
Maschinelles Lernen in der Anschlussdisposition

Machine Learning for connection dispatching

Bachelor-Thesis von Marc Arndt

Tag der Einreichung:

1. Gutachten: Prof. Dr. Johannes Fürnkranz
 2. Gutachten: Dr. Frederik Janssen, Anselmo Stelzer
-



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Fachgebiet Bahnsysteme und Bahntechnik
Fachgebiet Knowledge Engineering

Maschinelles Lernen in der Anschlussdisposition
Machine Learning for connection dispatching

Vorgelegte Bachelor-Thesis von Marc Arndt

1. Gutachten: Prof. Dr. Johannes Fürnkranz
2. Gutachten: Dr. Frederik Janssen, Anselmo Stelzer

Tag der Einreichung:

Erklärung zur Bachelor-Thesis

Hiermit versichere ich, die vorliegende Bachelor-Thesis ohne Hilfe Dritter nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

In der abgegebenen Thesis stimmen die schriftliche und elektronische Fassung überein.

Darmstadt, den 7. April 2016

(Marc Arndt)

Inhaltsverzeichnis

Abbildungsverzeichnis	5
Tabellenverzeichnis	6
1 Einleitung	7
1.1 Aufbau der Arbeit	7
2 Aktueller Forschungsstand und andere Arbeiten im Bereich „Automatisierte Anschlussdisposition“	9
2.1 Zielsetzung	9
2.2 Methodik	10
2.3 Andere Einsatzgebiete der Disposition	14
2.4 Einsatzgebiete des Maschinellen Lernens bei der Eisenbahn	15
3 Problemstellung	16
4 Grundlagen	17
4.1 Einführung in die Anschlussdisposition	17
4.1.1 Begriffsklärung	17
4.1.2 Maßnahmen für die Anschlusskonfliktlösung	19
4.1.3 Technische Informationssysteme	21
4.1.4 Betrachtete Attribute in der Praxis	22
4.2 Einführung in das Maschinelle Lernen	25
4.2.1 Definition einer Trainingsmenge	26
4.2.2 Regellerner	26
4.2.2.1 Definition einer Regel	27
4.2.2.2 Separate and Conquer Regellerner	28
4.2.2.3 RIPPER Algorithmus	31
4.2.3 Evaluierung	32
4.2.4 Erklärung des für das Maschinelle Lernen verwendeten Werkzeugs	34
5 Datenaufbereitung	37
5.1 Datenmodellierung	38
5.1.1 Betriebsstellen	38
5.1.2 Fahrt	39
5.1.3 Halte	39
5.1.4 Anschlüsse	40
5.1.5 Anschlusskonflikte	41
5.2 Aufbereitung der Trainingsmenge	41
5.2.1 Betriebsstellen	41
5.2.2 Fahrten	42
5.2.3 Fahrplandaten	43
5.2.4 Zusatzzüge (Sonder-, Ersatz- und Zusatzzüge)	44
5.2.5 Istdaten	45
5.2.6 Regelwartezeiten und Übergangszeiten	46

5.2.7	Anschlüsse und Anschlusskonflikte	51
5.2.8	Maßnahmen	53
5.3	Konvertierung des Datenmodells in das Attribute-Relation File Format	56
5.4	Exportierte Attribute der Anschlussdisposition	57
5.4.1	Gegebene Attribute	57
5.4.2	Weitere abgeleitete Attribute	58
6	Gewähltes Lernproblem	60
7	Ergebnisse	61
7.1	Versuchsaufbauten	61
7.1.1	Klassenzusammenstellungen	61
7.1.2	Featuremengen	65
7.1.3	Abkürzungen der Maßnahmen	67
7.2	Evaluations-Baseline	67
7.3	Evaluationsmethode	67
7.4	Evaluation	68
7.4.1	Klassenzusammenstellung \mathbb{K}_1	70
7.4.2	Klassenzusammenstellung \mathbb{K}_2	75
7.4.3	Klassenzusammenstellung \mathbb{K}_3	77
7.4.4	Klassenzusammenstellung \mathbb{K}_4	78
8	Fazit	80
9	Zukünftige Ansätze für die Automatisierte Anschlussdisposition	82
9.1	Betrachtungsebenen für Anschlusskonflikte	82
9.2	Einschränkungen bei der Verwendung von Maschinellen Lernverfahren bei der Anschlussdisposition	83
9.3	Gelabelte Anschlusskonflikte und Maßnahmen	83
9.4	Einführung weiterer Attribute	83
9.5	Weitere und umfangreichere Trainingsdaten	84
	Glossary	85
	Acronyms	86
	Literatur	87
A	Regelmengen	90
A.1	Klassenzusammenstellung \mathbb{K}_1	90
A.1.1	Featuremenge \mathbb{F}_1	90
A.1.2	Featuremenge \mathbb{F}_2	92
A.1.3	Featuremenge \mathbb{F}_3	94
A.1.4	Featuremenge \mathbb{F}_4	96
A.2	Klassenzusammenstellung \mathbb{K}_2	97
A.2.1	Featuremenge \mathbb{F}_1	98
A.2.2	Featuremenge \mathbb{F}_2	99
A.2.3	Featuremenge \mathbb{F}_3	100
A.2.4	Featuremenge \mathbb{F}_4	102

A.3	Klassenzusammenstellung \mathbb{K}_3	104
A.3.1	Featuremenge \mathbb{F}_1	104
A.3.2	Featuremenge \mathbb{F}_2	105
A.3.3	Featuremenge \mathbb{F}_3	106
A.3.4	Featuremenge \mathbb{F}_4	108
A.4	Klassenzusammenstellung \mathbb{K}_4	108
A.4.1	Featuremenge \mathbb{F}_1	109
A.4.2	Featuremenge \mathbb{F}_2	109
A.4.3	Featuremenge \mathbb{F}_3	110
A.4.4	Featuremenge \mathbb{F}_4	111

Abbildungsverzeichnis

4.1	Darstellungen eines Anschlusses	20
5.1	Ausschnitt der Übergangstabelle aus [2].	47
5.2	Tabelle mit den Regelwartezeiten für Anschlussfahrten.	48
5.3	Ausschnitt aus der Liste der Abweichungen von den Regelwartezeiten für den Hauptbahnhof Darmstadt.	49
5.4	Zustandsautomat für den Einlesevorgang der Abweichungen von den Regelwartezeiten.	49
7.1	Zeigt die Accuracy der Regelmengen aller Klassenzusammenstellungen unter Verwendung aller Featuregruppen.	69
7.2	Vergleich der richtig und falsch erkannten Maßnahmen in \mathbb{K}_1 unter Verwendung aller Featuregruppen.	70
7.3	Vergleich der richtig und falsch erkannten Maßnahmen in \mathbb{K}_2 unter Verwendung aller Featuregruppen.	75
7.4	Vergleich der richtig und falsch erkannten Maßnahmen in \mathbb{K}_3 unter Verwendung aller Featuregruppen.	76
7.5	Vergleich der richtig und falsch erkannten Maßnahmen in \mathbb{K}_4 unter Verwendung aller Featuregruppen.	78

Tabellenverzeichnis

4.5	Beispiel einer Trainingsmenge mit den Attributen: „Anzahl Ecken“, „Sind alle Winkel identisch?“ und „Sind alle Kanten gleich lang?“ sowie den Klassen: „Rechteck“, „Dreieck“ und „Stern“	26
4.6	Schema einer Konfusionsmatrix für die Klassen A, B und C.	32
5.6	Zeigt die direkt exportierten Attribute der jeweiligen Entitäten aus dem Datenmodell. . . .	58
5.7	Zeigt die zu berechnenden Attribute für die Entitäten aus dem Datenmodell.	59
7.4	Definition der evaluierten Klassenzusammenstellungen.	62
7.2	Verteilung der vorkommenden Maßnahmen auf die Anschlusskonflikte in \mathbb{K}_1	63
7.3	Verteilung der vorkommenden Maßnahmen auf die Anschlusskonflikte in \mathbb{K}_2 , \mathbb{K}_3 und \mathbb{K}_4	64
7.5	Definition der evaluierten Featuremengen.	65
7.6	Vergleich der in den Featuremengen \mathbb{F}_1 , \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 enthaltenen Attribute, gruppiert nach ihrer Objektzugehörigkeit im Datenmodell.	67
7.7	Abkürzungen der Maßnahmen in den Regelmengen.	67
7.8	Vergleich der Anzahl erzeugter Regeln für die verschiedenen Klassenzusammenstellungen in Abhängigkeit der einzelnen Featuregruppen.	68
7.9	Zeigt die Anzahl gelernter Regeln die mindestens eine Bedingung mit einer Linie, einer Fahrtrichtung oder einem Bahnhof enthalten.	68
7.10	Vergleich der Accuracy-Ergebnisse der verschiedenen Versuchsaufbauten im Vergleich zur Baseline.	68
7.11	Precision-Ergebnisse in Prozent für \mathbb{K}_1	71
7.12	Recall-Ergebnisse in Prozent für \mathbb{K}_1	72
7.13	Precision-Ergebnisse in Prozent für \mathbb{K}_2	74
7.14	Recall-Ergebnisse in Prozent für \mathbb{K}_2	75
7.15	Precision-Ergebnisse in Prozent für \mathbb{K}_3	76
7.16	Recall-Ergebnisse in Prozent für \mathbb{K}_3	77
7.17	Precision-Ergebnisse in Prozent für \mathbb{K}_4	78
7.18	Recall-Ergebnisse in Prozent für \mathbb{K}_4	78

1 Einleitung

Im Schienenverkehr möchte ein Fahrgast von einem Startbahnhof S zu einem Zielbahnhof Z gelangen. Häufig ist es dafür notwendig, dass der Fahrgast an einem Zwischenbahnhof A , auch Anschlussbahnhof genannt, in ein anderes Fahrzeug umsteigen muss, das ihn anschließend nach Z bringt. Dabei kann es vorkommen, dass das Fahrzeug, das den Fahrgast von S nach A bringen soll, verspätet ist, sodass der Fahrgast seinen Anschlusszug nicht mehr pünktlich erreichen kann.

Die Anschlussdisposition hat die Aufgabe Maßnahmen in einer solchen Situation zu finden, die es dem betroffenen Fahrgast ermöglichen sein Ziel weiterhin pünktlich zu erreichen. Diese Maßnahmen können von einem Warten des Anschlusszugs, bis zu einer Beauftragung eines Taxiunternehmens reichen, das die betroffenen Fahrgäste an ihr Ziel bringen soll. Jede dieser Maßnahmen hat Vor- und Nachteile, die vom Disponenten bei der Wahl der anzuwendenden Maßnahme in Betracht gezogen werden müssen. Im schlimmsten Fall kommen sonst durch eine Fehlentscheidung hohe Kosten auf den Bahnbetreiber zu, die andernfalls durch den Einsatz einer geeigneten Maßnahme hätten vermieden werden können.

Bei der Deutschen Bahn AG¹ wird die Anschlussdisposition von Mitarbeitern, den sogenannten Disponenten, durchgeführt. Dabei steht den Disponenten der Deutschen Bahn nur begrenzt technische Unterstützung zur Verfügung. Zudem ist es häufig schwierig nachzuvollziehen, aufgrund welcher Gegebenheiten ein Disponent eine Maßnahme getroffen hat. Ein weiteres Problem ist, dass ein Disponent nur einen Abschnitt des Streckennetzes überwacht und nicht weiß, was auf anderen Abschnitten passiert. Maßnahmen, die in einem Abschnitt getroffen werden, können allerdings auch Folgeauswirkungen auf andere Abschnitte haben, die der Disponent nur schwer abschätzen kann. Um diese Schwierigkeiten zu lösen, beziehungsweise zu verringern, ist es das Ziel dieser Arbeit dem Disponenten ein Werkzeug zur Verfügung zu stellen, das ihm dabei hilft, die bestmögliche Maßnahme für einen bestehenden Anschlusskonflikt zu finden.

Zu diesem Zweck wird untersucht, ob sich ein Regellerner aus dem Maschinellen Lernen dafür eignet, den Disponenten durch gelernte Regeln bei der Disposition zu unterstützen. Ein großer Vorteil bei der Verwendung von Regeln ist, dass diese von einem Menschen interpretiert werden können, was es dem Disponenten ermöglicht, die gefundene Entscheidung nachzuvollziehen.

1.1 Aufbau der Arbeit

Nach dieser Einleitung wird in Kapitel 2 (Aktueller Forschungsstand und andere Arbeiten im Bereich „Automatisierte Anschlussdisposition“) ein Überblick über ähnliche Arbeiten zum Thema „Automatisierte Anschlussdisposition“ gegeben. Darin wird zum einen gezeigt, welche aktuellen technischen Ansätze für die Anschlussdisposition existieren. Zum anderen wird darauf eingegangen, wo Disposition in vergleichbaren Bereichen noch eingesetzt wird, und wie sie dort technisch unterstützt wird. Außerdem wird ein Überblick darüber gegeben, wo Verfahren aus dem Maschinellen Lernen bei der Eisenbahn eingesetzt werden.

Darauf folgt das Kapitel 3, mit der in dieser Arbeit behandelten Problemstellung. Darin wird erläutert, warum welches Verfahren aus dem Maschinellen Lernen in dieser Arbeit eingesetzt wird.

¹ <https://www.deutschebahn.com/de/start.html>

Nachdem ein Überblick über die verwandten Arbeiten und die Problemstellung gegeben wurde, folgt das Kapitel 4 (Grundlagen), in dem die notwendigen Grundlagen dieser Arbeit erläutert werden. Die Grundlagen teilen sich in zwei Bereiche auf. Dies sind zum einen die Grundlagen der Anschlussdisposition und zum anderen die Grundlagen des Maschinellen Lernens.

In Kapitel 5 (Datenaufbereitung) wird darauf eingegangen, welche Daten für diese Arbeit zur Verfügung stehen und wie diese anschließend verarbeitet werden. Dabei wird zuerst auf das in dieser Arbeit verwendete Datenmodell eingegangen, in das die Daten eingelesen werden. Anschließend wird erläutert, wie die Informationen von der Deutschen Bahn in das Datenmodell eingelesen werden. Zuletzt wird erklärt, wie die Informationen aus dem Datenmodell exportiert werden, um sie in einer Evaluation zu verwenden.

Nachdem die zu Verfügung gestellten Daten eingelesen wurden, wird in Kapitel 6 das daraus resultierende Lernproblem erläutert.

Auf das Lernproblem wird in Kapitel 7 (Ergebnisse) eine Evaluation eines Regellerners auf den zuvor verarbeiteten Daten durchgeführt. Dabei werden mehrere Experimente zunächst erklärt und anschließend evaluiert. Während der Evaluation werden die gelernten Regeln analysiert und bewertet.

Anschließend wird in Kapitel 8 (Fazit), eine Antwort auf die Frage gegeben, ob sich der Einsatz eines Maschinellen Lernverfahrens, speziell eines Regellerners, für die Anschlussdisposition eignet.

Zuletzt wird in Kapitel 9 ein Ausblick darauf gegeben, welche Alternativansätze gewählt werden können, um den Einsatz von Maschinellen Lernverfahren für die Anschlussdisposition zu verbessern und in welche Richtung in Zukunft weiter geforscht werden kann.

2 Aktueller Forschungsstand und andere Arbeiten im Bereich „Automatisierte Anschlussdisposition“

Im Jahr 2014 hat die Deutsche Bahn jeden Tag mit über 40.000 Zügen insgesamt 5.676 Bahnhöfe angefahren. Das Gesamtstreckennetz der Deutschen Bahn AG umfasst dabei über 33.400 Kilometer. Insgesamt wurden dadurch etwa 6,2 Millionen Passagiere täglich an ihr Ziel gebracht [5].

Während des normalen Bahnbetriebs treten allerdings Probleme auf, wenn Züge sich verspäten, oder Teile des Streckennetzes blockiert sind. Eine Folge davon sind unter anderem Anschlusskonflikte, die entstehen, wenn Passagiere ihren geplanten Anschluss nicht wahrnehmen können. Für die Sicherung solcher Anschlüsse sind, nicht nur bei der Deutschen Bahn, sondern auch bei anderen Eisenbahnunternehmen, Disponenten zuständig. Ein Disponent hat dabei die Aufgabe, eine Maßnahme für einen gefährdeten Anschluss zu finden, mit dem Ziel, dass die Kunden ihr Ziel trotzdem pünktlich erreichen.

