffizienz.

Ein Überblick über die Zusammenhänge der entwickelten Lösungsansätze wurde in Kapitel 6 gegeben, indem die genutzten Basisoptimierungsmodelle, sowie die Kombination der Modelle zu Basislösungsschemata für kosteneffiziente und robuste Ressourceneinsatzplanung herausgearbeitet wurden. Dabei wurde die Interaktion der verschiedenen Komponenten in Lösungsschemata zur Umlaufplanung, sowie zur sequenziellen, teilintegrierten und integrierten Umlauf- und Dienstplanung abgebildet.

In Kapitel 7 wurde ein Lösungsansatz zur teilintegrierten Umlauf- und Dienstplanung entwickelt, der Verspätungspropagation auf beiden Ressourcenebenen betrachtet, insbesondere aber auch die Verknüpfung von Umläufen und Diensten steuern kann, um bei gleichbleibenden Plankosten potenzielle Verspätungspropagation zwischen den Ressourcenebenen zu verringern. Dabei wurden drei wesentliche Lösungsprinzipien umgesetzt, die in unterschiedlicher Ausprägung auch in die anderen Lösungsschemata eingebracht wurden: eine Kopplung zwischen flussbasierter und pfadbasierter Formulierung zur Beibehaltung der sich aus der Time-Space-Netzwerk-basierten Modellierung ergebenden Freiheitsgrade bei gleichzeitiger korrekter Abbildung von Verspätungspropagation in den Modellen, eine globale Dekomposition des Lösungsflusses auf den Time-Space-Netzwerken der Umlaufebene unter Einbeziehen potenziell resultierender Verspätungspropagation (im Zusammenspiel mit den erstellten Diensten), sowie ein Abbilden von Propagationseffekten beim Lösen ressourcenbeschränkter Kürzeste-Wege Probleme zum Erstellen von Umläufen und Diensten.

Basierend auf den einzelnen im teilintegrierten Ansatz verwendeten Lösungskomponenten wurde in Kapitel 8 ein integriertes Modell zur robusten Ressourceneinsatzplanung vorgeschlagen. Dieses erlaubt das Berücksichtigen von Verspätungspropagation simultan auf Umlauf- und Dienstebene bei gleichzeitiger Berücksichtigung von Kostenminimierung als Zielsetzung. Die Kopplung von gleichrangigen Planungsaufgaben im Modell und ein im Vergleich zur Literatur erstmals umgesetztes vollständig integriertes Lösungsverfahren zur Planung mit Verspätungspropagation und mehreren Ressourcenebenen ermöglichen es, Verspätungspropagation zu verhindern bzw. in den erstellten Plänen gleichlaufende Propagationseffekte auf Umlauf- und Dienstebene herbeizuführen, die nicht zwischen den Ressourcenebenen weitergegeben werden und die frühzeitig absorbiert werden können.

Der integrierte Lösungsansatz wurde in Kapitel 9 durch das Einbeziehen von Zeitfenstern für Servicefahrten in der Modellierung und im Lösungsverfahren erweitert. Durch die gemeinsame Betrachtung von Fahrplan-, Umlauf- und Dienstebene inklusive potenzieller Verspätungspropagation wird die Anpassung des Fahrplans in Abstimmung mit einer kosteneffizienten Ressourceneinsatzplanung ermöglicht. So können gleichzeitig durch eine leichte Modifikation des Ursprungsproblems sowohl Pufferzeiten für die Ausführung von Servicefahrten geschaffen werden, als auch neue Fahrtenverknüpfungen zur Verringerung des Ressourceneinsatzes entstehen.

Aufgrund der Problemkomplexität wurden verschiedene Techniken zur Reduktion der in den Lösungsansätzen jeweils zu betrachtenden Problemgrößen vorgeschlagen. So werden verspätungstolerante Umläufe im Rahmen der Column Generation Pricing-Probleme durch ein zweiphasiges Pricing-Schema erstellt, bei dem zunächst Teilumläufe und darauf basierend Umläufe generiert werden. Für die integrierte Planung mit Zeitfenstern kann ein problemspezifisches Preprocessing der einzubeziehenden Zeitfenstergrößen in Abhängigkeit potenzieller primärer Verspätungen den Umfang des zugrunde liegenden Netzwerkmodells vor dem Lösen verringern. Darüber hinaus führt die vorgeschlagene Kombination aus Kantengenerierung und Column Generation durch die Einschränkung der zu berücksichtigenden Netzwerkkanten während der integrierten Lösungsverfahren zu einer drastischen Laufzeitreduzierung bei guter Lösungsqualität.

In Kapitel 10 wurde zur Lösung der integrierten Umlauf- und Dienstplanungsprobleme ein hybrider evolutionärer Algorithmus basierend auf dem SPEA2-Verfahren vorgeschlagen, der entwickelte Komponenten aus den vorherigen Kapiteln aufgreift. Im Gegensatz zu den anderen vorgeschlagenen Lösungsansätzen, die eine explizite Steuerung des Zielkonflikts zwischen Robustheit und Kosteneffizienz erlauben, wird mit Hilfe der hybriden Metaheuristik direkt eine breite Menge an Ressourceneinsatzplänen generiert, die die Pareto-optimale Menge der beiden konkurrierenden Zielsetzungen approximiert.

Untersuchung der Ausprägungen Robuster Effizienz durch Variation der Planungsmethode: In Kapitel 11 wurden die entwickelten Lösungsansätze anhand ausführlicher Experimente evaluiert und dazu verwendet, den robusten und kosteneffizienten Ressourceneinsatz anhand aufeinander aufbauender Studien zu diskutieren. Durch Variation der Planungsmethode mit variierender Gewichtung der einzelnen Zielsetzungen wurden unterschiedlich ausgeprägte Ressourceneinsatzpläne erstellt, die einen Trade-Off zwischen Robustheit und Kosteneffizienz abbilden. Dabei wurden *sequenzielle*, *teilintegrierte* und *integrierte Planungsmethoden*, sowie das *nachträgli-*

che und das *integrierte Einbeziehen der Fahrplanungsebene* in die Ressourceneinsatzplanung miteinander verglichen.

Untersuchung der Ausprägungen Robuster Effizienz durch Variation der Simulationsmethodik: Zusätzlich zur Evaluation der Robustheit des Ressourceneinsatzes im Sinne von *Stabilität* wurde in Kapitel 11 die Betrachtung des Ressourceneinsatzes auf Simulationsebene um die Evaluation potenzieller *Flexibilität* erweitert. Mit der zusätzlichen Simulation kostenneutraler dispositiver Maßnahmen wurde eine Wechselwirkung zwischen den Aspekten Stabilität und Flexibilität untersucht.

Analyse erstellter Planstrukturen als zusätzliche Untersuchungsebene: Die Betrachtung der entstehenden Lösungen wurde in Kapitel 11 durch die zusätzliche Analyse der sich ergebenden Planstrukturen um eine weitere Untersuchungsebene vertieft. Zur Identifikation planimmanenter Ursachen für Verspätungspropagation bzw. -absorption wurde die potenzielle Verspätungspropagation durch Variation der in der Simulation einbezogenen Ressourcenebenen auf Dienststück-, Umlauf- sowie kombinierter Umlauf- und Dienstebene betrachtet. Dabei wurde der Bezug einzelner Propagationseffekte zur *Vernetzung der Ressourcenebenen* bzw. zur *Verdichtung von Aufgabensequenzen* untersucht.