Neben der Anschlussdisposition wird auch in anderen Bereichen bei der Deutschen Bahn Disposition durchgeführt. Dazu gehört zum Beispiel die Fahrzeugdisposition, bei der die Zuteilung von Fahrzeugen zu Fahrten geplant wird.

Ein wichtiges Thema bei der Anschlussdisposition und bei der Eisenbahndisposition im Allgemeinen ist den Disponenten eine technische Unterstützung zu Verfügung zu stellen. Eines dieser Hilfsmittel für die Eisenbahndisposition sind Expertensysteme. Daneben existieren auch andere aktuelle Ansätze zur technischen Unterstützung von Disponenten, wie zum Beispiel Optimierungsverfahren und Verfahren aus der Statistik. Diese Ansätze unterscheiden sich durch ihre Methodik und ihrer Zielsetzung voneinander.

2.1 Zielsetzung

In diesem Abschnitt werden die wissenschaftlichen Arbeiten nach ihrer Zielsetzung in unterschiedliche Gruppen eingeteilt. Dazu gehören zum einen die Gruppen der Kunden- und Fahrzeug-orientierten Ansätze sowie eine Unterscheidung aufgrund der betrachteten Problemstellung.

Kunden- und Fahrzeug-orientierte Ansätze

Eine Unterscheidung der Zielsetzung von technischen Ansätzen für die Eisenbahndisposition ist die Unterscheidung in kundenorientierte und fahrzeugorientierte Ansätze.

Ziel der kundenorientierten Ansätze wie in [10, 23] ist es, die Gesamtverspätung der einzelnen Fahrgäste zu minimieren. Dabei können einzelnen Fahrgästen auch Gewichtungen zugewiesen werden. Diese Gewichte geben an, welche Priorität der jeweilige Fahrgast besitzt. Hat ein Kunde eine hohe Gewichtung, wird dieser im System priorisiert. Dies kann sich zum Beispiel in der benötigten Zeit auswirken, die der Kunde unterwegs ist.

Das Ziel der fahrzeugorientierten Ansätze [15] hingegen ist die Minimierung der Gesamtverspätung aller Fahrzeuge im Eisenbahnnetz. Dabei werden die eigentlichen Fahrgäste meistens in den fahrzeugorientierten Modellen nicht berücksichtigt. Ebenso wie bei den kundenorientierten Ansätzen, können auch bei den fahrzeugorientierten Ansätzen Gewichtungen eingesetzt werden. Der Unterschied zu den kundenorientierten Ansätzen ist, dass die Fahrzeuge, beziehungsweise die zu fahrenden Strecken, eine Gewichtung erhalten. Diese Gewichtung kann zum Beispiel von den angefahrenen Bahnhöfen, der Uhrzeit oder der Länge der gefahrenen Strecke abhängen.

Problemstellungen

Eine weitere Unterscheidung der Dispositionsansätze liegt in ihrer Problemstellung. Eine häufige Betrachtungsweise der Anschlusskonflikte versucht die Anschlüsse auf ihre Einhaltbarkeit zu bewerten. Dabei muss zwischen den Maßnahmen Warten und Weiterfahren abgewogen werden, um zu bestimmen, bis wann es sich für einen Anschluss noch lohnt, den abbringenden Zug warten zu lassen. Dies läuft auf die Fragestellung hinaus, bis wann es sich lohnt einen Zug warten zu lassen und ab wann es besser ist, ihn weiterfahren zu lassen.

So versuchen die Autoren von [24] in ihrer Arbeit automatisch zu entscheiden, ob ein Anschluss definitiv gehalten werden kann, unsicher oder kritisch ist oder definitiv gebrochen wird. Wird ein Anschluss definitiv gehalten, lohnt es sich den Abbringer warten zu lassen. Gilt der Anschluss hingegen als unsicher, beziehungsweise kritisch, muss von Fall zu Fall entschieden werden ob der Abbringer warten oder weiterfahren soll. Wird ein Anschluss hingegen definitiv gebrochen, sollte der Abbringer weiterfahren, ohne zu warten. Dafür wird für jeden Anschluss ein Wahrscheinlichkeitswert berechnet. Dieser Wahrscheinlichkeitswert wird daraufhin dafür verwendet, um mithilfe von Grenzwertintervallen herauszufinden, ob der Anschluss definitiv gehalten wird, unsicher ist, kritisch ist oder definitiv gebrochen wird. Ein Ergebnis der Arbeit ist, dass in einem durchgeführten Experiment nur 0,96% aller vom System als sicher eingestuften Anschlüsse später gebrochen wurden.

2.2 Methodik

Auch die eigentlichen Lösungsansätze der verschiedenen Arbeiten, die sich mit dem Eisbahndispositionsproblem auseinandersetzen, sind sehr unterschiedlich. Der aktuelle Stand der Forschung bei der Eisenbahndisposition sind Ansätze auf Basis von Optimierungsverfahren. Daneben existieren auch Arbeiten, die sich mit anderen Ansätzen, wie mobilen Agenten auseinandersetzen.

Expertensysteme

In der Eisenbahndisposition werden häufig Expertensysteme [21, 36] zur Unterstützung der Disponenten eingesetzt. Diese Expertensysteme enthalten von Experten vorgefertigte Entscheidungsregeln, die dafür verwendet werden, einem Disponenten bei seiner Arbeit zu helfen.

Für das chinesische Eisenbahnnetz arbeiten die Autoren von [8] an einem regelbasierten Expertensystem, das den Disponenten bei der Eisenbahndisposition helfen soll. Das von den Autoren vorgestellte System besteht aus zwei Subsystemen, dem eigentlichen Expertensystem und einem sogenannten „Train Movement Simulation System“ (TMSS), das die Informationen zu vorhergesagten Konflikten und Zugsbewegungen erzeugt. Zuerst sucht das TMSS nach Konflikten, diese Konflikte werden anschließend dem Expertensystem zur Findung einer Konfliktlösung weitergegeben. Hat das Expertensystem eine Lösung gefunden, wird diese dem Disponenten zur Entscheidung vorgelegt.

Der Autor von [32] schreibt über das „vehicle and infrastructure control and operations system“ (VICOS), das unter anderem ein Unterstützungssystem für die Disposition enthält. VICOS ist zum Beispiel in der Lage, Anschlusskonflikte und Fahrplankonflikte, durch eine Vorhersage des Fahrplanverlaufs der nächsten maximal zwei Stunden, zu erkennen. Nach der Detektion von Konflikten kann VICOS, abhängig vom Konfliktyp, Vorschläge liefern, mit denen die erkannten Konflikte gelöst werden können, wie zum Beispiel durch das Verlangsamen oder Beschleunigen von Zügen. Die Bewertung der Maßnahmen wird anhand von Experten- und Technischem-Wissen getroffen. Dabei wird Expertenwissen dazu eingesetzt, die Maßnahmen in hilfreiche und weniger hilfreiche Maßnahmen zu unterteilen. Anschließend werden die gefundenen Maßnahmen anhand von Kostenfunktionen optimiert und zuletzt dem Disponenten zur Entscheidung vorgelegt.

In [13] schreibt der Autor über das Decision Support System „Railway traffic Optimization by Means of Alternative graphs“ (ROMA). Dieses System ist dafür entwickelt worden, um mit Echtzeitproblemen wie Zugverspätungen und blockierten Strecken umzugehen.

Optimierungsverfahren

Neben dem Einsatz von Expertensystemen wird auch versucht durch Verwendung von Optimierungsverfahren, wie Integer Linear Programming [11] oder Genetischen Algorithmen [16, 38], die Fahrplange-nerierung, und im Konfliktfall die Konfliktlösung, zu optimieren, beziehungsweise zu verbessern.

In den Arbeiten [39, 37] vergleichen die Autoren mehrere Optimierungsverfahren miteinander. Dabei stellen sie fest, dass sich die meisten Optimierungsverfahren nur auf kleine Streckennetzen mit bis zu ca. 100 Haltestellen und bis zu ca. 1.000 Zügen beziehen. Auf größeren Streckennetzen benötigen viele Verfahren einen sehr großen Rechenaufwand, um rechtzeitig für die Entscheidung des Disponenten zu einem Ergebnis zu kommen. Dies führt dazu, dass dem Disponenten, nachdem eine Lösung gefunden wurde, nicht ausreichend Zeit bleibt diese anzuwenden, ehe der alte Fahrplan zeitlich zu weit fortge-schritten ist.

Integer Linear Programming: Einige dieser Arbeiten verwenden Integer Linear Programming (ILP) um einen neuen Dispositionsfahrplan zu generieren. In diesem neuen Fahrplan sind die vorher aufgetretenen Konflikte gelöst, beziehungsweise deren Auswirkungen minimiert.

In den Arbeiten [15, 33, 34] implementieren die Autoren jeweils Modelle, die dazu dienen, die Gesamtverspätung aller Züge im Eisenbahnnetz zu minimieren. Dafür werden dem Modell aktuelle und prognostizierte Verspätungsdaten der Züge im Eisenbahnnetz gegeben. Aus diesen Informationen wird mithilfe des Modells ein neuer Fahrplan generiert, in dem die abzusehende Gesamtverspätung durch geschicktes Umleiten der Züge auf andere Gleise oder Strecken minimiert ist. Die Gesamtverspätung aller Züge ergibt sich durch die Summe der Einzelverspätungen aller Züge des Fahrplans, die aufgrund von Konflikten entstanden sind. Für die Modelle werden verschiedene Nebenbedingungen definiert, die Abhängigkeiten zwischen den verschiedenen Entitäten im Modell darstellen. Dazu gehören Bedingungen wie, dass ein Zug erst von einem Bahnhof abfährt, nachdem er dort angekommen ist, oder das bei einem gehaltenen Anschluss der Abbringer erst abfährt, nachdem der Zubringer am Anschlussbahnhof angekommen ist.

Häufig sind diese Modelle sehr detailliert und benötigen Informationen über einzelne Streckenabschnit-te, Zuglängen, Infrastruktur und Signale. Diese Informationen liegen allerdings in der Praxis nur sehr selten vor, weshalb diese Modelle nur auf ausgewählten Streckennetzen eingesetzt werden können.

Genetische Algorithmen: Ein weiterer Optimierungsansatz für die Disposition stellt der Einsatz von Ge-netischen Algorithmen dar. Genetische Algorithmen dienen dazu, eine optimale Parameterbelegung für eine gegebene Funktion zu finden. Im Fall der Anschlussdisposition sind die gesuchten Parameter die Belegung eines Fahrplans, dazu gehören die im Fahrplan eingesetzten Züge, die Ankunfts- und Abfahrts-zeiten sowie die angefahrenen Bahnhöfe.

Ein gutes Beispiel für die Verwendung eines Genetischen Algorithmus für die Eisenbahndisposition ist [16]. Dort versuchen die Autoren, durch Verwendung eines Genetischen Algorithmus die Gesamtver-spätung aller Züge zu minimieren. Das Ergebnis vergleichen die Autoren mit denen eines künstlichen neuronalen Netzwerks, dass das Verhalten eines menschlichen Disponenten widerspiegeln soll. Dabei schneidet die Implementierung des Genetischen Algorithmus, mit einer niedrigeren Gesamtverspätung, allgemein besser ab, als das neuronale Netzwerk.

Maschinelle Lernansätze

Neben Optimierungsansätzen wird auch an Ansätzen aus dem Maschinellen Lernen, wie zum Beispiel stochastischen Verfahren geforscht, um den Dispositionsprozess zu verbessern.

In [1] wird ein stochastisches Verfahren vorgestellt, das aufgrund von Erwartungswerten, einer Kovarianz-Matrix, einer Genauigkeit-Matrix und anhand von Dichte-Werten einen Dispositionsfahrplan erstellt, in dem die Gesamtverspätung der Züge im Eisenbahnnetzwerk minimiert worden ist.

Die Autoren von [9] stellen ein Verfahren vor, dass aufgrund von Wahrscheinlichkeiten die Ankunfts- und Abfahrtsverspätungen von Zügen an Bahnhöfen vorhersagen soll. Dabei werden auch die Anschlussbeziehungen für die Bestimmung der Verspätung miteinbezogen, weil sich diese auf die Gesamtverspätung von Abbringer-Zügen auswirken können. In ca. 66% aller Fälle schaffen es die Autoren, die Ankunfts- und Abfahrtszeit mit einer Abweichung von maximal 5 Minuten zu bestimmen. In einem weiteren Experiment haben die Autoren getestet, wie gut ihr System Verspätungen über verschiedene Zeitintervalle voraussagen kann. Dabei ist herausgekommen, dass das System Verspätungen mit einer Abweichung von 4 Minuten innerhalb der nächsten 30 Minuten und mit einer Abweichung von 7 Minuten innerhalb der nächsten 4 Stunden vorhersagen kann.

Mit dem Einsatz eines Maschinellen Lernverfahrens mit dem Ziel eine passende Maßnahme zu einem gegebenen Anschlusskonflikt zu finden, beschäftigt sich der Autor von [40]. Dabei werden die Maßnahmen „Anschluss gewähren“ (im folgenden: AG), „AG Abbringergleiswechsel“, „Anschluss nicht gewähren“ (im folgenden: AnG), „AnG Reisende werden von einem anderen Zug aufgenommen“, „AnG Bus für Reisende bestellen“, „AnG Taxi für Reisende bestellen“ und „AnG Hotelzimmer für Reisende organisieren“ betrachtet. Als Maschinelles Lernverfahren wird der k -Nearest-Neighbors Algorithmus [27, Kapitel 14] eingesetzt. Dieser Algorithmus versucht mithilfe einer Ähnlichkeitsfunktion die k ähnlichsten Expertenentscheidungen zu einem gegebenen Anschlusskonflikt, aus einer Menge bereits vorhandenen Expertenentscheidungen, zu finden. Anhand der so gefundenen k ähnlichsten Entscheidungen kann dem Disponenten anschließend vom System eine Dispositionsmaßnahme vorgeschlagen werden.

Evaluiert wurde das verwendete Verfahren auf einer Datenmenge mit insgesamt 61.197 Anschlüssen, die über einen Zeitraum von 24 Kalendertagen von der Deutschen Bahn gesammelt wurden. Davon wurden 44.676 Anschlüsse als Trainingsmenge und 16.521 Anschlüsse als Testmenge verwendet. Für die Evaluation wurden die Trainings- und Testdaten anhand der 20, für die Maßnahmen-Entscheidung zuständigen, Transportleitungen aufgeteilt.

Im Schnitt konnten von den 16.521 Anschlüssen aus der Testmenge 60% der Anschlussmaßnahmen jeder Transportleitung korrekt erkannt werden. Dies ist nach Aussage des Autors nicht zufriedenstellend, um zu entscheiden, ob ein Anschluss gehalten wird oder nicht, da die erhaltenen 60% nur knapp über einem Würfelwurf mit einer Wahrscheinlichkeit von 50% liegen.

Darüber hinaus hat der Autor einige weitere Erkenntnisse gewonnen. Eine dieser Erkenntnisse ist, dass kein Zusammenhang zwischen der Anzahl der Trainingsdaten und der resultierenden Trefferquote besteht. Außerdem besteht laut Aussage des Autors kein Zusammenhang zwischen der Verteilung Trainingsbeispiele auf die Klassen und der Trefferquote. Eine weitere Erkenntnis ist, dass die erhaltenen Daten Anschlüsse mit identischen Informationen enthalten, für die unterschiedliche Maßnahmen getroffen wurden, womit das gewählte Lernverfahren nicht zu recht kommt. Insgesamt liefert das implementierte System laut Aussage des Autors gute und in einem angebrachten Rahmen zuverlässige Ergebnisse, die sich grundsätzlich dazu eignen den Verkehrsdisponenten in ihrer Entscheidungsfindung zu unterstützen.

Agentenbasierte Systeme

Neben Optimierungsansätzen und Maschinellen Lernverfahren wird auch an Ansätzen mit autonomen Agenten in der Eisenbahndisposition geforscht. Bei diesen Ansätzen werden die Akteure, die im Dispositionsprozess aktiv sind, jeweils durch einen unabhängigen Agenten repräsentiert. Diese Agenten können miteinander kommunizieren und so ihre Absichten anderen Agenten mitteilen, um ihre Intentionen bekannt zu machen und durchzusetzen. Ein Beispiel für Entitäten, die als Agenten repräsentiert werden können, sind die Fahrgäste beziehungsweise die Bahnkunden und die Bahnhöfe. Als Agent ist es jedem Kunden möglich, sich von seinem Bahnhof eine Verbindung mitteilen zu lassen. Der Bahnagentsagent wiederum, kann aufgrund der eingegangenen Anfragen genaue Statistiken zu seiner Auslastung und der Auslastung der Züge die bei ihm ankommen erstellen. Dies ermöglicht es dem Kunden gezielt, die für ihn beste Route zu erfahren.