Erkenntnisorientierte Sicht auf diese Arbeit

Mit den durchgeführten Untersuchungen können die Forschungsfragen zusammenfassend wie folgt beantwortet werden.

Forschungsfrage F1: Der Trade-Off zwischen Kosteneffizienz (Plankosten) und Robustheit (Stabilität) lässt sich bei der Planung des mehrschichtigen Ressourceneinsatzes steuern, wobei die Auswahl der Planungsmethode einen klaren Einfluss auf die Lösungsqualität hat. Im Vergleich der sequenziellen, teilintegrierten und integrierten Planungsmethoden wurden deutliche Unterschiede in der erreichbaren Lösungsqualität in Bezug auf beide Zielsetzungen nachgewiesen und eine bedingt durch die zur Verfügung stehenden Freiheitsgrade abstufende Dominanz der Planungsmethoden untereinander aufgezeigt (vgl. Abschnitt 11.2).

Durch die gleichzeitige Betrachtung der gegenseitigen Abhängigkeiten von Umlauf- und Dienstplan sowie der gleichzeitigen Berücksichtigung von Verspätungspropagation auf Umlauf- und Dienstebene während der Optimierung erlaubt die integrierte Planung das Erstellen von sowohl robusten als auch hochgradig kosteneffizienten

Plänen. Im Vergleich zu reiner Kostenminimierung nahezu ohne Zusatzkosten, so dass sie den anderen Verfahren überlegen ist.

Das Einbeziehen der Fahrplanungsebene durch kostenneutrale nachträgliche Fahrplananpassungen basierend auf den erstellten Ressourceneinsatzplänen ermöglicht eine zusätzliche Steigerung der Robustheit. Darüber hinaus wurde gezeigt, dass die Ausweitung der Freiheitsgrade durch Einbeziehen von Verschiebeintervallen für Servicefahrten in der integrierten Planung zu einer weiteren Verbesserung führt. Im Vergleich zu nachträglichen Fahrplananpassungen erlaubt die integrierte Planung mit Zeitfenstern eine gleichzeitige Steigerung der Robustheit und Kosteneffizienz. Sie dominiert somit die anderen Planungsmethoden (vgl. Abschnitt 11.4).

Als Resultat kann mit den vorgeschlagenen Lösungsansätzen die Verspätungspropagation im Ressourceneinsatz minimiert werden unter gleichzeitiger Beibehaltung von Kosteneffizienz. Dabei führt das Abbilden von Propagationseffekten zu einer passenden Verteilung von Pufferzeiten sowie einer passenden Verknüpfung von Fahrten bzw. Aufgaben, um Verspätungen zu absorbieren, und hilft damit eine Robuste Effizienz des eingeplanten Ressourceneinsatzes zu erreichen.

Forschungsfrage F2: Im Vergleich zu rein kostenminimierender Planung kann eine höhere Robustheit kostenneutral erreicht werden. Es wurde gezeigt, dass kostengleiche und aufgabenidentische Ressourceneinsatzpläne existieren, bei denen eine geschickte andere Zuordnung von Aufgabensequenzen bzw. Dienststücken und Diensten zu Umläufen zur deutlichen Verringerung der Verspätungspropagation führt. So können auch kostenoptimale bzw. sehr kosteneffiziente Ressourceneinsatzpläne robuster im Sinne von Stabilität gestaltet werden, sofern Verspätungspropagation bei der Planung berücksichtigt wird (vgl. Abschnitt 11.2).

Darüber hinaus kann die Robustheit durch die Ausweitung der Planungsaufgabe um eine nachträgliche für den Ressourceneinsatz kostenneutrale Fahrplananpassung merklich erhöht werden (vgl. Abschnitt 11.4). Zusätzlich hat die Ausweitung des Robustheitsbegriffes um Flexibilität Einfluss auf die Bewertung der Robustheit. Es wurde gezeigt, dass auch in kostenoptimalen bzw. sehr kosteneffizienten Ressourceneinsatzplänen kostenneutrale Tauschmöglichkeiten existieren, so dass durch Eingriffe bei der Planausführung Verspätungspropagation reduziert werden kann (vgl. Abschnitt 11.5).

Forschungsfrage F3: In der Analyse der verschiedenartig erstellten Umlauf- und Dienstpläne wurde herausgearbeitet, dass die Verdichtung von Aufgabensequenzen in Dienststücken als wesentlicher Faktor für Verspätungspropagation vor einer Ver-

netzung der Ressourcenebenen durch eingeplante Fahrzeug- bzw. Fahrerwechsel zu beachten ist. Für die gesamte Minimierung von Verspätungspropagation ist jedoch das Zusammenspiel mit einer passenden Vernetzung entscheidend, da diese helfen kann, mögliche Propagation durch größere Pufferzeiten aufzufangen. So kann auch in hochgradig kosteneffizienten und (stabil) robusten Plänen Verspätungspropagation größtenteils in einzelnen Dienststücken isoliert und zwischen verschiedenen Ressourcen und unterschiedlichen Ressourcenebenen vermieden werden (vgl. Abschnitt 11.3, sowie ergänzend Abschnitt 11.4).

Mit Blick auf die Umlauf- und Dienstplangestaltung wurden überdies zwei Indikatoren zur einfachen Bewertung von Robustheit des kombinierten Ressourceneinsatzes herausgearbeitet, der *AP-Indikator* und der *DP-Indikator*. Beide Kennzahlen setzen Aufgaben- bzw. Fahrzeugwechsel in einen Zusammenhang mit zur Verfügung stehenden Pufferzeiten und bilden so die Bewertung der Verdichtung bzw. Vernetzung innerhalb der Pläne ab (vgl. Abschnitt 11.3).

Darüber hinaus konnte gezeigt werden, dass eine stabile Plangestaltung einen flexiblen Ressourceneinsatz nicht ausschließt und daher im Umgang mit alltäglichen Störungen vorzuziehen ist. Auch stabile Pläne beinhalten Freiheitsgrade zum Tausch von Ressourcen bzw. Aufgaben, gleichzeitig brauchen während der Planausführung nur wenige Tauschoperationen durchgeführt werden. Vielmehr können stabile Pläne durch geringe Eingriffe zum Teil noch robuster ausgeführt werden (vgl. Abschnitt 11.5).

Implikationen für die Praxis

Die entwickelten Optimierungsmodelle und -methoden erweitern den aktuellen Stand der Technik von Optimierungskomponenten für Planungssysteme. Mit dem Einbeziehen von Robustheitsaspekten in die Planungsverfahren wird ein entscheidendes praxisrelevantes Qualitätskriterium bei der Erstellung von Ressourceneinsatzplänen abgebildet – bei gleichzeitiger Berücksichtigung der traditionellen Forderung nach kosteneffizienten Plänen. So können die in dieser Arbeit entwickelten Artefakte dazu beitragen, die Akzeptanz der durch mathematische Optimierung zur Entscheidungsunterstützung vorgeschlagenen Ressourceneinsatzpläne im Planungsprozess zu erhöhen und somit zu einer Vereinfachung und Beschleunigung des komplexen Planungsprozesses führen.

Das Einbinden von (historischen) Verspätungsdaten in die Lösungsansätze ermöglicht es darüber hinaus, eine enge Rückkopplung zwischen Ausführung und Planung des Ressourceneinsatzes zu etablieren. Informationen über aktuelle Verspätungstrends könnten verstärkend bei der Ressourceneinsatzplanung berücksichtigt werden,

so dass durch eine kontinuierliche Anpassung an sich stark ändernde Verspätungssituationen die Leistungsfähigkeit des ÖPV-Systems beibehalten bzw. langfristig verbessert wird.

Die vorgestellten Ergebnisse zeigen deutlich, dass zur Planungsunterstützung eine kombinierte Optimierung von Umlauf- und Dienstplänen anderen Planungsvarianten vorzuziehen ist. Der eindeutig höhere Laufzeitaufwand der integrierten Planungsvariante rentiert sich aufgrund der erreichbaren Lösungsqualität in Bezug auf die beiden Zielsetzungen Robustheit und Kosteneffizienz.

Die robusten und kosteneffizienten Pläne können so erstellt werden, dass bei alltäglichen Störungen kaum in die Ausführung eingegriffen werden muss. Gleichzeitig beinhalten diese Pläne auch kostenneutrale Tauschmöglichkeiten. Um die Anzahl benötigter dispositiver Maßnahmen während der Planausführung weitestgehend einzuschränken, bietet sich vor diesem Hintergrund eine stabile vor einer ausdrücklich flexiblen Planunggestaltung an. Darüber hinaus ist durch die Anwendung der Verfahren zur integrierten Planung mit Zeitfenstern die Identifikation von Fahrtenmengen möglich, deren gemeinsame, geringfügige zeitliche Anpassung Potenzial zur Erhöhung von sowohl Robustheit als auch Kosteneffizienz des Ressourceneinsatzes bieten kann.

Zusätzlich bieten die herausgearbeiteten Kennzahlen zur Beschreibung des Verdichtungs- bzw. Vernetzungsgrades von Umlauf- und Dienstplänen einen Ansatzpunkt, um die potenzielle Robustheit (im Sinne von Stabilität) eines kombinierten Ressourceneinsatzes bewerten und einschätzen zu können. Sie sind ohne Rückgriff auf Simulationsstudien einfach anwendbar für den Vergleich von Ressourceneinsatzplänen.

Anknüpfende Forschungsmöglichkeiten

Die Arbeit bietet in vielfältiger Weise Ansatzpunkte für weitere Forschungsmöglichkeiten. Die Wesentlichen werden im Folgenden in Bezug auf unterschiedliche Betrachtungsebenen aufgegliedert.

In den vorgeschlagenen Lösungsansätzen wurden durch Einbeziehen von Verschiebeintervallen für Fahrten Teilentscheidungen der Fahrplanung mit einer integrierten Umlauf- und Dienstplanung kombiniert. Ein konsequenter nächster Schritt besteht nun darin, die in dieser Arbeit adressierte, produktionsorientierte Sichtweise der Planung noch stärker mit angebotsorientierten Aspekten zu verknüpfen. Im Rahmen der Angebotsplanung beschäftigen sich beispielsweise [Goerigk et al., 2013] und [Goerigk et al., 2014] mit dem Berücksichtigen von Robustheitsaspekten aus Passagiersicht bei der Linien- bzw. Fahrplanerstellung (im Bahnverkehr). Die Verbindung solcher angebotsorientierten Forschungsstränge mit denen bereits in dieser Arbeit gebündelten Untersuchungen bietet Möglichkeiten, eine übergreifende Optimierung von weiten

Teilen des ÖPV-Systems zu untersuchen. Durch das Einbeziehen von Informationen über z.B. Passagierströme könnten potenzielle Auswirkungen von Verspätungen in den Optimierungsansätzen der Ressourceneinsatzplanung zusätzlich bewertet werden. Lösungen solcher erweiterten Modelle und Verfahren könnten dazu beitragen, die Kundenzufriedenheit bei Nutzung eines gleichzeitig hoch effizienten ÖPV-Systems sicherzustellen.

Überdies bietet sich für weitere Untersuchungen eine Ausweitung der Evaluation des Ressourceneinsatzes an. Diese könnte auch im Hinblick auf Kundenzufriedenheit und mögliches Kundenverhalten erfolgen. Dafür sind Entscheidungen über dispositive Maßnahmen wie Anschlusssicherung oder Fahrtenausfälle in der Simulation des Ressourceneinsatzes abzubilden sowie individuelle Passagiere bzw. Passagiergruppen zu simulieren. In ersten Tests hat sich die Entwicklung einer passgenauen agentenbasierten Simulation aufbauend auf dem im Rahmen dieser Arbeit geschaffenen Simulationsframework für eine solche Untersuchung als brauchbar erwiesen. Diese kann – eine mesoskopische Sicht verfolgend – individuelles Passagierverhalten einbeziehen, ohne dabei den tiefen Detaillierungsgrad agentenbasierter Simulationssysteme wie MATSim zu erfordern.

In dieser Arbeit wurde u.a. gezeigt, dass (kosteneffizient und) stabile Pläne gleichzeitig so erstellt werden können, dass sie bei Bedarf flexibel in der Ausführung sind und kostenneutrale Tauschoperationen zur zusätzlichen Absorption von Verspätungen ermöglichen. Ergänzend könnte der Aspekt der Flexibilität des Ressourceneinsatzes explizit als Zielsetzung in der Planung berücksichtigt werden, um das erreichbare Ausmaß möglicher Flexibilität und einen etwaigen Einfluss auf Stabilität vor dem Hintergrund kosteneffizienter Planung zu untersuchen.

Darüber hinaus ist eine Überführung der Lösungsansätze auf Anwendungsfelder außerhalb der ÖPNV-Domäne vorstellbar. Ein Beispiel ist die Unterstützung der Touren- und Fahrereinsatzplanung in Logistiknetzwerken, in denen regelmäßig (zeitkritische) Transportaufgaben koordiniert werden. Dabei bedarf es eines kosteneffizienten und robusten Ressourceneinsatzes z.B. für Transporte in Netzen mit einer sehr engen Taktung bei Vor- und Hauptlauf oder für Begegnungsverkehre zum Tausch der Ladung.

Literaturverzeichnis

[Abbink et al., 2005] Abbink, E., Fischetti, M., Kroon, L., Timmer, G., und Vromans, M. (2005). Reinventing crew scheduling at netherlands railways. *Interfaces*, 35(5):393–401.

[Ahmadbeygi et al., 2010] Ahmadbeygi, S., Cohn, A., und Lapp, M. (2010). Decreasing airline delay propagation by re-allocating scheduled slack. *IIE Transactions*, 42(7):478–489.

[Ahuja et al., 2005] Ahuja, R. K., Liu, J., Orlin, J. B., Sharma, D., und Shughart, L. A. (2005). Solving real-life locomotive-scheduling problems. *Transportation Science*, 39(4):503–517.

[Ahuja et al., 1993] Ahuja, R. K., Magnanti, T. L., und Orlin, J. B. (1993). *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, New Jersey.

[Amberg, 2014] Amberg, B. (2014). *Ressourceneinsatzplanung im öffentlichen Personennahverkehr: Ähnlichkeitsaspekte in der fahrplanübergreifenden Optimierung*. PhD thesis, Universität Paderborn.

[Amberg et al., 2011] Amberg, B., Amberg, B., und Kliewer, N. (2011). Approaches for increasing the similarity of resource schedules in public transport. *Proceedings of the 14th Euro Working Group on Transportation (EWGT) Meeting - Procedia Social and Behavioral Sciences*, 20:836–845.