In [10] stellen die Autoren Hilfsmittel für die Disposition vor, mit dem Ziel die Anschlussdisposition kundenorientiert durchzuführen. Dabei hat die Pünktlichkeit der Fahrgäste eine höhere Priorität, als die Pünktlichkeit der Züge. Dafür stellen die Autoren einen Ansatz vor, der auf mobilen Agenten basiert. Darin werden sowohl die Züge, die Bahnhöfe, die Disponenten und die Kunden als Agenten dargestellt. Außerdem enthält das Modell einen sogenannten Passagierrouter, der die Aufgabe hat die Routen für die Passagiere zu bestimmen. Dabei hat er sowohl die Aufgabe Routen für neue Kunden zu bestimmen, als auch die Routen von Fahrgästen, die ihren Anschluss verpasst haben, zu korrigieren. Die von den Autoren betrachteten Dispositionsmaßnahmen für Anschlusskonflikte beziehen sich auf Wartestrategien für einzelne Abbringer, die auf ihren Zubringer warten.

Insgesamt wurden in [10] 14 Warte-Strategien umgesetzt, die anschließend durch mehrere Simulationen getestet wurden. Der Zeithorizont dieser Simulationen betrug jeweils sechs Stunden, in denen ca. eine Million Fahrgäste mit knapp 9000 Zügen transportiert wurden. Ein Simulationslauf kann auf einem kleinen Cluster mit drei Computern mit je knapp 1.4 GHz und 512 MB RAM in weniger als zwei Stunden durchgeführt werden. Dabei kann eine Dispositionsentscheidung in wenigen Sekunden berechnet werden. Bei der Simulation unter Verwendung von passagier-orientierten Strategien konnten insgesamt circa 90% der Fahrgäste pünktlich an ihr Ziel gebracht werden. Während bei der Simulation unter Verwendung der üblicherweise von der Deutschen Bahn eingesetzten Regelwartezeiten [2] hingegen nur 70% der Fahrgäste pünktlich an ihrem Ziel ankommen.

Die Vorteile dieses Ansatzes sind zum einen, dass alle für einen Anschluss wichtigen Parteien als Agenten modelliert werden können, dadurch können sie direkt an der Entscheidung mitwirken. Zum anderen lässt sich dieser Ansatz auch, wie anhand der Simulation in [10] gezeigt wurde, auf größeren Streckennetzen einsetzen. Dafür müssen die Agenten ausreichend miteinander verbunden sein, um eine schnelle Kommunikation zwischen den einzelnen Agenten zu gewährleisten.

Der Nachteil dieser agenten-basierten Ansätze ist, dass sich einige Entitäten, wie zum Beispiel die Fahrgäste und die Fahrtstrecken, nur mit viel Aufwand als Agenten realisieren lassen. Der Grund dafür ist, dass die Fahrtstrecken dafür mit Sensoren ausgestattet werden müssen und die Fahrgäste bereit sein müssten aktiv im Dispositionsprozess mitzuhelfen. Außerdem wurden in den beiden oben genannten Arbeiten nur Wartemaßnahmen betrachtet. Damit lassen sich allerdings nicht alle Anschlusskonflikte lösen, wie zum Beispiel die aufgrund von Unfällen oder anderen Störungen verursachten Anschlusskonflikte, da dort ein Warten nur begrenzt hilft.

2.3 Andere Einsatzgebiete der Disposition

Die Disposition wird nicht nur im Bereich der Eisenbahn verwendet, sondern auch in anderen Bereichen.

Betondisposition

Ein solcher Bereich ist die Betonindustrie, dort müssen täglich viele Tonnen Ready Mixed Concrete (RMC) an Kunden verteilt werden. Diese Aufgabe ist, wie auch bei der Eisenbahn, häufig sehr Zeitkritisch, da RMC sehr schnell fest, und damit unbrauchbar wird. Aus diesem Grund kann ein Betonwerk meistens nur einen sehr kleinen Radius mit RMC bedienen. Tritt bei der Auslieferung zusätzlich eine Verspätung auf, zum Beispiel aufgrund eines Staus oder eines Fahrzeugdefekts, kann es zu Problemen kommen. Hinzu kommt, dass viele Kunden nur spezielle Zeitfenster zur Verfügung haben, in denen sie frischen Beton annehmen können. Außerdem wird häufig zusätzlich zum eigentlichen RMC ein Pumpfahrzeug gebraucht, das den Beton an die Stellen pumpt, an der er gebraucht wird. Auch dies muss im Zeitplan der Fahrzeuge berücksichtigt, und im Konfliktfall angepasst werden. In der Forschung werden für die Betondisposition häufig, wie auch bei der Eisenbahn-Disposition, Optimierungsverfahren eingesetzt.

Die Autoren von [26] verfolgen einen Ansatz, der ein Maschinelles Lernverfahren einsetzt. Dort wird mithilfe eines Regellerners und eines Entscheidungsbaums [28, Kapitel 3], der in eine Regelmenge umgewandelt wird, versucht ein Expertensystem für die Disposition in der Betonindustrie zu bauen. Dieses Expertensystem kann anschließend anhand der gelernten Regeln eine Empfehlung geben, welcher Kunde als nächstes angefahren werden soll.

Für die Evaluation der vorgestellten Methode haben die Autoren ein simuliertes Modell mit je drei Betonmischanlagen und Kunden erzeugt. Dafür wurden anschließend Kundenaufträge für 200 Tage generiert. Daraufhin wurde ein Disponent gebeten sich die simulierten Daten anzusehen, und eine Entscheidung über die Priorität der Kunden für jeden Tag zu treffen. Anschließend wurde aus den simulierten Daten und den Entscheidungen des Disponenten ein Trainingsmodell erzeugt, aus dem zuletzt ein Entscheidungsbaum generiert wurde. Insgesamt konnte das erzeugte Expertensystem in knapp 80% aller Fälle den nächsten anzufahrenden Kunden korrekt identifizieren. Außerdem ist während der Evaluation des Verfahrens eine Interpretation der erzeugten Regeln durch die Autoren erfolgt, bei der sie die Plausibilität der erzeugten Regeln bewerten.

Öffentlicher Nahverkehr

Neben der Betonindustrie und der Eisenbahn wird auch im öffentlichen Nahverkehr die Disposition eingesetzt. Auch dort ist es üblich, dass Busse und Straßenbahnen aufeinander warten, um Fahrgästen den Umstieg zu ermöglichen. Dabei kann es vorkommen, dass Busse und Straßenbahnen aufgrund eines zu großen Fahrgastaufkommens Verspätung haben, oder dass Fahrgäste nicht mehr zusteigen können, wenn ein Bus oder eine Straßenbahn bereits voll ist. Ein weiteres Problem stellen Verkehrsunfälle oder Staus dar. In all diesen Fällen muss ein Disponent eine Lösung finden, um die Zufriedenheit der Kunden zu gewährleisten. Interessant ist auch, dass der öffentliche Nahverkehr häufig mit dem Fernverkehr, egal ob mit der Bahn oder dem Flugzeug, zusammenarbeitet, um Fahrgäste schneller und besser an ihr Ziel zu bringen.

In [14] vergleichen die Autoren die Verwendung von Dispositionsregeln für die Anschlussdisposition von Bussen in Kombination mit und ohne sogenannter Intelligent Transportation Systems (ITS). Zu den ITS gehören unter anderem Technologien für die Positionsbestimmung von Fahrzeugen und für die Erfassung

der Anzahl der eingestiegenen Fahrgäste. Außerdem werden von den ITS Verspätungen von Bussen in Betracht gezogen. Eine der Erkenntnisse in [14] ist, dass die Verwendung von ITS zu kürzeren Wartezeiten für die Fahrgäste an den Bushaltestellen und zu geringeren Verspätungen der Busse führt.

2.4 Einsatzgebiete des Maschinellen Lernens bei der Eisenbahn

Maschinelle Lernverfahren werden von Eisenbahnunternehmen bereits für das „Predictive Maintenance“ [25] eingesetzt. Dabei wird versucht, mithilfe von Sensordaten Probleme mit der Infrastruktur, egal ob Schienen oder Fahrzeuge, vor Auftreten eines schwerwiegenden Fehlers zu erkennen und damit frühzeitig zu beheben.

In [25] versuchen die Autoren, mithilfe einer Support Vector Machine (SVM) [27, Kapitel 15], 3 bis 7 Tage vor Auftreten eines Eisenbahnschienendefekts, diesen vorherzusagen, um Entgleisungen zu verhindern. Für die Vorhersage von Defekten werden Daten verwendet, die zuvor von verschiedenen Sensoren erfasst wurden. Dazu gehören Temperaturmesssensoren, die sowohl die Temperatur an den Schienen als auch an den Rädern der darüber fahrenden Eisenbahnen messen. Außerdem werden Sensoren zur Messung der von Zügen auf die Schienen übertragenen Kräfte eingesetzt.

Evaluiert wurde das in [25] entwickelte Verfahren auf historischen Daten. Bei der Evaluation wird ein Decision Tree [28, Kapitel 3] als Baseline-Verfahren verwendet, das mit den Ergebnissen des erzeugten SVM-Modells verglichen wird. Mit der Verwendung des Decision Trees können die Autoren insgesamt 91,5% der Defekte in den nächsten 7 Tagen korrekt vorhersagen, wobei sie 6,8% der Defekte falsch vorhersagen. Mit der Verwendung einer SVM wurden hingegen insgesamt 97,5% der Defekte in den nächsten 7 Tagen korrekt und nur 5,6% falsch vorhergesagt. Ein weiteres Ergebnis der Evaluation ist, dass bei maximal 20 zugelassenen falsch prognostizierten Defekten, insgesamt ca. 45% der Gleisdefekte, der nächsten drei Tage, korrekt vorhergesagt werden konnten.

3 Problemstellung

Wie in Kapitel 2 erläutert wurde, wurden bereits mehrere Ansätze zur Unterstützung des Dispositionsprozesses erforscht. Expertensysteme haben dabei den Nachteil, dass sie von Experten erstellte Regeln benötigen. Diese Regeln sollten zudem sehr umfangreich sein, um möglichst viele Sonderfälle abzudecken. Die Erzeugung solcher Regeln kostet allerdings viel Zeit und Aufwand. Agentenbasierte Systeme haben hingegen den Nachteil, dass sie voraussetzen, dass alle an der Disposition beteiligten oder betroffenen Entitäten als Agenten modelliert werden müssen. Dies ist in der Regel nur mit viel technischem und finanziellem Aufwand möglich und setzt zudem die Mitarbeit der Fahrgäste voraus. Diese Verfahren haben allerdings häufig den Nachteil, dass sie im Fall der Optimierungsverfahren sehr rechenintensiv und damit nur auf kleinen Streckennetzen anwendbar sind. Dies hat zur Folge, dass sie nicht auf einem Streckennetz mit der Größe von dem der Deutschen Bahn anwendbar sind. Außerdem haben Optimierungsansätze den Nachteil, dass sie bei jeder neuen Verspätung einen vollständig neuen Fahrplan berechnen müssen.

Ansätze aus dem Bereich des Maschinellen Lernens haben diese Probleme im Regelfall nicht. Ein Maschinelles Lernverfahren erzeugt in der Regel einmal ein Modell, das anschließend dafür verwendet werden kann, Entscheidungen zur Lösung von einem oder mehreren Anschlusskonflikten zu treffen. Dabei muss insgesamt nur ein Modell berechnet werden, unabhängig von der Anzahl aufgetretener Verspätungen und anderer Störungen. Dies macht Verfahren aus dem Maschinellen Lernen sehr attraktiv für die Verwendung in der Anschlussdisposition.

Ein weiteres Problem mit den meisten Ansätzen, wie sie in Abschnitt 2 vorgestellt wurden, ist, dass ihre Ergebnisse von einem Menschen meist nur schwierig bis gar nicht interpretierbar sind. So bestehen die Ergebnisse von vielen Maschinellen Lernverfahren häufig aus einem Modell, das für einen Anschluss beziehungsweise für einen Anschlusskonflikt eine Liste von Zahlen erzeugt. Diese Zahlen geben an, welche Maßnahmen sich am besten zur Lösung des Anschlusskonflikts eignen. Die Interpretation dieser Zahlen und deren Bedeutung ist dabei allerdings häufig von einem Menschen nur schwer bis gar nicht nachvollziehbar.

Um das Problem der Verständlichkeit der Ergebnisse zu lösen, eignen sich insbesondere Regellernalgorithmen aus dem Bereich des Maschinellen Lernen. Diese Algorithmen liefern als Ergebnis eine Menge von Regeln. Diese Regeln lassen sich von einem Menschen lesen und nachvollziehen. Dies ermöglicht es auch dem Disponenten Rückschlüsse auf die Gründe, die zu einer Entscheidung geführt haben, zu schließen. Aus diesen Gründen wird in dieser Arbeit an einem Ansatz geforscht, der einen Regellernalgorithmus aus dem Bereich des Maschinellen Lernens für die Anschlusskonfliktlösung einsetzt, mit dem Ziel für einen gegebenen Anschlusskonflikt einen Maßnahmenvorschlag zu erhalten, der anschließend von einem Disponenten zur Konfliktlösung durchgeführt werden kann.

4 Grundlagen

In diesem Kapitel werden die Grundlagen, die später benötigt werden, erläutert. Die Grundlagen lassen sich dabei in zwei Bereiche unterteilen:

- Grundlagen der Anschlussdisposition (4.1)
- Grundlagen des Maschinellen Lernens (4.2)

Unter den Grundlagen der Anschlussdisposition wird alles benötigte Wissen zusammengefasst, das aus dem Bereich der Anschlussdisposition stammt und in dieser Arbeit eingesetzt wird. Unter den Grundlagen des Maschinellen Lernens wird alles benötigte Wissen zusammengefasst, das aus dem Gebiet des Maschinellen Lernens der Informatik stammt, und für die Bearbeitung dieser Arbeit relevant ist.

4.1 Einführung in die Anschlussdisposition

Für das Verständnis dieser Arbeit sind einige Themen aus der Anschlussdisposition wichtig, die in diesem Abschnitt eingeführt werden. Dazu gehören Definitionen wichtiger Begriffe, wie einer Betriebsstelle, einer Fahrt, eines Halts, eines Anschlusses und eines Anschlusskonflikts. Außerdem wird ein Überblick über verschiedene Maßnahmen gegeben, die in der Praxis von Disponenten zur Lösung von Anschlusskonflikten eingesetzt werden. Zum Abschluss dieser Einführung wird ein Überblick über die dem Disponenten vorliegenden Informationen gegeben. Diese Informationen sollen ihm dabei helfen, die beste Maßnahme für einen Anschlusskonflikt zu finden.

4.1.1 Begriffsklärung

Hier werden die in dieser Arbeit verwendeten Begriffe aus der Anschlussdisposition eingeführt und erklärt, wie diese zusammenhängen.

Betriebsstelle

In dieser Arbeit werden unter Betriebsstellen Haltepunkte und Bahnhöfe verstanden. Jede Betriebsstelle besitzt einen Namen, eine eindeutige Abkürzung, den sogenannten Richtlinie 100 (RIL100)- beziehungsweise Druckschrift 100 (DS100)-Code, sowie einem Kurznamen. In dieser Arbeit werden nur Betriebsstellen betrachtet, an denen Züge halten können und daher für einen Anschluss direkt interessant sind. Dies sind hauptsächlich Bahnhöfe.

Fahrt

Eine Fahrt beschreibt eine Zugfahrt mit einer Folge von Halten, die in einer vorher bekannten zeitlichen Reihenfolge durch den Zug angefahren werden. Die Haltefolge beginnt mit dem Starthalt des Zuges, an dem die Fahrt beginnt, gefolgt von einer beliebigen Anzahl Zwischenhalte und endet mit dem Endhalt, an dem die Fahrt endet. Insgesamt besteht eine Fahrt aus mindestens zwei Halten, dem Starthalt und dem Endhalt. Eine Fahrt kann eindeutig durch die Kombination ihrer Zugnummer und ihrem Start- und Endhalt identifiziert werden.