[Arikan et al., 2013] Arikan, M., Deshpande, V., und Sohoni, M. (2013). Building reliable air-travel infrastructure using empirical data and stochastic models of airline networks. *Operations Research*, 61(1):45–64.

[Atkinson et al., 2012] Atkinson, S. E., Ramdas, K., und Williams, J. W. (2012). An empirical investigation into the impact of robust scheduling practices in the u.s. airline industry. *Working Paper Series*. Available at SSRN: http://ssrn.com/abstract=2201648 or http://dx.doi.org/10.2139/ssrn.2201648.

[Balas und Padberg, 1976] Balas, E. und Padberg, M. W. (1976). Set partitioning: A survey. *SIAM Review*, 18(4):710–760.

[Ball et al., 1983] Ball, M., Bodin, L., und Dial, R. (1983). A matching based heuristic for scheduling mass transit crews and vehicles. *Transportation Science*, 17(1):4–31.

[Beasly und Chu, 1996] Beasly, J. und Chu, P. (1996). A genetic algorithm for the set covering problem. *European Journal of Operational Research*, 94(2):392–404.

[Ben-Tal et al., 2009] Ben-Tal, A., El Ghaoui, L., und Nemirovski, A. (2009). *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press.

[Benchimol et al., 2012] Benchimol, P., Desaulniers, G., und Desrosiers, J. (2012). Stabilized dynamic constraint aggregation for solving set partitioning problems. *European Journal of Operational Research*, 223:360–371.

[Berliner Verkehrsbetriebe AöR, 2017] Berliner Verkehrsbetriebe AöR (2017). Zahlenspiegel 2016. Verfügbar unter http://unternehmen.bvg.de/de/unternehmen/medien/publikationen (letzter Abruf 30.3.2017).

[Bertossi et al., 1987] Bertossi, A. A., Carraresi, P., und Gallo, G. (1987). On some matching problems arising in vehicle scheduling models. *Networks*, 17:271–281.

[Birge und Louveaux, 1997] Birge, J. R. und Louveaux, F. (1997). *Introduction to stochastic programming*. Springer series in operations research. Springer, New York.

[Borndörfer et al., 2008] Borndörfer, R., Löbel, A., und Weider, S. (2008). A bundle method for integrated multi-depot vehicle and duty scheduling in public transit. In Hickmann, M., Mirchandani, P., und Voß, S., editors, *Computer-Aided Scheduling of Public Transport*, Volume 600, Seiten 3–24.

[Borndörfer et al., 2002] Borndörfer, R., Löbel, A., und Weider, S. (2002). Integrierte Umlauf- und Dienstplanung im Öffentlichen Nahverkehr. *Technischer Bericht ZIB Report 02-10*.

[Borndörfer et al., 2004] Borndörfer, R., Löbel, A., und Weider, S. (2004). A bundle method for integrated multi-depot vehicle and duty scheduling in public transit. *Technischer Bericht ZIB Report 04-14*.

[Branke et al., 2008] Branke, J., Meisel, S., und Schmidt, C. (2008). Simulated annealing in the presence of noise. *Journal of Heuristics*, 14(6):627–654.

[Bundesministerium für Verkehr, Bau- und Wohnungswesen, 2005] Bundesministerium für Verkehr, Bau- und Wohnungswesen (2005). Verordnung zur Durchführung des Fahrpersonalgesetzes (Fahrpersonalverordnung - FPersV). Letzte Änderung durch Artikel 1 der Verordnung vom 22. Mai 2013 (BGBl. S. 1395).

[Bunte, 2009] Bunte, S. (2009). *Lösungen für Anwendungsfälle der Fahrzeugeinsatzplanung im öffentlichen Personennahverkehr*. PhD thesis, Universität Paderborn.

[Bunte und Kliewer, 2009] Bunte, S. und Kliewer, N. (2009). An overview on vehicle scheduling models. *Public Transport*, 1(4):299–317.

[Bunte et al., 2006] Bunte, S., Kliewer, N., und Suhl, L. (2006). Mehrdepot-Umlaufplanung: Berücksichtigung von Verschiebeintervallen für Fahrten in einem Time-Space-Netzwerk-basierten Modell. *Operations Research Proceedings 2005: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), Bremen, September 7–9, 2005*, Seiten 49–54.

[Burke et al., 2010] Burke, E. K., de Gausemaecker, P., de Maere, G., Mulder, J., Paelinck, M., und Berghe, G. V. (2010). A multi-objective approach for robust airline scheduling. *Computers & Operations Research*, 37(5):822–832.

[Cacchiani et al., 2012] Cacchiani, V., Caprara, A., Galli, L., Kroon, L., Maroti, G., und Toth, P. (2012). Railway rolling stock planning: Robustness against large disruptions. *Transportation Science*, 46(2):213–232.

[Cacchiani et al., 2014] Cacchiani, V., Huisman, D., Kidd, M., Kroon, L., Toth, P., Veelenturf, L., und Wagenaar, J. (2014). An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B*, 63:15–37.

[Cacchiani und Toth, 2012] Cacchiani, V. und Toth, P. (2012). Nominal and robust train timetabling problems. *European Journal of Operational Research*, 219:727–737.

[Cadarso und Marin, 2011] Cadarso, L. und Marin, A. (2011). Robust rolling stock in rapid transit networks. *Computers & Operations Research*, 38(8):1131–1142.

[Cadarso und Marin, 2012] Cadarso, L. und Marin, A. (2012). Integration of timetable planning and rolling stock in rapid transit networks. *Annals of Operations Research*, 199(2):113–135.

[Caprara et al., 2007] Caprara, A., Kroon, L., Monaci, M., Peeters, M., und Toth, P. (2007). Chapter 3 - Passenger Railway Optimization. In Barnhart, C. und Laporte, G., editors, *Transportation*, Volume 14 of *Handbooks in Operations Reserach & Management Science*, Kapitel 3, Seiten 69–127. Elsevier.

[Ceder, 2007] Ceder, A. (2007). *Public Transit Planning and Operation - Theory, Modelling and Practice*. Elsevier, Butterworth-Heinemann, Oxford.

[Center Nahverkehr Berlin, 2015] Center Nahverkehr Berlin (2015). *Qualität im Berliner ÖPNV - Jahresbericht 2013*. Jahresberichte. Januar 2015, Berlin.

[Center Nahverkehr Berlin, 2017] Center Nahverkehr Berlin (2017). Verkehrsvertrag - Qualitätsvorgaben. Verfügbar unter http://www.cnb-online.de/bvg-verkehrsvertrag/qualitaetsvorgaben/ (letzter Abruf 30.3.2017).

[Chvatal, 1983] Chvatal, V. (1983). *Linear Programming*. Series of books in the mathematical sciences. W. H. Freeman, Wiesbaden.

[Clausen et al., 2010] Clausen, J., Larsen, A., Larsen, J., und Rezanova, N. J. (2010). Disruption management in the airline industry - concepts, models and methods. *Computers & Operations Research*, 37(5):809–821.

[Daduna und Paixão, 1995] Daduna, J. R. und Paixão, J. M. P. (1995). Vehicle scheduling for public mass transit - an overview. *Proceedings of the Sixth International Workshop on Computer-Aided Scheduling of Public Transport*, 430:76–90.

[DeAlmeida et al., 2008] DeAlmeida, D., Marcos, N., Vianey, S., und Weber, C. (2008). Robustness of rolling stock schedules. *Technical Report ARRIVAL-TR-0129*. SNCF - Innovation and Research Department, Paris, France.

[Dell'Amico et al., 1993] Dell'Amico, M., Fischetti, M., und Toth, P. (1993). Heuristic algorithms for the multiple depot vehicle scheduling problem. *Management Science*, 39(1):115–125.

[Desaulniers et al., 2002] Desaulniers, G., Desrosiers, J., und Solomon, M. (2002). *Essays and Surveys in Metaheuristics*, Kapitel Accelerating Strategies in Column Generation Methods for Vehicle Routing and Crew Scheduling Problems, Seiten 309–324. Springer US, Boston, MA.

[Desaulniers und Hickman, 2007] Desaulniers, G. und Hickman, M. D. (2007). Chapter 2 - public transit. In Barnhart, C. und Laporte, G., editors, *Transportation*,

Volume 14 of *Handbooks in Operations Reserach & Management Science*, Kapitel 2, Seiten 69–127. Elsevier.

[Desrosiers et al., 1995] Desrosiers, J., Dumas, Y., Solomon, M. M., und Soumis, F. (1995). Time constrained routing and scheduling. *Handbooks in Operations Research and Management Science*, 8:35–139.

[Dréo et al., 2006] Dréo, J., Pétrowski, A., Siarry, P., und Taillard, E. (2006). *Metaheuristics for Hard Optimization: Methods and Case Studies*. Springer, Berlin Heidelberg.

[Dueck, 2010] Dueck, V. (2010). *Increasing Stability of Aircraft and Crew Schedules*. PhD thesis, Universität Paderborn.

[Dueck et al., 2012] Dueck, V., Ionescu, L., Kliewer, N., und Suhl, L. (2012). Increasing stability of crew and aircraft schedules. *Transportation Research Part C*, 20(1):47–61.

[Dunbar et al., 2012] Dunbar, M., Froyland, G., und Wu, C. L. (2012). Robust airline schedule planning: Minimizing propagated delay in an integrated routing and crewing framework. *Transportation Science*, 46(2):204–216.

[Dunbar et al., 2014] Dunbar, M., Froyland, G., und Wu, C. L. (2014). An integrated scenario-based approach for robust aircraft routing, crew pairing and re-timing. *Computers & Operations Research*, 45:68–86.

[Ehrgott, 2005] Ehrgott, M. (2005). *Multicriteria Optimization*, Volume 491 of *Lectures notes in economics and mathematical systems*. Springer.

[Ehrgott und Ryan, 2002] Ehrgott, M. und Ryan, D. M. (2002). Constructing robust crew schedules with bicreteria optimization. *Journal of Multi-Criteria Decision Analysis*, 11(3):139–150.

[Falkner und Ryan, 1992] Falkner, J. und Ryan, D. M. (1992). Express: Set partitioning for bus crew scheduling in christchurch. *Computer-Aided Scheduling*, 386:359–378.

[Feo und Resende, 1995] Feo, T. und Resende, M. (1995). Greedy randomized adaptive search procedures. *Journal of Global Optimization*, 6:109–133.

[Fioole et al., 2006] Fioole, P., Kroon, L., Maroti, G., und Schrijver, A. (2006). A rolling stock circulation model for combining and splitting of passenger trains. *European Journal of Operational Research*, 172(2):1281–1297.

[Fischetti et al., 1987] Fischetti, M., Lodi, A., Martello, S., und Toth, P. (1987). The fixed job schedule problem with spread-time constraints. *Operations Research*, 35(6):849–858.

[Fischetti et al., 1989] Fischetti, M., Lodi, A., Martello, S., und Toth, P. (1989). The fixed job schedule problem with working-time constraints. *Operations Research*, 37(3):395–403.

[Freling, 1997] Freling, R. (1997). *Models and Techniques for Integrating Vehicle and Crew Scheduling*. PhD thesis, Erasmus University of Rotterdam.

[Friberg und Haase, 1999] Friberg, C. und Haase, K. (1999). An exact branch and cut algorithm for the vehicle and crew scheduling problem. *Computer-Aided Transit Scheduling*, 471:63–80.

[Gabrel et al., 2014] Gabrel, V., Murat, C., und Thiele, A. (2014). Recent advances in robust optimization: An overview. *European Journal of Operational Research*, 235(3):471 – 483.

[Gaffi und Nonato, 1999] Gaffi, A. und Nonato, M. (1999). An integrated approach to ex-urban crew and vehicle scheduling. *Computer-Aided Transit Scheduling*, 471:103–128.

[Gandibleux et al., 2004] Gandibleux, X., Sevaux, M., Sörensen, K., und T'kindt, V., editors (2004). *Metaheuristics for Multiobjective Optimisation*, Volume 535 of *Lectures notes in economics and mathematical systems*. Springer.

[Gentle, 2003] Gentle, J. E. (2003). *Random Number Generation and Monte Carlo Methods*. Springer-Verlag, New York, 2. Edition.

[Gintner, 2008] Gintner, V. (2008). *Modelle und Lösungsverfahren für die integrierte Ressourceneinsatzplanung im öffentlichen Personennahverkehr*. PhD thesis, Universität Paderborn.

[GIRO Inc., 2017] GIRO Inc. (2017). HASTUS. Verfügbar unter http://www.giro.ca/de/loesungen/bus-u-und-s-bahn-strassenbahn (letzter Abruf 30.3.2017).

[Goerigk et al., 2013] Goerigk, M., Schachtebeck, M., und Schöbel, A. (2013). Evaluating line concepts using travel times and robustness. *Public Transport*, 5(3):267–284.

[Goerigk et al., 2014] Goerigk, M., Schmidt, M., Schöbel, A., Knoth, M., und Müller-Hannemann, M. (2014). The price of strict and light robustness in timetable information. *Transportation Science*, 48(2):225–242.

[Goldberg, 1989] Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1. Edition.

[Haase et al., 2001] Haase, K., Desaulniers, G., und Desrosiers, J. (2001). Simultaneous vehicle and crew scheduling in urban mass transit systems. *Transportation Science*, 35(1):286–303.

[Hamburger Hochbahn AG, 2017] Hamburger Hochbahn AG (2017). Zahlen-Daten-Fakten. Verfügbar unter https://www.hochbahn.de/hochbahn/hamburg/de/ home/unternehmen/unser_job_fuer_hamburg/zahlen-+daten-+fakten (letzter Abruf 30.3.2017).

[Hammersley und Handscomb, 1964] Hammersley, J. M. und Handscomb, D. C. (1964). *Monte Carlo Methods*. Springer.

[Hevner et al., 2004] Hevner, A. R., March, S. T., Park, J., und Ram, S. (2004). Design science in informations systems research. *MIS Quarterly*, 28(1):75–105.

[Horni et al., 2016] Horni, A., Nagel, K., und Axhausen, K. W., editors (2016). *The Multi-Agent Transport Simulation MATSim*. Ubiquity Press, London, 1. Edition.

[Huisman, 2004] Huisman, D. (2004). *Integrated and Dynamic Vehicle and Crew Scheduling*. PhD thesis, Erasmus University of Rotterdam.

[Huisman et al., 2004] Huisman, D., Freling, R., und Wagelmans, A. P. (2004). A robust solution approach to the dynamic vehicle scheduling problem. *Transportation Science*, 38(4):447–458.