Halt

Ein Halt ist in dieser Arbeit ein Haltepunkt einer Fahrt an einer Betriebsstelle. Dabei haben die Fahrgäste die Möglichkeit aus dem Zug aus- oder einzusteigen. Neben der Fahrt, die den Halt durchführt, einer Betriebsstelle, an der der Halt stattfindet, gehören zu einem Halt Soll-, Prognose- und Istdaten. Die Solldaten geben an, zu welcher Zeit und an welchem Gleis der Zug laut Fahrplan an der Betriebsstelle ankommen, beziehungsweise abfahren soll. Die Prognosedaten geben an, zu welcher Zeit und an welchem Gleis im laufenden Betrieb prognostiziert wird, dass der Zug an der Betriebsstelle ankommt, beziehungsweise abfährt. Die Istdaten geben an, zu welcher Zeit und an welchem Gleis der Zug tatsächlich an der Betriebsstelle angekommen, beziehungsweise abgefahren ist. Durch Verspätungen und andere Störungen können die Prognose- und die Istdaten von den Solldaten abweichen. Abhängig davon, welche Daten für einen Halt vorliegen, gilt: Istdaten über Prognosedaten über Solldaten. Dies bedeutet, dass sobald die Istdaten für einen Halt vorliegen, diese vor den Prognosedaten verwendet werden. Liegen keine Istdaten vor, dafür Prognosedaten werden diese anstelle von Solldaten verwendet. Liegen weder Istdaten noch Prognosedaten vor, werden die Solldaten betrachtet.

Durchfahrt

Neben den Halten können auch sogenannte Durchfahrten an Betriebsstellen durchgeführt werden. Eine Durchfahrt unterscheidet sich von einem Halt dadurch, dass der Zug nicht an der Betriebsstelle hält, sondern nur hindurchfährt. Daher gehört zu einer Durchfahrt, genauso wie zu einem Halt, ein Zug mit einer Fahrt, eine Betriebsstelle, sowie Soll-, Prognose- und Istdaten, die angeben, wo und wann der betroffene Zug an der Betriebsstelle durchfahren wird. Im Rahmen dieser Arbeit werden Durchfahrten als Sonderfälle von Halten betrachtet, bei denen die Züge nicht an der Betriebsstelle halten und die Abfahrtszeiten gleich den Ankunftszeiten sind.

Übergangszeit

Die Übergangszeit (ÜZ) [2, S. 2] ist die zum Umsteigen zwischen zwei Zügen erforderliche Zeit. Die Übergangszeiten der Betriebsstellen sind in der Übersicht der Übergangszeiten (Planungsparameter des Fahrplans) festgeschrieben.

Regelwartezeit

Die Regelwartezeit (RWZ) ist die Zeit, die für ein Zugpaar, bestehend aus einem zubringenden Zug und einem abbringenden Zug, zur Anschlussicherung vorgesehen ist (siehe auch [35, S. 8]). Die Regelwartezeit werden allgemein aufgrund der Kombination der Gattungen des Zubringers und des Abbringers bestimmt, und können einer Tabelle in [2] entnommen werden. Zusätzlich zu den Standardwartezeitregeln existieren auch Abweichungen dazu, welche die allgemeinen Regeln in [2] ergänzen. Eine Übersicht der Abweichungen von den Regelwartezeiten kann ebenfalls aus [2] entnommen werden.

Wartezeitüberschreitung

Die Wartezeitüberschreitung (WZÜ) ist die Zeit, die neben der Regelwartezeit zusätzlich zur Sicherung eines Anschlusses benötigt wird, sollte die Regelwartezeit nicht ausreichen. Wenn eine Wartezeitüberschreitung zur Sicherung eines Anschlusses benötigt wird, muss sie zuerst vom Eisenbahnverkehrsunternehmen und der DB Netz AG genehmigt werden (siehe auch [35, S. 8, 9]).

Anschlüsse

Ein Anschluss stellt die Übergangsmöglichkeit von einem oder mehreren Fahrgästen an einer Betriebsstelle, dem Anschlussbahnhof, aus einem ankommenden Zug, dem sogenannten Zubringer, in einen abfahrenden Zug, dem sogenannten Abbringer, unter Einhaltung der Übergangszeit dar (siehe auch [35, 7]).

Anschlusskonflikt

Ist es den Fahrgästen zeitlich nicht möglich von einem Anschlusszubringer in dessen Abbringer umzusteigen, weil die Mindestübergangszeit zwischen der Ankunftszeit des Zubringers und der Abfahrtszeit des Abbringers nicht mehr gewährleistet ist, gilt der dazugehörige Anschluss als gefährdet. Dies kann zum Beispiel aufgrund einer Verspätung des Zubringers geschehen, wenn dessen Verspätung so hoch ist, dass die Fahrgäste nicht mehr genug Zeit haben aus dem Zubringer in den Abbringer umzusteigen. In einer solchen Situation wird von einem Anschlusskonflikt gesprochen (siehe auch [35, 7]).

Die Abbildung 4.1 zeigt drei Unterabbildungen, die insgesamt drei Anschlusszenarien enthalten. Das erste Anschlusszenario aus Abbildung 4.1a zeigt einen gehaltenen Anschluss, deren Zubringer pünktlich am Anschlussbahnhof ankommt, und dessen Abbringer pünktlich abfährt. Die zweite Abbildung 4.1b zeigt einen Anschluss, bei dem der Zubringer verspätet am Anschlussbahnhof ankommt, der Abbringer den Anschluss allerdings noch in Regelwartezeit halten kann. Die dritte Abbildung 4.1c zeigt einen Anschluss, dessen Zubringer, genau wie in Abbildung 4.1b, eine Verspätung aufweist, der Abbringer allerdings, anders als zuvor, mit Wartezeitüberschreitung warten muss, um den Anschluss zu halten.

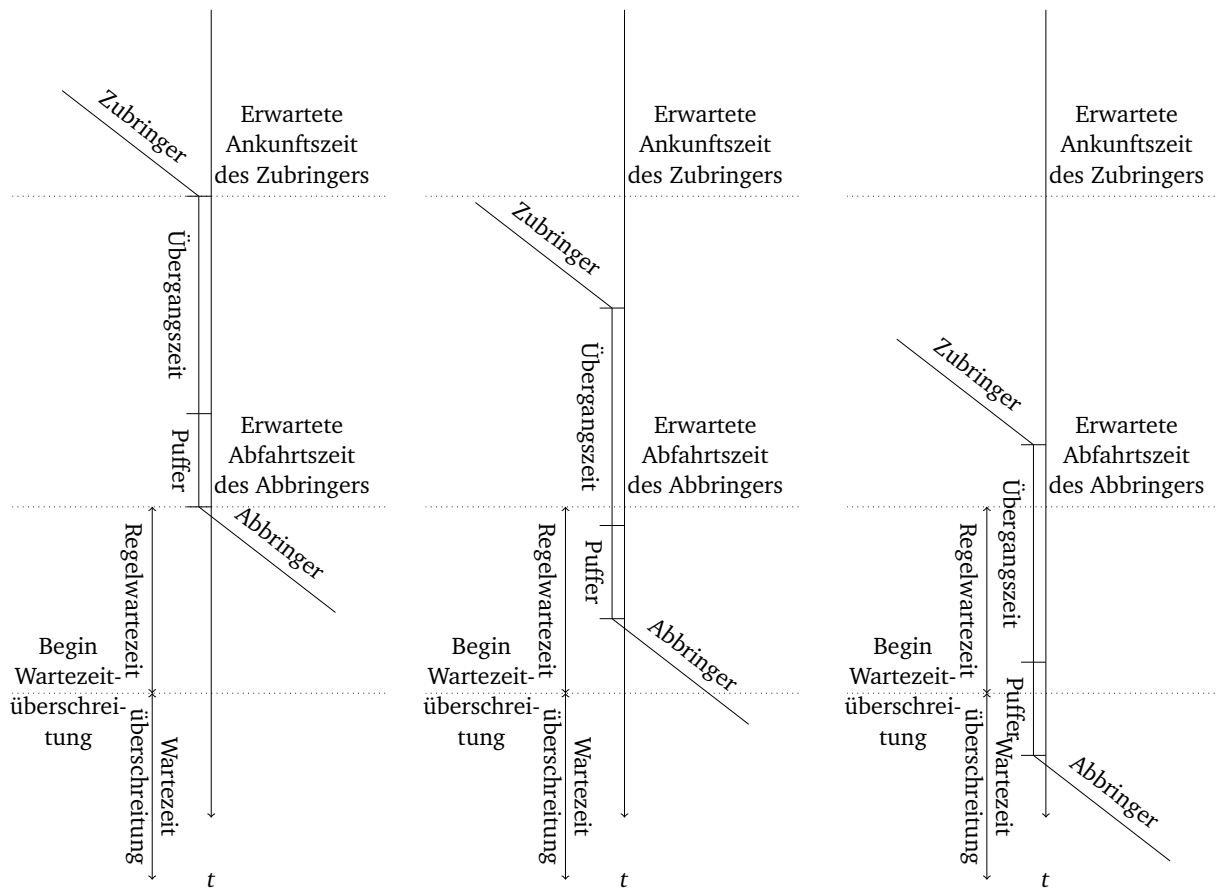
Tritt eine solche Situation ein, ist es die Aufgabe des Disponenten den gefährdeten Anschlusskonflikt, durch das Treffen einer geeigneten Maßnahme zu sichern.

4.1.2 Maßnahmen für die Anschlusskonfliktlösung

Nachdem ein Anschlusskonflikt aufgetreten ist, hat der Disponent die Aufgabe den Anschlusskonflikt zu lösen. Um einen Anschlusskonflikt zu sichern, muss der Disponent eine Maßnahme treffen, die dazu führen soll, dass die betroffenen Fahrgäste des Zubringers, die den Anschluss wahrnehmen wollen, trotzdem möglichst pünktlich an ihr Ziel gelangen. Für diesen Zweck kann der Disponent auf eine Palette von Maßnahmen zurückgreifen, die alle zum Ziel haben, den Fahrgästen den Umstieg zu ermöglichen, beziehungsweise die Fahrgäste möglichst pünktlich an ihr Ziel zu bringen.

Warten innerhalb einer definierten Regelwartezeit

Eine der einfachsten Maßnahmen ist das Wartenlassen des Abbringers auf den Zubringer am Anschlussbahnhof. Diese Maßnahme kann abhängig davon, wie lange der Abbringer warten muss, in zwei Untermaßnahmen untergliedert werden. Ist es dem Abbringer möglich, noch innerhalb der Regelwartezeit abzufahren, wird von der Maßnahme „Warten innerhalb einer definierten Regelwartezeit“ gesprochen. Mit dieser Maßnahme kann der Abbringer innerhalb seiner Regelwartezeit auf die Ankunft des Zubringers warten, um so den Fahrgästen, die den gefährdeten Anschluss nutzen wollen, dies weiterhin zu ermöglichen.



(a) Ein normaler gehaltener Anschluss ohne Verspätungen (b) Ein Anschlusskonflikt in dem der Abbringer in der Regelwartezeit gewartet hat. (c) Ein Anschlusskonflikt mit Wartezeitüberschreitung des Abbringers

Abbildung 4.1: Darstellungen eines Anschlusses

Warten mit Wartezeitüberschreitung

Die zweite Untermaßnahme neben der Maßnahme „Warten innerhalb einer definierten Regelwartezeit“ ist die Maßnahme „Warten mit Wartezeitüberschreitung“. Die Maßnahme „Warten mit Wartezeitüberschreitung“ kann von dem Disponenten getroffen werden, wenn die zu wartende Zeit des Abbringers für die Anschlusssicherung, die eingeplante Regelwartezeit überschreiten würde. In diesem Fall kann der Disponent die Maßnahme „Warten mit Wartezeitüberschreitung“ treffen, was es dem Abbringer ermöglicht, auch über die Regelwartezeit hinaus am Anschlussbahnhof zu warten. In diesem Fall wird von einer Wartezeitüberschreitung gesprochen.

Weiterfahren ohne zu Warten

In manchen Fällen ist es nicht gewollt, dass der Abbringer auf den Zubringer wartet. Zum Beispiel wenn die Gleise am Anschlussbahnhof belegt wären, wenn der Zubringer ankommen würde. In solchen Fällen kann sich der Disponent dagegen entscheiden, eine Wartemaßnahme zu treffen. Diese Entscheidung wird in dieser Arbeit als Maßnahme „Weiterfahren ohne zu Warten“ bezeichnet und beschreibt das nicht-Treffen einer Wartemaßnahme für den Anschlusskonflikt. Häufig wird diese Maßnahme mit einer anderen Maßnahme kombiniert, wie zum Beispiel einem Alternativanschluss. Dadurch können sowohl die durch den Anschlussausfall betroffenen Fahrgäste über einen Alternativanschluss als auch die Reisenden, die nicht von dem Konflikt betroffen sind, zeitnah an ihr Ziel gelangen.

Zubringer- und Abbringergleiswechsel

Unabhängig von der Entscheidung ob der Abbringer auf den Zubringer wartet, oder direkt weiterfährt, besteht auch die Möglichkeit sowohl den Zubringer, sowie den Abbringer an einem anderen Gleis des Anschlussbahnhofs halten zu lassen. Diese Maßnahme wird „Gleiswechsel“ genannt, und kann aus unterschiedlichen Gründen gewählt werden. Ein möglicher Grund für einen Gleiswechsel kann eine Zeiterparnis, durch die Rotation der Übergangszeit, für die Fahrgäste, durch das Annähern des Gleises des Abbringers an das des Zubringers, beim Umsteigen sein.

Alternativanschluss

Eine weitere Möglichkeit auf einen Anschlusskonflikt zu reagieren ist, den betroffenen Fahrgästen einen Alternativanschluss zu empfehlen. Der Alternativanschluss ermöglicht den Fahrgästen, die ursprünglich aus dem Zubringer des gefährdeten Anschlusses in dessen Abbringer umsteigen sollten, in einen alternativen Abbringer eines anderen Anschlusses umzusteigen. Diese Maßnahme wird „Alternativanschluss“ genannt. Dies kann allerdings für die Fahrgäste, aufgrund der eventuell abweichenden Fahrtroute des neuen Abbringers, andere Verspätungen zur Folge haben.

Zusatzzug

Eine weitere Maßnahme für die Lösung eines Anschlusskonflikts stellt der Einsatz eines Zusatzzugs dar. Bei dieser Maßnahme wird ein zusätzlicher Zug eingesetzt, der auf mindestens einem Teil der Strecke des ursprünglichen Anschlusszugs verkehrt. Die Fahrgäste können dann in den neu eingesetzten Zusatzzug umgeleitet werden und so weiter in Richtung ihres Ziels transportiert werden. Diese Maßnahme nennt sich daher „Zusatzzug“. Mit dem Einsatz eines zusätzlichen Zuges für den Abbringer ist es denkbar, die Fahrgäste flexibler weiter zu befördern, weil dieser, je nach gewählter Route, auch später als der normale Fahrplan abfahren kann.

4.1.3 Technische Informationssysteme

Die Deutsche Bahn verwendet unter anderem Systeme zur Schaffung eines Überblicks über die aktuelle Betriebslage. Dafür kommen unter anderem die Systeme Informationssystem Transportleitung Personenverkehr (ISTP) und Reisenden Informations System (RIS) bei der Deutschen Bahn zum Einsatz.

Informationssystem Transportleitung Personenverkehr

Das Informationssystem Transportleitung Personenverkehr (ISTP) [3] ist ein System, das von den Transportleitungen verwendet wird und das jederzeit den Regelbetrieb und die aktuelle Betriebslage der Bahn darstellt und vergleicht. Dafür erzeugt ISTP zum Beispiel bei Planabweichungen von Zügen automatisch Verspätungsprognosen, sowie eventuell daraus resultierende Anschlüsse und Anschlusskonflikte. ISTP besteht aus mehreren Teilkomponenten, diese sind:

Produktionsmodell Die Produktionsmodell-Anwendung dient dazu, aktuelle Zuglaufinformationen in verschiedenen Darstellungsarten tabellarisch und grafisch darzustellen.

Kundeninformation bei Großstörungen Die Kundeninformation bei Großstörungen, dient den Transportleitungen zur Kundeninformation. Damit ist es möglich, eine Vielzahl an Zügen und Verkehrsstationen zu informieren.