[Huisman et al., 2005] Huisman, D., Freling, R., und Wagelmans, A. P. (2005). Multiple-depot integrated vehicle and crew scheduling. *Transportation Science*, 39(4):491–502.

[Huisman und Wagelmans, 2006] Huisman, D. und Wagelmans, A. P. (2006). A solution approach for dynamic vehicle and crew scheduling. *European Journal of Operational Research*, 172:453–471.

[Ibarra-Rojas et al., 2015] Ibarra-Rojas, O. J., Delgado, F., Giesen, R., und Munoz, J. C. (2015). Planning, operation, and control of bus transport systems: A literature review. *Transportation Research Part B*, 77:38–75.

[Ide und Schöbel, 2015] Ide, J. und Schöbel, A. (2015). Robustness for uncertain multi-objective optimization: a survey and analysis of different concepts. *OR Spectrum*, 38(1):235–271.

[INIT innovation in traffic systems AG, 2017] INIT innovation in traffic systems AG (2017). MOBILE. Verfügbar unter http://www.init-ka.de/de/produkte/produkte.php (letzter Abruf 30.3.2017).

[Ionescu und Kliewer, 2011] Ionescu, L. und Kliewer, N. (2011). Increasing flexibility of airline schedules. *Proceedings of the 14th Euro Working Group on Transportation (EWGT) Meeting - Procedia Social and Behavioral Sciences*, 20:1019–1028.

[Irnich und Desaulniers, 2004] Irnich, S. und Desaulniers, G. (2004). Shortest path problems with ressource constraints. Technical report, Les Cahiers du GERAD, HEC Montréal.

[IVU Traffic Technologies AG, 2017] IVU Traffic Technologies AG (2017). IVU.suite. Verfügbar unter http://www.ivu.de/produkte-loesungen/ivusuite.html (letzter Abruf 30.3.2017).

[Jespersen-Groth et al., 2009] Jespersen-Groth, J., Potthoff, D., Clausen, J., Huisman, D., Kroon, L., Maroti, G., und Nielsen, M. N. (2009). Disruption management in passenger railway transportation. In Ahuja, R. K., Moehring, R. H., und Zaroliagis, C. D., editors, *Robust and Online Large-Scale Optimization*, Volume 5868 of *Lecture Notes in Computer Science*, Seiten 399–421. Springer Berlin Heidelberg.

[Kliewer, 2005] Kliewer, N. (2005). *Optimierung des Fahrzeugeinsatzes im öffentlichen Personennahverkehr - Modelle, Methoden und praktische Anwendungen.* PhD thesis, Universität Paderborn.

[Kliewer et al., 2012] Kliewer, N., Amberg, B., und Amberg, B. (2012). Multiple depot vehicle and crew scheduling with time windows for scheduled trips. *Public Transport*, 3(3):213–244.

[Kliewer et al., 2008] Kliewer, N., Gintner, V., und Suhl, L. (2008). Line change considerations within a time-space network based multi-depot bus scheduling model. In Hickmann, M., Mirchandani, P., und Voß, S., editors, *Computer-Aided Scheduling of Public Transport*, Volume 600, Seiten 57–70.

[Kliewer et al., 2006] Kliewer, N., Mellouli, T., und Suhl, L. (2006). A time-space network based exact optimization model for multi-depot bus scheduling. *European Journal of Operational Research*, 175(3):1616–1627.

[Kramkowski et al., 2009] Kramkowski, S., Kliewer, N., und Meier, C. (2009). Heuristic methods for increasing delay-tolerance of vehicle schedules in public bus transport. *MIC 2009 Proceedings*. Metaheuristic International Conference VIII, Hamburg, Germany.

[Kroon und Fischetti, 2000] Kroon, L. und Fischetti, M. (2000). Crew Scheduling for Netherland Railways - Destination: Customer. *Working Paper ERS-2000-56-LIS*. Erasmus Universität Rotterdam, Rotterdam, The Netherlands.

[Lan et al., 2006] Lan, S., Clarke, J. P., und Barnhart, C. (2006). Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions. *Transportation Science*, 40(1):15–28.

[Laumanns et al., 2002] Laumanns, M., Thiele, L., Deb, K., und Zitzler, E. (2002). Combining convergence and diversity in evolutionary multiobjective optimization. *Evolutionary computation*, 10(3):263–282.

[Law, 2014] Law, A. M. (2014). *Simulation Modeling and Analysis*. McGraw-Hill, New York, 5. Edition.

[Löbel, 1997] Löbel, A. (1997). *Optimal Vehicle Scheduling in Public Transit*. PhD thesis, Technische Universität Berlin.

[Löbel, 1998] Löbel, A. (1998). Vehicle scheduling in public transit and lagrangean pricing. *Management Science*, 44(12):1637–1649.

[Liebchen et al., 2009] Liebchen, C., Lübbecke, M. E., Möhring, R. H., und Stiller, S. (2009). The concept of recoverable robustness, linear programming recovery, and railway applications. In Ahuja, R. K., Möhring, R. H., und Zaroliagis, C. D., editors, *Robust and Online Large-Scale Optimization*, Volume 5868 of *Lecture Notes in Computer Science*, Seiten 1–27. Springer.

[Lourenço et al., 2001] Lourenço, H., Paixão, J., und Portugal, R. (2001). Multiobjective metaheuristics for the bus driver scheduling problem. *Transportation Science*, 35(3):331–343.

[Macal und North, 2010] Macal, C. M. und North, M. J. (2010). Tutorial on agent-based modelling and simulation. *Journal of Simulation*, 4(3):151 – 162.

[Maher, 2016] Maher, S. J. (2016). Solving the integrated airline recovery problem using column-and-row generation. *Transportation Science*, 50(1):216–239.

[MAIOR srl., 2017] MAIOR srl. (2017). MTRAM. Verfügbar unter http://www.maior.it/soluzioni/tpl (letzter Abruf 30.3.2017).

[Münchner Verkehrsgesellschaft mbH, 2017] Münchner Verkehrsgesellschaft mbH (2017). MVG Pünktlichkeitswerte. Verfügbar unter https://www.mvg.de/services/fahrgastservice/mvg-puenktlichkeitswerte.html (letzter Abruf 30.3.2017).

[Moudani et al., 2001] Moudani, W. E., Cosenza, C. A. N., Coligny, M. d., und Mora-Camino, F. (2001). A bi-criterion approach for the airlines crew rostering problem. In *Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization*, EMO '01, Seiten 486–500, London, UK, UK. Springer-Verlag.

[Moz et al., 2009] Moz, M., Respício, A., und Pato, M. (2009). Bi-objective evolutionary heuristics for bus driver rostering. *Journal of Public Transport*, 1(3):189–210.

[Naumann et al., 2011] Naumann, M., Suhl, L., und Kramkowski, S. (2011). A stochastic programming approach for robust vehicle scheduling in public bus transport. *Proceedings of the 14th Euro Working Group on Transportation (EWGT) Meeting - Procedia Social and Behavioral Sciences*, 20:826–835.

[Nemhauser und Wolsey, 1988] Nemhauser, G. und Wolsey, L. (1988). *Integer and combinatorial optimization*. Wiley-Interscience series in discrete mathematics and optimization. Wiley, New York.

[O'Keefe, 2014] O'Keefe, R. (2014). Design science, the design of systems and Operational Research: back to the future? *Journal of the Operational Research Society*, 65(5):673–684.