Verspätungserfassung und -Auswertung Die Verspätungserfassung und -Auswertung dient den Transportleitungen zur Verspätungserfassung von Zügen und zur Erstellung von tabellarischen Auswertungen der Störungsinformationen.

elektronischer Zugbericht Der elektronischer Zugbericht dient der Kommunikation von Zugmängeln. Damit hat das Zugpersonal die Möglichkeit per Handycode via SMS oder manuell über die Transportleitungen Mängel zu erfassen.

TP Mailbox Die Transportleitung Personalverkehr (TP) Mailbox dient dazu den Transportaustausch zwischen den Transportleitungen zu ermöglichen.

Reisenden Informations System

Das Reisenden Informations System (RIS) [4] dient als Datendrehscheibe, um aus betrieblichen Daten der Bahn Reisendeninformationen zu erzeugen. Dabei soll das RIS sowohl Kunden als auch Mitarbeiter mit konsistenten Reisendeninformationen versorgen. Das RIS versorgt die Kunden unter anderem durch folgende Informationskanäle:

- Anhand von Anzeiger und Ansagen im Zug und am Bahnhof.
- Individuell durch DB-Mitarbeiter im Zug und am Bahnhof.
- Über das Internet oder über mobile Endgeräte.

4.1.4 Betrachtete Attribute in der Praxis

Den Disponenten stehen für die Auswahl der besten Maßnahme im Falle eines Anschlusskonflikts einige Informationen zur Verfügung. Die in diesem Abschnitt genannten Informationen stammen aus dem Modul-Handbuch für ISTP [6], und stehen daher dem Disponenten für die Entscheidungsfindung zur Verfügung. Die folgenden Attribute sind in unterschiedliche Kategorien unterteilt, die angeben, ob die Attribute zu der Zubringerfahrt, der Abbringerfahrt, dem Anschlussbahnhof, einen Zughalt oder den Anschlusskonflikt selbst beschreiben.

Allgemeine Attribute des Zubringers und Abbringers

Parametername	Beschreibung
Zugnummer	Die Zugnummer ist die Kennung eines Zugs. Sie ist eindeutig, und kann damit dafür verwendet werden, ein Fahrzeug eindeutig zu identifizieren.
Gattung	Die Zuggattung wird zur Kategorisierung der Züge eingesetzt. Einige mögliche Zuggattungen sind: ICE, CE, IC, RE und RB.
Zugposition	Die Zugposition gibt die zuletzt gemeldete Position eines Zuges an. Diese Position ist die RIL100 des zuletzt angefahrenen Bahnhofs.
Aktuelle Verspätung	Die aktuelle Verspätung gibt die Zeitdauer an, um die ein Zug hinter der Fahrplanzeit verspätet ist. Muss ein Zug zum Beispiel an einem Bahnhof 5 Minuten länger halten als im Fahrplan eingeplant ist, hat er eine Verspätung von 5 Minuten.
Startbahnhof	Der Startbahnhof ist der erste Halt den ein Zug anfährt. Häufig ist der Startbahnhof der Endbahnhof des gleichen Zugs bei einer früheren Fahrt in die entgegengesetzte Richtung.

Parametername	Beschreibung
Endbahnhof	Der Endbahnhof ist der letzte Halt, den ein Zug anfährt. Häufig ist der Endbahnhof auch der Startbahnhof des gleichen Zugs bei einer späteren Fahrt in die entgegengesetzte Richtung.
Zuletzt getroffene Dispositionentscheidung	Die zuletzt getroffene Dispositionsentscheidung ist die chronologisch letzte Maßnahme, die für einen Zug angewendet wurde, um einen aufgetretenen Anschlusskonflikt zu lösen.

Allgemeine Bahnhofsinformationen

Parametername	Beschreibung
Name des Bahnhofs	Der Name eines Bahnhofs gibt an, um welchen Bahnhof es sich handelt.
Mindestübergangszeit	Die Mindestübergangszeit eines Bahnhofs ist die Zeitdauer, die für einen Fahrgast erforderlich ist, um von einem ankommenden Zug auf einen abbringenden Zug zu wechseln. [35, 2]

Allgemeine Informationen zu einem Zughalt

Parametername	Beschreibung
Angefahrener Bahnhof	Der angefahrene Bahnhof gibt den Namen des Bahnhofs an, der mit diesem Zughalt angefahren wird.
Der fahrende Zug	Der fahrende Zug ist das Fahrzeug, das diesen Zughalt durchführt. Angegeben wird der fahrende Zug durch seine Zugnummer.
Anzahl aussteigender Personen	Die Anzahl der aussteigenden Personen gibt an, wie viele Fahrgäste an diesem Zughalt den Zug verlassen werden. Hier werden nur die Fahrgäste eingerechnet, die zuvor eine Zuggebundenefahrkarte gekauft haben.
Anzahl zusteigender Personen	Die Anzahl der zusteigenden Personen gibt an, wie viele Fahrgäste an diesem Zughalt in den Zug einsteigen werden. Hier werden nur die Fahrgäste eingerechnet, die zuvor eine Zuggebundenefahrkarte gekauft haben.
Soll-Ankunftszeit	Die Soll-Ankunftszeit gibt an, wann der Zug laut Fahrplan an diesem Zughalt ankommen soll.
Prognose-Ankunftszeit	Die Prognose-Ankunftszeit gibt an, wann der Zug wahrscheinlich an diesem Zughalt ankommen wird. Die Prognose-Ankunftszeit wird entweder von einem Disponenten eingetragen, oder vom System selbst berechnet.
Ist-Ankunftszeit	Die Ist-Ankunftszeit gibt an, wann der Zug tatsächlich an diesem Zughalt angekommen ist. Die Ist-Ankunftszeit eines Halts kann sich aufgrund von Verspätungen von der Soll-Ankunftszeit im Fahrplan unterscheiden.
Soll-Abfahrtszeit	Die Soll-Abfahrtszeit gibt an, wann der Zug laut Fahrplan von diesem Zughalt abfahren soll.
Prognose-Abfahrtszeit	Die Prognose-Abfahrtszeit gibt an, wann der Zug wahrscheinlich von diesem Zughalt abfahren wird. Die Prognose-Abfahrtszeit wird entweder von einem Disponenten eingetragen, oder vom System selbst berechnet.

Parametername	Beschreibung
Ist-Abfahrtszeit	Die Ist-Abfahrtszeit gibt an, wann der Zug tatsächlich von diesem Zughalt abgefahren ist. Genauso wie die Ist-Ankunftszeit kann auch die Ist-Abfahrtszeit von der im Fahrplan genannten Zeit aufgrund von Verspätungen abweichen.
Soll-Gleis	Das Soll-Gleis gibt an, an welchem Gleis des Bahnhofs der Zug laut Fahrplan halten soll.
Ist-Gleis	Das Ist-Gleis gibt an, an welchem Gleis des Bahnhofs der Zug tatsächlich gehalten hat. Das Ist-Gleis kann aufgrund von zuvor getroffenen Maßnahmen von dem Soll-Gleis abweichen, zum Beispiel wenn das Soll-Gleis aufgrund einer Verspätung inzwischen von einem anderen Zug belegt wird.

Informationen zum Anschlusskonflikt

Parametername	Beschreibung
Soll-Anschlussart	Bei der Soll-Anschlussart handelt es sich um den geplanten Anschluss-typ.
Anzahl Umsteiger	Anzahl der Umsteiger mit Sonderfahr-scheinen für diesen Anschluss
Anzahl vorgemeldete Umsteiger	Anzahl der vom Zugchef vorgemeldeten Reisenden für diesen Anschluss
Bahnhof	Der Bahnhof eines Anschlusskonflikts ist der Bahnhof, an dem der Anschlusskonflikt auftritt. Dieser Bahnhof wird im folgenden auch Anschlussbahnhof genannt. Der Bahnhof wird durch seinen Namen angegeben.
Zubringer	Der Zubringer eines Anschlusses ist der Zug, der die Fahrgäste zum Anschlussbahnhof befördert.
Abbringer	Der Abbringer eines Anschlusses ist der Zug, der die Fahrgäste vom Anschlussbahnhof abholt, um sie von dort aus weiter in Richtung ihres eigentlichen Ziels zu befördern.
Soll-Ankunftszeit des Zubringers	Die Soll-Ankunftszeit des Zubringers ist die Zeit, zu der der Zubringer laut Fahrplan am Anschlussbahnhof ankommen soll.
Prognose-Ankunftszeit des Zubringers	Die Prognose-Ankunftszeit des Zubringer ist die Zeit, zu der angenommen wird, dass der Zubringer am Anschlussbahnhof ankommt. Die Prognose-Ankunftszeit des Zubringers wird benötigt, um eine Abschätzung treffen zu können, ob ein Anschluss ohne Eingreifen eines Disponenten gehalten werden kann oder ob der Disponent eine Maßnahme treffen muss.
Ist-Ankunftszeit des Zubringers	Die Ist-Ankunftszeit des Zubringers ist die tatsächliche Zeit, zu der der Zubringer am Anschlussbahnhof angekommen ist.
Soll-Gleis des Zubringers/Abbringers	Das Soll-Gleis des Zubringers beziehungsweise des Abbringers ist das Gleis, an dem der Zubringer beziehungsweise Abbringer laut Fahrplan am Anschlussbahnhof halten soll.
Ist-Gleis des Zubringers/Abbringers	Das Ist-Gleis des Zubringers beziehungsweise des Abbringers ist das Gleis, an dem der Zubringer beziehungsweise Abbringer tatsächlich am Anschlussbahnhof gehalten hat.
Soll-Abfahrtszeit des Abbringers	Die Soll-Abfahrtszeit des Abbringers ist die Zeit, zu der der Abbringer laut Fahrplan vom Anschlussbahnhof abfahren soll.

Parametername	Beschreibung
Prognose-Abfahrtszeit des Abbringers	Die Prognose-Abfahrtszeit des Abbringers ist die Zeit, zu der angenommen wird, dass der Abbringer vom Anschlussbahnhof abfährt. Die Prognose-Abfahrtszeit des Abbringers wird benötigt, um eine Abschätzung treffen zu können, ob ein Anschluss ohne Eingreifen eines Disponenten gehalten werden kann oder ob der Disponent eine Maßnahme treffen muss.
Ist-Abfahrtszeit des Abbringers	Die Ist-Abfahrtszeit des Abbringers ist die Zeit, zu der der Abbringer tatsächlich vom Anschlussbahnhof abgefahren ist.

4.2 Einführung in das Maschinelle Lernen

Ziel des Maschinellen Lernens ist das Konstruieren von Algorithmen, die lernen können, ein gegebenes Problem zu lösen. Um Maschinelles Lernen einzusetzen, wird zuerst eine Beschreibung von dem benötigt, was gelernt werden soll. Diese Beschreibung ist das sogenannte Lernproblem. Formal wird das Lernproblem nach Mitchell [28] wie folgt beschrieben:

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E .

Um ein Lernproblem nach Mitchell zu definieren, werden die folgenden Informationen benötigt:

T: Eine Aufgabe (Task), die von dem Lernalgorithmus gelernt werden soll.

E: Eine Anzahl vorliegender Daten (Experience), die dazu verwendet werden das Lösen der gestellten Aufgabe T zu erlernen.

P: Ein Messverfahren, mit dem bewertet werden kann, ob der Algorithmus gelernt hat die gestellte Aufgabe T zu lösen.

Mithilfe eines Lernproblems kann ein Algorithmus aus dem Maschinellen Lernen anhand von gegebenen Daten ein Modell erlernen, das angibt, wie eine Instanz, die vorher noch nie gesehen wurde, zu klassifizieren ist. Diese gegebenen Daten werden im folgenden Trainingsdaten genannt, weil sie dazu dienen den Lernalgorithmus auf ein gegebenes Lernproblem zu trainieren.

Insgesamt existieren vier verschiedene Typen von Trainingsdaten, die sich durch das Vorhandensein von vorliegenden Klasseninformationen zu den Trainingsdaten unterscheiden:

Supervised Learning: Beim Supervised Learning liegen dem Lernalgorithmus die Information, zu welcher Klasse ein Trainingsbeispiel gehört, vor.

Unsupervised Learning: Beim Unsupervised Learning liegen keine weiteren Informationen zu den Trainingsdaten vor.

Semi-supervised Learning: Beim Semi-supervised Learning liegen dem Lernalgorithmus, im Gegensatz zum Supervised Learning, nur für einen Teil der Trainingsbeispiele die gelabelten Klasseninformationen vor.

Reinforcement Learning: Beim Reinforcement Learning wird dem Lernalgorithmus nach Anwendung des Algorithmus auf eine vorher noch nicht gesehene Instanz Feedback gegeben. Das Feedback soll dabei helfen, die Ergebnisse des Lernalgorithmus zu verbessern.

In den folgenden Unterabschnitten werden Themen behandelt, die für die Verwendung eines Maschinellen Lernalgorithmus, insbesondere eines Regellerners, in dieser Arbeit vorausgesetzt werden. Dazu gehören die folgenden Themen:

1. Definition einer Trainingsmenge
2. Definition einer Regel
3. Definition eines Regellerners
4. Erklärung des hier für das Maschinelle Lernen verwendeten Werkzeugs

4.2.1 Definition einer Trainingsmenge

Eine Trainingsmenge ist eine Ansammlung von Trainingsbeispielen, die von einem Lernalgorithmus dazu eingesetzt werden können, eine Aufgabe zu lösen. Jedes Trainingsbeispiel besteht aus einer Menge von Attributen, die auch Feature genannt werden. Ein Attribut ist eine Eigenschaft des Trainingsbeispiels und hat die Aufgabe das Trainingsbeispiel zu beschreiben. Ein Beispiel für ein Attribut wäre die Anzahl der Ecken und die dazugehörigen Winkel eines Polygons. Zusätzlich ist beim Supervised und Semi-supervised Learning für jedes Trainingsbeispiel die dazugehörige Klasse bekannt, die vom Lernalgorithmus erkannt werden soll. Dabei wird auch von einer gelabelten Trainingsmenge gesprochen. Alternativ kann beim Supervised Learning eine Trainingsmenge als eine Menge von Klassen betrachtet werden, denen jeweils mehrere Trainingsbeispiele zugeordnet sind.

Anzahl Ecken	Sind alle Winkel identisch?	Sind alle Kanten gleich lang?	Klasse
4	ja	ja	Rechteck
4	ja	nein	Rechteck
3	ja	ja	Dreieck
3	nein	nein	Dreieck
5	ja	ja	Stern
7	ja	ja	Stern

Tabelle 4.5: Beispiel einer Trainingsmenge mit den Attributen: „Anzahl Ecken“, „Sind alle Winkel identisch?“ und „Sind alle Kanten gleich lang?“ sowie den Klassen: „Rechteck“, „Dreieck“ und „Stern“.

Ein Beispiel für eine Trainingsmenge ist in Tabelle 4.5 gegeben. Diese Tabelle enthält in jeder Zeile ein Trainingsbeispiel eines Datensets für geometrische Formen. Für jedes Trainingsbeispiel liegen die Attribute: „Anzahl Ecken“, „Sind alle Winkel identisch“ und „Sind alle Kanten gleich lang“ gegeben. Außerdem ist für jedes Trainingsbeispiel die dazugehörige Klasse gegeben. Die möglichen Klassen sind: „Rechteck“, „Dreieck“ und „Stern“.

4.2.2 Regellerner

Im Maschinellen Lernen existieren viele Lerntechniken. Eine dieser Lerntechniken sind Klassifikations-Regeln, die von Regellernern gelernt werden können.

Ein Regellerner ist ein Algorithmus aus dem Maschinellen Lernen, der anhand einer gegebenen Trainingsmenge ein Modell, bestehend aus einer Menge von Regeln, lernt. Mithilfe dieser gelernten Regeln kann anschließend eine gegebene Instanz einer Klasse aus der Trainingsmenge zugeordnet werden [20, 18].

Als Eingabe erhält ein Regellerner eine Trainingsmenge \mathbb{T} mit Trainingsbeispielen. Außerdem bekommt ein Regellerner eine Menge von Attributen \mathbb{A} , die in den Prämissen (siehe folgender Abschnitt 4.2.2.1) der resultierenden Regeln enthalten sein dürfen. Als Ausgabe gibt ein Regellerner eine Menge erzeugter Regeln \mathbb{R} zurück.