[Oliver et al., 1987] Oliver, I. M., Smith, D. J., und Holland, J. R. C. (1987). A study of permutation crossover operators on the traveling salesman problem. In *Proceedings of the Second International Conference on Genetic Algorithms on Genetic algorithms and their application*, Seiten 224–230. L. E. Associates, Inc.

[Patrikalakis und Xerocostas, 1992] Patrikalakis, I. und Xerocostas, D. (1992). A new decomposition scheme of the urban public transport scheduling problem. *Computer-Aided Transit Scheduling: Proceedings of the Fifth International Workshop*, Seiten 407–425.

[Pepin et al., 2009] Pepin, A.-S., Desaulniers, G., Hertz, A., und Huisman, D. (2009). A comparison of five heuristics for the multiple depot vehicle scheduling problem. *Journal of Scheduling*, 12(1):17–30.

[Petersen et al., 2012] Petersen, J., Sölveling, G., Clark, J. P., Johnson, E., und Shebalov, S. (2012). An optimization approach to airline integrated recovery. *Transportation Science*, 46(4):482–500.

[Scholl, 2001] Scholl, A. (2001). *Robuste Planung und Optimierung: Grundlagen - Konzepte und Methoden - Experimentelle Untersuchungen*, Kapitel Robuste Planung, Seiten 89–172. Physica-Verlag, Heidelberg.

[Shebalov und Klabjan, 2006] Shebalov, S. und Klabjan, D. (2006). Robust airline crew pairing: Move-up crews. *Transportation Science*, 40(3):300–312.

[Shukla und Deb, 2007] Shukla, P. K. und Deb, K. (2007). On finding multiple pareto-optimal solutions using classical and evolutionary generating methods. *European Journal of Operational Research*, 181(3):1630 – 1652.

[Siebers et al., 2010] Siebers, P. O., Macal, C. M., Garnett, J., Buxton, D., und Pidd, M. (2010). Discrete-event simulation is dead, long live agent-based simulation! *Journal of Simulation*, 4(3):204 – 210.

[Soyster, 1973] Soyster, A. L. (1973). Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):1154–1157.

[Statistisches Bundesamt, 2016] Statistisches Bundesamt (2016). *Verkehr aktuell - 05/ 2016*. Fachserie 8 Reihe 1.1. Wiesbaden.

[Steinzen, 2007] Steinzen, I. (2007). *Topics in Integrated Vehicle and Crew Scheduling in Public Transport*. PhD thesis, Universität Paderborn.

[Steinzen et al., 2010] Steinzen, I., Gintner, V., Suhl, L., und Kliewer, N. (2010). A time-space network approach for the integrated vehicle- and crew-scheduling problem with multiple depots. *Transportation Science*, 44(3):367–382.

[Österle et al., 2010] Österle, H., Becker, J., Frank, U., Hess, T., Karagiannis, D., Krcmar, H., Loos, P., Mertens, P., Oberweis, A., und Sinz, E. (2010). Memorandum zur gestaltungsorientierten Wirtschaftsinformatik. *Schmalenbachs Zeitschrift für betriebswirtschaftlichliche Forschung*, 62(9):662–672.

[Tam et al., 2011] Tam, B., Ehrgott, M., Ryan, D., und Zakerl, G. (2011). A comparison of stochastic programming and bi-objective optimisation approaches to robust airline crew scheduling. *OR Spectrum*, 33:49–75.

[Trapeze Software Inc., 2017] Trapeze Software Inc. (2017). Verfügbar unter http://www.trapezegroup.de/solution/public-transport-operators (letzter Abruf 30.3.2017).

[Veelenturf et al., 2016] Veelenturf, L., Potthoff, D., Huisman, D., Kroon, L., Maroti, G., und Wagelmans, A. (2016). A quasi-robust optimization approach for crew rescheduling. *Transportation Science*, 50(1):204–215.

[Verband Deutscher Verkehrsunternehmen VDV, 2011] Verband Deutscher Verkehrsunternehmen VDV (2011). *VDV - Schrift 730: Funktionale Anforderungen an ein itcs - Leitfaden für die itcs-Ausschreibung.* VDV-Schriften. beka Verlag, Köln.

[Verein Deutscher Ingenieure VDI, 2013] Verein Deutscher Ingenieure VDI (2013). *VDI-Richtlinie 3633: Simulation von Logistik-, Materialfluss- und Produktionssystemen - Grundlagen.* VDI-Richtlinien. VDI Verlag, Düsseldorf.

[Visentini et al., 2014] Visentini, M., Borenstein, D., Li, J. Q., und Mirchandani, P. (2014). Review of real-time vehicle schedule recovery methods in transportation services. *Transportation Research Part B*, 17(6):541–567.

[Weide et al., 2010] Weide, O., Ryan, D. M., und Ehrgott, M. (2010). An iterative approach to robust and integrated aircraft routing and crew scheduling. *Computers & Operations Research*, 37:833–844.

[Weider, 2007] Weider, S. (2007). *Integration of Vehicle and Duty Scheduling in Public Transport.* PhD thesis, Technische Universität Berlin.

[Wolsey, 1998] Wolsey, L. A. (1998). *Integer Programming.* John Wiley & Sons, New York.

[Wren, 2004] Wren, A. (2004). Scheduling Vehicles and their Drivers – Forty Years' Experience. *Technical Report 2004.03, University of Leeds - School of Computing, UK.*

[Xie et al., 2012] Xie, L., Amberg, B., und Suhl, L. (2012). A vehicle-crew-scheduling-and-rostering approach for improving driver satisfaction in public bus transport. *Technischer Bericht 08/2012, Universität Paderborn, DS&OR Lab.*

[Xie und Suhl, 2015] Xie, L. und Suhl, L. (2015). Cyclic and non-cyclic crew rostering problems in public bus transit. *OR Spectrum*, 37(1):99–136.

[Yen und Birge, 2006] Yen, J. W. und Birge, J. R. (2006). A stochastic programming approach to the airline crew scheduling problem. *Transportation Science*, 40(1):3–14.

[Zhou et al., 2011] Zhou, A., Qu, B., Li, H., Zhao, S., Suganthan, P., und Zhang, Q. (2011). Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, 1(1):32–49.

[Zitzler et al., 2002] Zitzler, E., Laumanns, M., und Thiele, L. (2002). Spea2: Improving the strength pareto evolutionary algorithm for multiobjective optimization. *Evolutionary Methods for Design, Optimisation, and Control*, Seiten 19–26. Barcelona, Spain. CIMMNE.

Anhang A

Anhang - Beispiel Konvergenz der Simulation

Ergänzend zu den in Kapitel 11 beschriebenen Experimenten veranschaulichen die folgenden Abbildungen exemplarisch den Konvergenzverlauf der Simulation (vgl. Abschnitt 11.1.2) zur Bewertung der Robustheit eines berechneten Umlauf- und Dienstplan zur Testinstanz B. Beispielhalft für drei Servicefahrten aus der gesamten Fahrtenmenge werden die jeweils erwartete propagierte bzw. sekundäre Verspätung (E(SD)) in Abhängigkeit des Simulationsverlaufes dargestellt.