Je nach gewähltem Lernalgorithmus kann es vorkommen, dass einige Trainingsbeispiele aus \mathbb{T} von den gelernten Regeln falsch zugeordnet werden. Dies ist in einigen Fällen in Ordnung. Zum Beispiel, wenn das Datenset einige falsch gelabelte Trainingsbeispiele enthält, die dafür sorgen, dass eine Regel auch Negativbeispiele abdeckt. Ist dies nicht gewollt, muss für alle Trainingsbeispiele $B \in \mathbb{T}$, die von einer Regel in \mathbb{R} abgedeckt werden, gelten, dass auch die zu B gehörende Klasse gleich der Konklusion der abgedeckten Regel sein muss.

4.2.2.1 Definition einer Regel

Eine Regel ist ein Tupel, bestehend aus einer Prämisse und einer Konklusion.

$$\text{Prämisse} \rightarrow \text{Konklusion}$$

Die Prämisse einer Regel besteht hier aus einer Konjunktion von Bedingungen in folgender Form:

$$\text{Prämisse} = T_1 \wedge T_2 \wedge \dots \wedge T_n$$

Dabei beschreibt jedes T_i eine einzelne Bedingung, die einen Attributtest enthält.

Attributtests bestehen aus einem Attribut A_i , einem Wert V_i und einem Operator σ , mit dem A_i und V_i miteinander verglichen werden können. Ein solcher Attributtest hat den Aufbau:

$$T_i = A_i \sigma V_i$$

Ein Attributtest mit dem Gleichheitsvergleich hat damit die folgende Form:

$$T_i = (A_i = V_i)$$

Dieser Gleichheitstest wird erfüllt, wenn der für das Attribut A_i eingesetzte Wert gleich V_i ist. Alternativ zu dem Gleichheitsoperator können auch andere Operatoren für σ eingesetzt werden, um ein Attribut zu testen, wie zum Beispiel die Operatoren: $\neq, <, >, \leq, \geq$.

Werden alle diese Bedingungen einer Regel erfüllt, gilt automatisch die Konklusion der Regel. Die Konklusion gibt eine Eigenschaft an, die gelten muss, wenn die Prämisse der Regel erfüllt ist.

Beim Regellernen besteht die Konklusion einer Regel aus einer Klasse, die in der Trainingsmenge enthalten ist. Damit gilt für ein Trainingsbeispiel B , dessen Attribute die Prämisse einer Regel R erfüllt, automatisch, dass die zu B gehörende Klasse gleich der Konklusion von R ist.

Ein Beispiel einer Regel für das in Tabelle 4.5 definierte Datenset ist:

$$\text{Anzahl Ecken} = 4 \wedge \text{Sind alle Winkel identisch} = \text{ja} \rightarrow \text{Klasse} = \text{Rechteck}$$

4.2.2.2 Separate and Conquer Regellerner

Eine Regellernergruppe bilden die „Separate and Conquer“-Regellerner [18, 22]. Die Separate and Conquer Regellernalgorithmen zeichnen sich dadurch aus, dass sie nacheinander einzelne Regeln auf einer Trainingsmenge lernen, die in ihrer Summe die fertige Regelmenge bilden. Dafür wird mehrfach hintereinander der „Separate“- und der „Conquer“-Schritt ausgeführt.

Zuerst beginnt der Regellernalgorithmus damit eine Regel zu finden, die möglichst viele positive und wenige negative Trainingsbeispiele der gegebenen Trainingsmenge abdeckt. Dieser Schritt, bei dem eine einzelne Regel gelernt wird, ist der „Conquer“-Schritt. Die gelernte Regel wird anschließend der Menge der gelernten Regeln hinzugefügt.

Nachdem zuvor im Conquer-Schritt eine neue Regel gelernt wurde, müssen die von der Regel abgedeckten Trainingsbeispiele aus der Trainingsmenge entfernt werden. Dadurch ist es dem Regellerner möglich, eine neue Regel zu lernen, die andere Trainingsbeispiele abdeckt. Würden die abgedeckten Trainingsbeispiele nicht aus der Trainingsmenge entfernt werden, würde der Regellernalgorithmus die gleiche Regel mehrfach lernen. Das Entfernen der von der Regel abgedeckten Trainingsbeispiele wird „Separate“-Schritt genannt. Anschließend wiederholt der Separate and Conquer Regellerner den Separate- und den Conquer-Schritt auf der reduzierten Regelmenge, um weitere Regeln zu lernen. Dies wird so häufig wiederholt, bis der Regellernalgorithmus eine vorher definierte Endbedingung erfüllt.

Ein Trainingsbeispiel wird von einer Regel „abgedeckt“, wenn die Attribute des Trainingsbeispiels die Prämisse der Regel erfüllen. Ein „positives“ Trainingsbeispiel ist ein Trainingsbeispiel, das zu der gesuchten Klasse gehört. In einem 2-Klassenproblem, das zwei Klassen enthält, wird die gesuchte Klasse als positive Klasse bezeichnet und die andere Klasse als negative Klasse. Um eine bisher noch nicht gesehene Instanz entweder der positiven oder der negativen Klasse zuzuordnen, wird eine Regel in der gelernten Regelmenge gesucht, die die Attribute der Instanz abdeckt. Konnte eine solche Regel gefunden werden, gehört die Instanz der positiven Klasse an. Andernfalls gilt automatisch die Defaultregel, die alle nicht abgedeckten Klassen auffängt und die Instanz gehört der negativen Klasse an. In einem Mehrklassenproblem, in dem Regeln für die Erkennung aller Klassen gesucht werden, wird die aktuell betrachtete Klasse als positive Klasse bezeichnet. Um Regeln für alle vorhandenen Klassen zu lernen, wird vom Regellerner nacheinander für jede Klasse eine Regelmenge gelernt. Dabei wird vorgegangen wie bei einem 2-Klassenproblem, bei dem die Trainingsbeispiele der aktuell betrachtete Klasse zu der positiven Klasse gehören, und alle Trainingsbeispiele der anderen Klasse zu der negativen Klasse gehören. Nachdem für eine Klasse eine zufriedenstellende Regelmenge gelernt wurde, werden die gelernten Regeln der Menge aller gelernten Regeln hinzugefügt. Diese Regelmenge enthält am Ende auch die Regeln der anderen Klassen.

Um eine einzelne Regel zu lernen, existieren zwei Ansätze: der „top-down“ Ansatz sowie der „bottom-up“ Ansatz. In beiden Ansätzen versucht der Regellerner, durch iterative Verfeinerung eine Regel zu verbessern. Die beiden Ansätze unterscheiden sich dabei zum einen durch die Anfangsregel und zum anderen durch die Art wie die Regel anschließend verfeinert wird.

Beim top-down Ansatz beginnt der Regellerner mit einer Regel, die alle Trainingsbeispiele abdeckt, egal ob diese der positiven oder der negativen Klasse angehören. Anschließend versucht der Algorithmus iterativ die Regel zu spezialisieren, um weniger negative Beispiele abzudecken. In dieser Arbeit werden Regeln durch konjunktives Hinzufügen weiterer Bedingungen zu der Regel spezialisiert. Ein Beispiel für die Spezialisierung einer Regel auf Basis des in Abschnitt 4.2.2.1 gezeigten Regelbeispiels ist:

$$\begin{array}{ccc} \text{Alte Regel} & & \text{Neue spezialisierte Regel} \\ \underbrace{\text{Anzahl Ecken} = 4} & & \underbrace{\text{Anzahl Ecken} = 4 \wedge \text{Sind alle Winkel identisch} = \text{Ja}} \end{array}$$

Dabei wurde die Prämisse $\text{Anzahl Ecken} = 4$ durch Hinzufügen einer weiteren Bedingung spezialisiert.

Beim bottom-up Ansatz beginnt der Regellerner mit einer Regel, die kein Trainingsbeispiel abdeckt. Anschließend versucht der Algorithmus iterativ die Regel zu generalisieren, um zusätzliche positive Beispiele abzudecken. Eine Regel kann zum Beispiel durch Weglassen einer Bedingung aus einer Konjunktion von Bedingungen generalisiert werden. Ein Beispiel für die Generalisierung einer Regel auf Basis des in Abschnitt 4.2.2.1 gezeigten Regelbeispiels ist:

$$\overbrace{\text{Anzahl Ecken} = 4 \wedge \text{Sind alle Winkel identisch} = \text{Ja}}^{\text{Alte Regel}} \qquad \overbrace{\text{Anzahl Ecken} = 4}^{\text{Neue generalisierte Regel}}$$

Dabei wurde die Prämisse $\text{Anzahl Ecken} = 4 \wedge \text{Sind alle Winkel identisch} = \text{Ja}$ durch Weglassen einer Bedingung generalisiert.

Wichtig für das Lernen einer Regel ist die Auswahl der nächsten Bedingung, mit der die Regel verfeinert werden soll. Für die Auswahl der nächsten Bedingung wird eine Heuristik eingesetzt. Diese Heuristik erhält die möglichen Attribute, ihre anzunehmenden Werte und die noch nicht von der bisher gebauten Regel abgedeckten Trainingsbeispiele. Daraus berechnet sie, auf Basis der Eingaben, welches Attribut mit welcher Bedingung die Regel am besten verbessert.

Eine Heuristik die hierfür eingesetzt wird ist der *Foil Gain* [30]:

$$\text{Gain}(R_0, R_1) = t \cdot \left(\log \left(\frac{p_1}{p_1 + n_1} \right) - \log \left(\frac{p_0}{p_0 + n_0} \right) \right)$$

Dabei ist R_0 die Regel ohne die neue Bedingung, R_1 die Regel mit der neuen Bedingung, t die Anzahl positiver Trainingsbeispiele, die sowohl von R_0 als auch von R_1 abgedeckt werden, p_1 die Anzahl von positiven Trainingsbeispielen, die von R_1 abgedeckt werden, p_0 die Anzahl positiver Trainingsbeispiele, die von R_0 abgedeckt werden, n_1 die Anzahl negativer Beispiele, die von R_1 abgedeckt werden und n_0 die Anzahl negativer Beispiele, die von R_0 abgedeckt werden.

Neben dem Hinzufügen von Bedingungen zu einer Regel kann es auch sinnvoll sein, Bedingungen aus einer Regel zu entfernen. Zum Beispiel, wenn die Regel durch den top-down Ansatz zu sehr spezialisiert wurde. In einer solchen Situation, kann durch gezieltes Entfernen von Bedingungen, die Regel verallgemeinert werden, um mehr positive Trainingsbeispiele abzudecken. Dieser Vorgang des Verallgemeinerns von zu spezialisierten Regeln wird als „Prunen von Bedingungen“ bezeichnet. Dabei wird, nachdem eine weitere Bedingung für eine Regel gelernt wurde, für jede in der Regel enthaltene Bedingung überprüft, ob durch das Weglassen der Bedingung die Regel verbessert werden kann. Die Bewertung, wie gut eine Regel performt, kann ebenfalls, wie bei der Auswahl der nächsten Bedingung für die Verfeinerung der Regel, mithilfe einer Heuristik durchgeführt werden. Alternativ kann auch eine Metrik eingesetzt werden. Liefert diese Metrik für die neue geprunte Regel ein besseres Ergebnis als für die alte ungeprunte Regel, wird die geprunte Regel verwendet.

Eine Beispielmetrik ist:

$$M(R) = \left(\frac{p - n}{p + n} \right)$$

Dabei beschreibt R die zu überprüfende Regel, p die Anzahl der positiven von R abgedeckten Beispiele und n die Anzahl der negativen von R abgedeckten Beispiele.

Um die gelernte Regel besser prunen zu können, sollte die Trainingsmenge vor Beginn des eigentlichen Lernprozesses in eine sogenannte „Growing“- und in eine „Pruning“-Menge aufgeteilt werden. Dadurch werden die in der Regel enthaltenen Bedingungen nicht aufgrund derselben Trainingsdaten weggeprunt, mit denen sie der Regel zuerst hinzugefügt wurden. Die Growing-Menge wird vom Regellernalgorithmus dafür verwendet, die Regel um weitere Bedingungen zu erweitern und verfeinern. Die Pruning-Menge

wird dazu verwendet, die verfeinerte Regel falls notwendig zu generalisieren, um mehr Beispiele abzudecken, wenn dadurch die Qualität der Regel nicht abnimmt. Diese Strategie aus Growing und Pruning wird auch „Incremental Reduced Error Pruning“ genannt [19] (abgekürzt: I-REP).

Nach jeder Regelverfeinerung überprüft der Algorithmus, ob die Regel weiter verfeinert werden muss, oder ob es besser ist, eine weitere Regel zu lernen. Diese Überprüfung erfolgt ebenfalls mithilfe einer Heuristik und wird auch Stopping Criterion genannt. Dabei wird überprüft, ob durch Hinzufügen einer Bedingung die Regel verbessert werden kann. Ist dies nicht möglich, gilt die Regel als fertig gelernt, und der Regellerner kann mit dem Lernen der nächsten Regel beginnen. Zusätzlich dazu können weitere Kriterien gesetzt werden, um die Qualität der gelernten Regeln zu verbessern. Zum Beispiel empfiehlt es sich, eine Mindestanzahl abgedeckter Beispiele zu definieren. Diese Zahl von Beispielen müssen mindestens von einer gelernten Regel abgedeckt werden, bevor die Verfeinerung der Regel terminiert.

Neben dem Prunen von Bedingungen einer Regel können auch ganze Regeln aus der Regelmenge entfernt werden. Zum Beispiel, wenn diese sich mit anderen Regeln zu stark überlappen und vermehrt negative Trainingsbeispiele abdecken. Das Entfernen von Regeln aus der Regelmenge wird „Prunen von Regeln“ genannt. Dabei wird, genau wie beim Prunen von Bedingungen, eine Heuristik oder eine Metrik eingesetzt, um zu bewerten, ob der Einsatz einer Regelmenge mit einer ausgewählten Regel sich als besser erweist, als der Einsatz einer Regelmenge ohne diese Regel. Wird eine Heuristik eingesetzt, kann die Regel, durch deren Weglassen sich das Ergebnis der Heuristik am meisten verbessert, entfernt werden. Dafür wird eine Metrik eingesetzt, um für jede Regel in der gelernten Regelmenge zu berechnen, wie gut die Regelmenge ohne die Regel, und mit der Regel ist. Liefert die Metrik ein besseres Ergebnis unter Verwendung der Regelmenge ohne die betrachtete Regel, wird diese weggelassen.

Nachdem eine Regel gelernt wurde, muss der Regellernalgorithmus entscheiden, ob eine weitere Regel gelernt werden muss oder ob die gelernten Regeln ausreichen. Dafür wird überprüft, ob die gelernten Regeln eine vorher definierte Endbedingung erfüllen, üblicherweise müssen die gelernten Regeln alle positiven Trainingsbeispiele abdecken, ist dies der Fall, terminiert der Regellernalgorithmus. Andernfalls wird mit dem Lernen der nächsten Regel begonnen.

Nachdem eine Regelmenge gelernt wurde, kann diese dafür genutzt werden, die Klasse einer gegebenen Instanz herauszufinden. Dafür wird in der gelernten Regelmenge nach einer Regel gesucht, welche die Instanz abdeckt. Wurde eine solche Regel gefunden, gehört die Instanz zu der dazugehörigen Klasse. Konnte keine Regel gefunden werden, welche die Instanz abdeckt, gilt eine Defaultregel, die alle Instanzen auffangen soll, die von keiner anderen Regel abgedeckt werden. Eine sehr verbreitete Defaultregel ist die automatische Zuweisung zu der Klasse mit den meisten Trainingsbeispielen, in der Trainingsmenge auf der die Regelmenge gelernt wurde.

Im Rahmen dieser Arbeit sind die Regelmengen geordnet, womit von einer geordneten Regelmenge, beziehungsweise einer Decisionlist gesprochen wird. Bei einer geordneten Regelmenge wird nicht nach einer beliebigen Regel in der Regelmenge gesucht, die eine Instanz abdeckt. Sondern es werden der Reihe nach die einzelnen Regeln, mit der zuerst gelernten Regel angefangen, durchgegangen, bis eine Regel gefunden wurde, die die Instanz abdeckt. Deckt keine der Regeln in der Decisionlist die Instanz ab, gilt wie beim normalen Vorgehen eine Defaultregel, die angibt, welcher Klasse die Instanz angehört. Neben geordneten Regelmengen existieren auch ungeordnete Regelmengen, die in dieser Arbeit nicht verwendet werden.

4.2.2.3 RIPPER Algorithmus

Eine Implementierung eines Separate and Conquer Regellerners ist der RIPPER Algorithmus [12]. RIPPER steht dabei für „Repeated Incremental Pruning to Produce Error Reduction“.

Der RIPPER Algorithmus wird in dieser Arbeit für das Lernen von Anschlussdispositions-Regeln verwendet. Er ist noch immer ein State of the Art Regellernalgorithmus, der die Vorteile bietet, dass er sich zum einen für die Lösung von Mehrklassenproblemen eignet, und zum anderen zu sehr guten Ergebnissen mit sehr wenigen Falscherkennungen führt. Die hier angegebene Erklärung des RIPPER Algorithmus bezieht sich auf die Beschreibung aus [22, S. 65].

Der RIPPER Algorithmus ist in der Lage, eine Regelmenge für ein Mehrklassenproblem zu lernen. Dabei wird, wie in Abschnitt Separate and Conquer Regellerner beschrieben, für jede Klasse aus der Trainingsmenge nacheinander eine geordnete Regelmenge gelernt. Anschließend werden die gelernten Regelmengen der einzelnen Klassen zu einer großen geordneten Regelmenge zusammengefügt. Die Klassen werden dabei in aufsteigender Reihenfolge, nach der Anzahl vorhandener Trainingsbeispiele für die Klasse, bearbeitet, weil sie in dieser Reihenfolge gelernt wurden.

Der RIPPER Algorithmus unterteilt die Trainingsmenge in eine Growing- und in eine Pruning-Menge. Dabei enthält die Growingmenge $\frac{2}{3}$, und die Pruningmenge $\frac{1}{3}$ der Trainingsbeispiele. Die Growingmenge wird dafür verwendet neue Regeln zu lernen, wohingegen die Pruningmenge dafür eingesetzt wird, die gelernten Regel anschließend zu prunen.

Für das Lernen einer neuen Regel verwendet der RIPPER Algorithmus einen top-down Ansatz, bei dem die *foil gain* Heuristik, als Bewertungsfunktion für die nächste hinzuzufügende Bedingung eingesetzt wird. Eine Regel wird vom RIPPER Algorithmus iterativ so lange verfeinert, bis diese keine negativen Beispiele mehr abdeckt. Zudem setzt der RIPPER Algorithmus voraus, dass von jeder gelernten Regel mindestens zwei Trainingsbeispiele abgedeckt werden. Der RIPPER Algorithmus verwendet außerdem ein weiteres Stopping Criterion, das auf der Minimum Description Length (MDL) [31] basiert. Dieses Stopping Criterion berechnet für die „neue“ Regelmenge, die sich aus der alten Regelmenge inklusive der neuen Regel zusammensetzt, die MDL. Anschließend wird die neue Regel in die Regelmenge eingefügt, wenn die neu berechnete Description Length mindestens 64 Bit kürzer ist, als die bisher kleinste Description Length. Mit der Verwendung der Minimum Description Length wird versucht, eine Regelmenge sowie deren abgedeckte Trainingsbeispiele zu komprimieren. Dabei wird davon ausgegangen, dass eine bessere Regelmenge weniger Bytes zu deren Darstellung benötigt.

Die vom RIPPER Algorithmus gelernten Regeln können die Vergleichsoperatoren $=$, \leq und \geq enthalten. Wobei für nominale Vergleiche der Gleichheitsoperator $=$ eingesetzt wird und für numerische Vergleiche der Größer-Gleich- sowie der Kleiner-Gleich-Operator \geq und \leq eingesetzt werden.

Neben dem Optimieren einzelner Regeln versucht der RIPPER Algorithmus auch die für die einzelnen Klassen gelernten Regelmengen zu optimieren. Dabei testet der Algorithmus zum einen, ob die gesamte Regelmenge durch ein Austauschen einer Regel, durch eine neu gelernte Regel verbessert werden kann. Zum anderen testet der Algorithmus, ob die gesamte Regelmenge durch das Hinzufügen weiterer Bedingungen an eine vorhandene Regel verbessert werden kann. Sorgt eine der beiden Optimierungen für eine Verbesserung der ursprünglichen Regelmenge, wird die alte Regelmenge durch die optimierte Regelmenge ersetzt. Die Bewertung, ob eine optimierte Regelmenge besser ist, als eine nicht optimierte Regelmenge wird ebenfalls mithilfe eines Vergleichs, der berechneten Minimum Description Length, bestimmt. Falls nach einer Optimierung der Regelmenge nicht mehr alle positiven Trainingsbeispiele abgedeckt werden, muss der RIPPER Algorithmus weitere Regeln lernen, um die nicht mehr abgedeckten Trainingsbeispiele erneut abzudecken.

		Klassifiziert als			
		A	B	C	
Tatsächliche Klasse	A	$n_{A,A}$	$n_{B,A}$	$n_{C,A}$	$n_A = \sum_{x \in \{A,B,C\}} n_{x,A}$
	B	$n_{A,B}$	$n_{B,B}$	$n_{C,B}$	$n_B = \sum_{x \in \{A,B,C\}} n_{x,B}$
	C	$n_{A,C}$	$n_{B,C}$	$n_{C,C}$	$n_C = \sum_{x \in \{A,B,C\}} n_{x,C}$
		$\bar{n}_A = \sum_{y \in \{A,B,C\}} n_{A,y}$	$\bar{n}_B = \sum_{y \in \{A,B,C\}} n_{B,y}$	$\bar{n}_C = \sum_{y \in \{A,B,C\}} n_{C,y}$	$ E = \sum_{x,y \in \{A,B,C\}} n_{x,y}$

Tabelle 4.6: Schema einer Konfusionsmatrix für die Klassen A, B und C.

4.2.3 Evaluierung

Im Maschinellen Lernen ist die Evaluierung eines gelernten Modells sehr wichtig. Dies zeigt bereits die Frage nach einem Bewertungsverfahren beziehungsweise Messverfahren zu Beginn der Definition des Maschinellen Lernens. Im Maschinellen Lernen ist es bei der Evaluation das Ziel, ein für ein Lernproblem gelerntes Modell zu bewerten. Damit kann abgeschätzt werden, wie gut sich das gelernte Modell zur Lösung des Lernproblems in der Praxis eignet. Eine einfache Art der Evaluierung ist die Bewertung eines gelernten Modells durch einen Experten. Dies hat den Nachteil, dass die Bewertung eines Experten meistens subjektiv ist und die Evaluation sehr zeitintensiv ist.

Eine Alternative, zur Bewertung durch einen Experten, ist die Bewertung des Modells anhand von Testdaten. Dafür wird eine separate Testmenge mit gelabelten Trainingsdaten benötigt, die nicht in der Trainingsmenge enthalten sind. Auf dieser Testmenge kann ein gelerntes Modell anschließend bewertet werden, indem überprüft wird welchen Klassen die Testdaten durch das Modell zugewiesen werden. Anschließend können die Ergebnisse mit den bereits vorhandenen Labeln der Testdaten verglichen werden, um zu sehen, wie gut das gelernte Modell performt. Ein häufiges Problem bei der Evaluierung ist, dass nicht genügend Daten zu Verfügung stehen, um eine Testmenge zurückzuhalten. Aus diesem Grund sollte in einer solchen Situation ein Evaluierungsverfahren gewählt werden, dass auch mit wenigen Trainingsdaten zurechtkommt.

Cross-Validation

Ein häufig eingesetztes Verfahren für die Evaluierung anhand von Testdaten, das versucht das Problem des Zurückhaltens von Trainingsdaten zu kompensieren, ist die *Cross-Validation*. Bei der Cross-Validation wird die vorhandene Trainingsmenge in mehrere Teile, sogenannte „Folds“, aufgeteilt. Außerdem wird die Evaluierung mehrfach durchgeführt, um dafür zu sorgen, dass die Folds jeweils andere Trainingsbeispiele enthalten. Häufig wird bei der Cross-Validation daher von einer $n \times k$ Cross-Validation gesprochen. Dabei gibt n die Anzahl der verwendeten Folds an und k die Anzahl, wie häufig die Evaluation durchgeführt werden soll.

Insgesamt werden bei der $n \times k$ Cross-Validation n Testläufe durchgeführt. Auf jedem der n Teile der Trainingsmenge wird ein Testlauf ausgeführt. Dieser $\frac{1}{n}$ Teil der ursprünglichen Trainingsmenge stellt die Testmenge dar, die übrigen $n - 1$ Teile stellen zusammen die neue Trainingsmenge dar. Daraufhin wird auf jeder der n neuen Trainingsmengen ein Modell gelernt. Anschließend wird das Modell auf der dazugehörigen Testmenge evaluiert.

Das Ergebnis der Cross-Validation ist eine Liste von gefundenen Klassen für alle Instanzen aus der Trainingsmenge. Abhängig von der für eine Instanz aus der Trainingsmenge gefundenen Klasse können die

Instanzen anschließend in eine Konfusionsmatrix, wie in Tabelle 4.6 gezeigt, einsortiert werden. In dieser Tabelle steht ein Wert $n_{x,y}$ für die Anzahl Instanzen die zu Klasse y gehören und als Klasse x erkannt wurden. Für $x = y$ ist $n_{x,y}$ daher die Anzahl korrekt erkannten Instanzen der Klasse y . Die Werte der Konfusionsmatrix aus Tabelle 4.6 lassen sich anschließend dafür verwenden, um die Qualität des gelernten Modells zu bestimmen. Hierfür können die folgenden Funktionen verwendet werden:

Accuracy

Eine Funktion mit der die Qualität eines Modells bestimmt werden kann ist die sogenannte Accuracy. Die Accuracy berechnet sich mit der Funktion:

$$\text{accuracy} = \frac{\sum_{x \in \text{Alle Klassen}} n_{x,x}}{|E|}$$

Damit gibt die Accuracy an, wie viel Prozent der Trainingsinstanzen bei der Evaluierung korrekt erkannt werden konnten.

Precision

Eine weitere Funktion mit der die Qualität eines Modells gemessen werden kann ist die Precision. Die Precision muss für jede mögliche Klasse individuell berechnet werden. Dafür berechnet sie sich für eine Klasse C wie folgt:

$$\text{precision} = \frac{n_{C,C}}{\bar{n}_C}$$

Recall

Eine weitere Funktion für die Bestimmung der Qualität eines Modells ist der Recall. Auch der Recall muss, wie bereits die Precision, für jede Klasse C der Trainingsmenge individuell berechnet werden:

$$\text{recall} = \frac{n_{C,C}}{n_C}$$

Baseline

Um die Qualität eines gelernten Modells zu bewerten, wird ein Vergleichswert benötigt. Dieser Vergleichswert wird auch als Baseline bezeichnet und wird dazu verwendet, um zu bestimmen, ob der Einsatz des gelernten Modells zu guten oder zu schlechten Ergebnissen führt. Damit kann eingeschätzt werden, ob mithilfe eines gewählten Maschinellen Lernverfahrens ein passendes Modell für ein gegebenes Lernproblem gelernt werden kann.

Häufig wird für die Berechnung der Baseline ein Modell eingesetzt, das davon ausgeht, dass jede Testinstanz der am häufigsten in der Trainingsmenge enthaltenen Klasse angehört. Für dieses Modell wird anschließend, genau wie für das zu evaluierende Modell, ein Messwert, wie zum Beispiel die Accuracy, bestimmt. Anschließend kann durch einen Vergleich der Baseline-Werte, mit denen des zu evaluierenden Modells bestimmt werden, ob in dem Modell etwas gelernt wurde.

Beispiel: Bei einer Trainingsmenge mit insgesamt 500 Beispielen, die den vier Klassen A , B , C und D , mit $|A| = 100$, $|B| = 100$, $|C| = 200$ und $|D| = 100$, zugewiesen sind, ist die am häufigsten vorkommende Klasse C . Demnach würde der Baseline-Klassifizierer für das Beispiel immer die Klasse C vorhersagen und eine Baseline-Accuracy von 40% auf dem Trainingsset erreichen.

4.2.4 Erklärung des für das Maschinelle Lernen verwendeten Werkzeugs

Für die Evaluation wird im Rahmen dieser Arbeit, die Software WEKA eingesetzt. Waikato Environment for Knowledge Analysis (WEKA) [17] ist eine Softwaresuite für das Maschinelle Lernen und Data Mining. WEKA ist in der Programmiersprache Java geschrieben und enthält Implementierungen für Verfahren aus dem Maschinellen Lernen, wie zum Beispiel diverse Regellerner Algorithmen sowie Algorithmen zur Generierung von Entscheidungsbäumen. Einer der enthaltenen Regellerner ist eine Implementation des zuvor in Abschnitt 4.2.2.3 vorgestellten RIPPER Algorithmus. Die Implementierung des RIPPER-Algorithmus in WEKA nennt sich JRIP [41].

In dieser Arbeit wird WEKA speziell für die Erzeugung und die Evaluation der Dispositionsregeln anhand von aufbereiteten Daten eingesetzt. Diese Entscheidung ist darin begründet, dass WEKA bereits eine sehr große Funktionsvielfalt besitzt, in Java implementiert ist und eine frei zugängliche API bietet, was den Zugriff auf Funktionen von WEKA für diese Arbeit vereinfacht.

Attribute-Relation File Format

WEKA verwendet für die Speicherung von Trainings- und Testdaten das sogenannte Attribute-Relation File Format (ARFF). Eine ARFF-Datei (siehe auch: [29]) besteht aus zwei Abschnitten:

1. Der erste Abschnitt einer ARFF-Datei ist der sogenannte Header. Dieser Header enthält:

Den Name der Relation Dabei handelt es sich um eine Bezeichnung beziehungsweise einen Namen für die Art der Relation in der die Attribute zu der Klasse stehen. Der Relationsname wird mit dem Tag `@RELATION` eingeleitet

Die Attributdefinitionen Nach der Definition des Relationsnamens folgt eine Liste mit den Definitionen der in der Relation enthaltenen Attributen. Diese Definitionen bestehen aus dem Namen des Attributs und dessen Typ. Der Datentyp eines Attributs kann entweder ein numerischer Wert, ein String, ein Enum oder ein Datum sein. Jede Definition eines Attributs wird mit dem `@ATTRIBUTE` Tag eingeleitet.

2. Im zweiten Abschnitt einer ARFF-Datei werden die Trainings- beziehungsweise Testdaten angegeben. Dieser Abschnitt wird mit einem `@DATA` Tag eingeleitet. Die Datensätze, beziehungsweise Relationen, werden im gleichen Format wie das einer CSV-Datei angegeben. Das bedeutet, dass die Relationen durch Zeilenumbrüche, und einzelne Attribute durch Kommata voneinander getrennt werden.

Attribut-Datentypen: Wie bereits geschildert, können in einer ARFF-Datei jedes Attribut einen von vier Datentypen annehmen. Diese Datentypen sind „Zahlen“, „Strings“, „Datumsangaben“ und „Enums“:

Datentyp	Beschreibung	Beispielangabe
Numerische Werte	Numerische Werte können im Attribute-Relation File Format sowohl ganze Zahlen als auch Kommazahlen sein. Um ein Attribut als numerisches Attribut zu definieren muss bei der Definition des Attributs der Attributtyp „numeric“ im Header angegeben werden.	7
Strings	Strings sind im Attribute-Relation File Format unstrukturierte Texte. Ein Stringattribut kann im Header mit dem Attributtyp „string“ in der Definition des Attributs definiert werden. Die Angabe eines Stringwerts im Datenteil muss anschließend entweder in einfachen oder doppelten Anführungszeichen eingefasst werden.	"Beispieltext"
Datumsangaben	Neben numerischen Werten und Strings können auch Datumsangaben verwendet werden. Um ein Attribut als Datum zu definieren muss in der Definition des Attributs im Header-Abschnitt der Attributtyp „date“ gefolgt von der Formatierung der Datumsangabe angegeben werden. Diese Formatierung muss nach dem ISO-8601 Standard angegeben werden. Anschließend können die Datumswerte im gegebenen Format, und mit Anführungszeichen umgeben, in den Datensätzen als Werte für das Datumsattribut verwendet werden. Intern werden die Datumsangaben anhand ihrer Millisekunden-Repräsentation verwendet, wodurch sie intern in WEKA wie Zahlen behandelt werden.	'2015-06-15 15:06'
Enums	Als letzten Datentyp werden im Attribute-Relation File Format Enums unterstützt. Enums sind nominale Attribute. Um ein Attribut als Enum zu definieren, müssen in der Definition des Attributs alle möglichen Werte, die das Enum annehmen kann, in geschweiften Klammern und mit Komma getrennt aufgelistet werden. Zudem dürfen Enums keine Leerzeichen enthalten. Anschließend können die zuvor in geschweiften Klammern definierten Werte als mögliche Annahmewerte für das Attribut in den Relationen im Datenteil eingesetzt werden.	GLEISWECHSEL

Übertragung der Javadatentypen in das Attribute-Relation File Format: Um die in dieser Arbeit aufbereiteten Attribute im Abschnitt Datenaufbereitung in WEKA verwenden zu können, müssen diese in die Datentypen des Attribute-Relation File Format übertragen werden. Dabei können die meisten Datentypen aus Java sehr einfach in das Attribute-Relation File Format übertragen werden:

- Alle auf Zahlen basierten Daten aus Java wie Integer-, Long-, Float- oder Double-Werte können als numerische Werte in das Attribute-Relation File Format übernommen werden.
- Die Zeitangaben aus Java wie LocalDate-, LocalTime- oder LocalDateTime-Objekte, die zum Beispiel für die Realisierung von Ankunfts- und Abfahrtszeiten verwendet werden, können als Datumsangaben in das Attribute-Relation File Format übertragen werden. Dafür empfiehlt es sich, je nach Java-Datentyp ein anderes Format für die verwendeten Daten anzugeben. Zum Beispiel reicht für reine Uhrzeitinformatoren wie zum Beispiel in einem LocalTime-Objekt die Formatierung „HH:mm“.