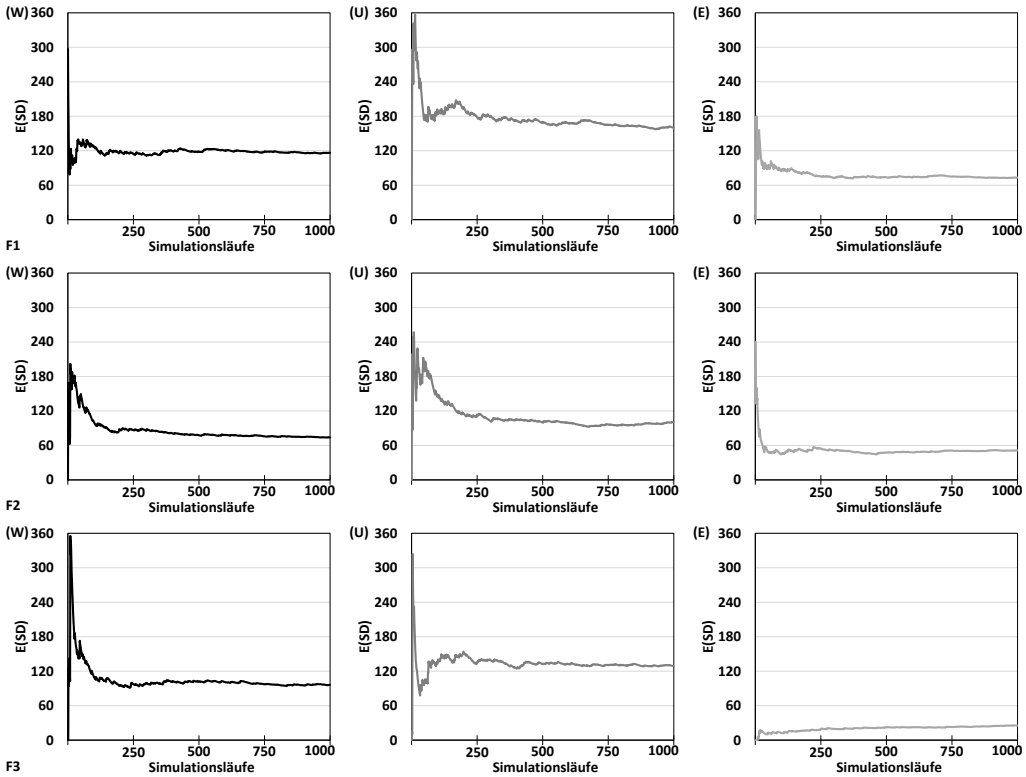

Abbildung A.1: Erwartete sekundäre Verspätung in Sekunden je Fahrt (F1, F2, F3) bei primären Verspätungen generiert mit verschobener Weibull-Verteilung (W), Gleichverteilung (U) und Exponentialverteilung (E)

Anhang B

Anhang - Beispiel Approximation pareto-optimale Lösungsmenge

Ergänzend zu den in Kapitel 11 durchgeführten Untersuchungen zeigt Abbildung B.1 die Approximation einer pareto-optimalen Lösungsmenge bezüglich der Zielsetzungen Kosteneffizienz und Robustheit durch das in Kapitel 10 vorgeschlagene hybride Lösungsverfahren. Die dargestellten Lösungen wurden für eine Testinstanz mit 549 Fahrten berechnet, bei der nur im Depot Pausen- und Ablösemöglichkeiten vorgesehen sind. Die beim Lösen verwendeten Diensttypen sowie die wesentlichen Parameter der problemspezifischen Optimierungsverfahren entsprechen den in Abschnitt 11.1 beschriebenen. Für den evolutionären Verfahrensverlauf wurden als Parameter zur Archivgröße $N = 25$, Populationsgröße $P = 40$, Mutationswahrscheinlichkeit $p_M = 0,01$ und für die Generationenanzahl $G = 120$ gewählt. So kann der Lösungsraum ausreichend durchsucht werden. Die initiale Population besteht aus vier mit dem beschriebenen Eröffnungsverfahren erstellten Lösungen. Als Bezugsgröße ist eine durch rein kostenminimierende integrierte Planung erstellte Lösung aufgeführt.

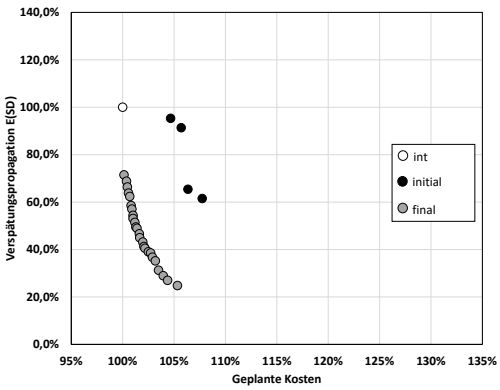

Abbildung B.1: Gegenüberstellung geplanter Kosten und erwarteter Verspätungspropagation der initialen Population (initial) und der finalen Population (final), sowie bei rein kostenminimierender Planung (int)

Veröffentlichungen

- Amberg Ba., Amberg Bo., Kliewer N. (2017), *Robust efficiency in urban public transportation: Minimizing delay propagation in cost-efficient resource schedules*, Transportation Science - Special Issue Recent Advances in Urban Transportation through Optimization and Analytics, im Druck.
- Amberg Ba., Ionescu L., Kliewer N. (2017), *Robust Efficiency in Public Bus Transport and Airline Resource Scheduling*, Operations Research Proceedings 2015, Wien, S. 259-264.
- Amberg Ba., Kliewer N., Beck M. (2012), *Robuste Effizienz bei der Einsatzplanung für Bus und Fahrer*, Der Nahverkehr 1/2 2012, S. 15-19.
- Kliewer N., Amberg Ba., Amberg Bo. (2012), *Multiple depot vehicle and crew scheduling with time windows for scheduled trips*, Public Transport, Vol. 3(3), S. 213-244.
- Amberg Bo., Amberg Ba., Kliewer N. (2012), *Integrated vehicle and crew scheduling in public transport with simultaneous consideration of timetables of different days*, Proceedings of the International Conference on Applied Mathematical Optimization and Modelling APMOD 2012, Paderborn, S. 297-302.
- Amberg Ba., Amberg Bo., Kliewer N. (2011), *Increasing delay-tolerance of vehicle and crew schedules in public transport by sequential, partial-integrated and integrated approaches*, Procedia - Social and Behavioral Sciences: The State of the Art in the European Quantitative Oriented Transportation and Logistics Research - 14th Euro Working Group on Transportation, Vol. 20, S. 292-301.
- Amberg Bo., Amberg Ba., Kliewer N. (2011), *Approaches for increasing the similarity of resource schedules in public transport*, Procedia - Social and Behavioral Sciences: The State of the Art in the European Quantitative Oriented Transportation and Logistics Research - 14th Euro Working Group on Transportation, Vol. 20, S. 836-845.
- Kliewer N., Amberg Ba., Amberg Bo. (2009), *Optimierungssysteme im ÖPNV: Mehrdepot Umlauf- und Dienstplanung mit Zeitfenstern für geplante Fahrten*,

Tagungsband (2) 9. Internationale Tagung Wirtschaftsinformatik, Wien, 2009, S. 55-64.

Erklärung

Hiermit erkläre ich an Eides Statt, dass ich die vorliegende Arbeit selbstständig und ohne unerlaubte fremde Hilfe angefertigt, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und die den benutzten Quellen und Hilfsmittel wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.
Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen.

Berlin, 4. April 2017

Bastian Amberg