-
- Java-Enums, oder andere Daten für die vorher bekannt ist, welche Werte sie annehmen können, sollten in Attribute-Relation File Format Enums übertragen werden.
 - Die meisten in Java verwendeten Texte, die dort als Strings gespeichert werden, wie zum Beispiel der Bahnname, lassen sich am besten als Enum darstellen, da bereits vorher bekannt ist, welche Bahnnamen existieren. Aus diesem Grund müssen diese zuerst in Java aufbereitet werden, bevor die Strings in das Attribute-Relation File Format übertragen werden können. Dafür werden die Strings zu einem Attribute-Relation File Format Enum umgewandelt. Im ersten Schritt werden alle vorkommenden Strings in einer Menge gesammelt. Diese String-Menge bildet die Menge an verschiedenen Werten, die das neue Enum annehmen kann. Im zweiten Schritt werden alle Leer- und Trennzeichen aus den Strings entfernt, damit die Strings als Enum in Weka akzeptiert werden.

Unbekannte Werte: Ist ein Wert für ein Attribut nicht bekannt, kann für das Attribut in der Attribute-Relation File Format Datei ein „?“ in der Relation eingetragen werden. Das „?“ gibt dem Klassifizierer an, dass die erwartete Information in diesem Datensatz nicht vorliegt.

5 Datenaufbereitung

Ein wichtiger Bestandteil dieser Arbeit ist die Aufbereitung von Trainingsdaten. Dies erfolgt mit dem Ziel, einen Regellerner zu trainieren, der mit den aufbereiteten Trainingsdaten eine Lösung für die Aufgabe eines gegebenen Lernproblems erlernen kann.

Zu diesem Zweck wurden von der Deutschen Bahn Daten zur Verfügung gestellt. Diese Daten bestehen aus einer Vielzahl von unterschiedlich formatierten Tabellen. Es liegt je eine Tabelle vor mit:

- den Betriebsstellen
- den Zügen
- einem ausgewählten Fahrplan der Deutschen Bahn
- passenden Istdaten zu dem gegebenen Fahrplan
- den Ausfällen
- den Gleiswechseln
- den Umleitungen
- den Zusatzzügen, Ersatzzügen und Sonderzügen
- Anschlüssen aus Informationssystem Transportleitung Personenverkehr
- Anschlüssen aus Reisenden Informations System
- den Übergangszeiten an den Betriebsstellen
- den Regelwartezeiten

Interessant für diese Arbeit sind davon allerdings nicht alle enthaltenen Informationen. Wichtig sind Informationen, die direkt für die Beschreibung von Anschlüssen gebraucht werden und die zur Erkennung von getroffenen Maßnahmen für Anschlusskonflikte dienen. Für die Beschreibung von Anschlusskonflikten werden die Tabellen mit den Anschlüssen, den Betriebsstellen, den Zügen, den Zusatz-/Ersatz-/Sonderzügen, dem Fahrplan, den Istdaten, den Übergangszeiten sowie den Regelwartezeiten gebraucht, da diese Tabellen die in Abschnitt 4.1.1 erläuterten Informationen zu einem Anschluss enthalten. Für die Maßnahmenerkennung können die Gleiswechsel-Tabelle sowie die Informationen aus den Istdaten und den Fahrplandaten verwendet werden. Die Informationen aus der Gleiswechsel-Tabelle können für die Erkennung von Gleiswechsel-Maßnahmen eingesetzt werden, während die Anschluss-Informationen aus der Anschlussstabelle, den Istdaten und den Fahrplandaten für die Erkennung von Alternativanschlüssen, Zusatzzügen und Wartemaßnahmen eingesetzt werden können.

Die übrigen Tabellen werden in dieser Arbeit nicht verwendet, da sie entweder keine Maßnahmen beschreiben, oder nicht mit absoluter Sicherheit mit einer Maßnahme in Verbindung gebracht werden können. Die Ausfälle stellen weniger eine eigene Maßnahme, sondern eher eine mögliche Ursache für einen Anschlusskonflikt dar. Einige Umleitungen können zwar als Maßnahmen betrachtet werden, allerdings ist nicht immer sichergestellt, dass eine Umleitung aufgrund eines Anschlusskonflikts durchgeführt wird. So könnte eine Umleitung auch eine Streckenstörung als Ursache haben, was eine Verbindung zu einem Anschlusskonflikt ausschließen würde. Ebenso ist es schwierig eine Umleitung einem Anschlusskonflikt zuzuordnen, da die Umleitung in der Regel nicht an der Betriebsstelle stattfindet, wo der Anschluss durchgeführt wird, sondern davor oder danach.

Gliederung

Damit die gelieferten Daten als Trainingsmenge verwendet werden können, müssen sie zuerst aufbereitet werden. Zu der Aufbereitung der gegebenen Daten gehören die folgenden, in diesem Abschnitt erläuterten Schritte:

- Zu Beginn dieses Abschnitts wird das in dieser Arbeit verwendete Datenmodell beschrieben. Dieses Datenmodell enthält einige der bereits in Abschnitt 4.1.1 eingeführten Elemente.
- Anschließend wird, im Hinblick auf das Datenmodell, erläutert, welche Informationen die Daten von der Bahn enthalten, und wie diese in das Datenmodell übertragen werden können. Dieser Abschnitt gliedert sich in das Einlesen der für einen Anschlusskonflikt relevanten Daten und in das eigentliche Einlesen und Erkennen der Maßnahmen für die zuvor eingelesenen Anschlusskonflikte.
- Nach dem Einlesen der Daten in das Datenmodell wird die Übertragung der Inhalte des Datenmodells in das in Abschnitt 4.2.4 beschriebene Attribute-Relation File Format erläutert.
- Zuletzt werden die Attribute genannt, die in das Attribute-Relation File Format exportiert werden sollen. Diese Attribute werden in die Gruppen der gegebenen und der abgeleiteten Attribute unterteilt.

5.1 Datenmodellierung

Um die von der Deutschen Bahn erhaltenen Daten aus unterschiedlichen Formaten und Quellen verwenden zu können, müssen sie zuerst in ein einheitliches Datenmodell eingelesen werden, in dem die Relationen zwischen den verschiedenen Datentabellen klar ersichtlich sind. Das Datenmodell, welches im Rahmen dieser Arbeit erzeugt und benutzt wird, enthält die folgenden Objekte: Betriebsstellen, Fahrten, Halte, Anschlüsse und Anschlusskonflikte. Den Objekten des Datenmodells fehlen im Vergleich zu deren Gegenstücken aus Abschnitt 4.1.4 einige Attribute. Der Grund dafür ist, dass die zur Verfügung gestellten Daten nur für die in diesem Abschnitt genannten Attribute Werte enthalten. Daher werden die Attribute in Abschnitt 4.1.4, für die keine Daten vorliegen, in der nachfolgenden Modellierung weggelassen.

5.1.1 Betriebsstellen

Eine Betriebsstelle B (siehe auch Abschnitt 4.1.1) aus der Menge aller Betriebsstellen \mathbb{B} wird durch folgendes Tupel beschrieben:

$$B = (\text{RIL100}, \text{Name}, \text{Name16}, \text{Bahnhofstyp}, t_{\text{mindestübergangszeit}}, t_{\text{mindestübergangszeit bg}}, t_{\text{min. haltezeit}}) \quad (5.1)$$

Parametername	Bedeutung
RIL100	Eine eindeutige Abkürzung der Betriebsstelle.
Name	Der Name der Betriebsstelle in Langform.
Name16	Der Kurzname der Betriebsstelle, bestehend aus maximal 16 Zeichen.
Bahnhofstyp	Der Typ der Betriebsstelle, zum Beispiel Hauptbahnhof, Bahnhof, Haltepunkt oder Umspannwerk.
$t_{\text{mindestübergangszeit}}$	Die Zeit die ein Fahrgast mindestens benötigt, um von einem Bahnsteig der Betriebsstelle zu einem anderen zu gelangen.
$t_{\text{mindestübergangszeit bg}}$	Die Zeit die ein Fahrgast mindestens benötigt, um von einem Zug zu einem anderen am selben Bahnsteig zu gelangen.
$t_{\text{min. haltezeit}}$	Die aus den an dieser Betriebsstelle haltenden Fahrten ermittelte minimale Haltezeit.

5.1.2 Fahrt

Eine Fahrt F aus der Menge aller Fahrten \mathbb{F} beschreibt eine Sequenz von durch einen Zug angefahrenen Halten. Für weitere Informationen zu einer Fahrt siehe auch Abschnitt 4.1.1. Eine Fahrt wird durch folgendes Tupel dargestellt:

$$F = (\text{Zugnummer}, \text{Linie}, \text{Richtung}, \text{Gattung}, B_{\text{Start}}, B_{\text{Ziel}}, t_{\text{Start}}, t_{\text{Ziel}}, t_{\text{min linie haltezeit}}, \text{Typ}, \mathbb{H}_F) \quad (5.1)$$

Parametername	Bedeutung
Zugnummer	Eine Identifikationsnummer der Fahrt. Diese Nummer ist in Kombination mit der Start- und Endzeit sowie dem Start- und Endbahnhof eindeutig.
Linie	Die Linie dieser Fahrt.
Richtung	Die Richtung die diese Fahrt fährt in der Form: (Von - Nach).
Gattung	Die Zuggattung dieser Fahrt. Dazu gehören: S, IC, ICE, RB, RE, usw.
B_{Start}	Der Startbahnhof dieser Fahrt.
B_{Ziel}	Der Zielbahnhof dieser Fahrt.
t_{Start}	Datum und Uhrzeit, zu der diese Fahrt laut Fahrplan vom Startbahnhof abfahren soll.
t_{Ziel}	Datum und Uhrzeit, zu der diese Fahrt laut Fahrplan am Zielbahnhof ankommen soll.
$t_{\text{min linie haltezeit}}$	Die minimale Haltezeit aller Fahrten mit der gleichen Linie dieser Fahrt an allen angefahrenen Halten unabhängig von deren Betriebsstelle.
Typ	Der Typ dieser Fahrt. Dabei kann es sich entweder um einen normalen Zug, der im Standardfahrplan angegeben wurde, einen Ersatzzug, einen Zusatzzug oder einen Sonderzug handeln.
\mathbb{H}_F	Die Menge aller Halte, die von dieser Fahrt angefahren werden.

5.1.3 Halte

Ein Halt H aus der Menge aller Halte \mathbb{H} beschreibt einen Haltepunkt einer Fahrt an einer Betriebsstelle. Für weitere Informationen zu einem Halt siehe auch Abschnitt 4.1.1. Im Datenmodell werden zu den Halten auch die Durchfahrten gezählt. Ein normaler Halt kann durch den Soll- beziehungsweise Ist-Fahrttyp von einer Durchfahrt unterschieden werden, je nachdem ob laut Fahrplan oder tatsächlich eine Durchfahrt durchgeführt wurde. Bei einer Durchfahrt werden zudem die Ankunftszeiten gleich den Abfahrtszeiten gesetzt. Ein Halt wird durch folgendes Tupel dargestellt:

$$H = (B, F, \text{Halttyp}, \text{Fahrtyp}_{\text{soll}}, \text{Fahrtyp}_{\text{ist}}, \text{Ausfallstatus}, t_{\text{soll ankunft}}, t_{\text{soll abfahrt}}, t_{\text{ist ankunft}}, t_{\text{ist abfahrt}}) \quad (5.1)$$

Parametername	Bedeutung
B	Die Betriebsstelle, an der dieser Halt stattfindet.
F	Die Fahrt, die diesen Halt durchführt.
Halttyp	Die Art dieses Halts. Die möglichen Haltearten sind: Starthalt, Zwischenhalt, Zielhalt. Dieser Parameter wird dazu verwendet um anzugeben, ob Ankunftszeiten und Abfahrtszeiten für diesen Halt vorausgesetzt werden.
Fahrttyp _{soll}	Dieser Parameter gibt an, ob es sich laut Fahrplan bei diesem Halt um eine Durchfahrt oder einen normalen Halt handelt.
Fahrttyp _{ist}	Dieser Parameter gibt an, ob für diesen Halt eine Durchfahrt oder ein Halt durchgeführt wurde. Dieser Parameter ist unabhängig vom Sollfahrttyp.
Ausfallstatus	Dieser Parameter gibt an, ob dieser Halt ausgefallen ist. Ist der Halt ausgefallen, wird er, im Gegensatz zu einer Durchfahrt, nicht von der Fahrt angefahren.
$t_{\text{soll ankunft}}$	Die im Fahrplan angegebene Ankunftszeit von F an B für diesen Halt.
$t_{\text{soll abfahrt}}$	Die im Fahrplan angegebene Abfahrtszeit von F an B für diesen Halt. Ist $\text{Fahrttyp}_{\text{soll}} = \text{Durchfahrt}$, gilt automatisch auch $t_{\text{soll ankunft}} = t_{\text{soll abfahrt}}$ an diesem Halt.
$t_{\text{ist ankunft}}$	Die tatsächlich eingetretene Ankunftszeit von F an B für diesen Halt.
$t_{\text{ist abfahrt}}$	Die tatsächlich eingetretene Abfahrtszeit von F an B für diesen Halt. Ist $\text{Fahrttyp}_{\text{ist}} = \text{Durchfahrt}$, gilt automatisch auch $t_{\text{ist ankunft}} = t_{\text{ist abfahrt}}$ an diesem Halt.

5.1.4 Anschlüsse

Ein Anschluss A aus der Menge aller Anschlüsse \mathbb{A} beschreibt die geplante Übergangsmöglichkeit von einem zubringenden zu einem abbringenden Zug. Für weitere Informationen zu einem Anschluss siehe auch Abschnitt 4.1.1. Ein Anschluss wird durch folgendes Tupel beschrieben:

$$A = (B, F_{\text{Zubringer}}, F_{\text{Abbringer}}, H_{B, F_{\text{Zubringer}}}, H_{B, F_{\text{Abbringer}}}, t_{\text{Eingang}}, t_{\text{Regelwartezeit}}, \text{Status}) \quad (5.1)$$

Parametername	Bedeutung
B	Der Anschlussbahnhof.
$F_{\text{Zubringer}}$	Die Zubringerfahrt dieses Anschlusses.
$F_{\text{Abbringer}}$	Die Abbringerfahrt dieses Anschlusses.
$H_{B, F_{\text{Zubringer}}}$	Der Halt des Zubringers an diesem Anschluss.
$H_{B, F_{\text{Abbringer}}}$	Der Halt des Abbringers an diesem Anschluss.
t_{Eingang}	Die Zeit, zu der dieser Anschluss gemeldet wurde.
$t_{\text{Regelwartezeit}}$	Die Regelwartezeit, die der Abbringer dieses Anschlusses zum Warten auf den Zubringer hat.
Status	Der Status dieses Anschlusses. Dies kann entweder „Gehalten“, „nicht Gehalten“ oder „Alternativanschluss“ sein.

5.1.5 Anschlusskonflikte

Ein Anschlusskonflikt A_k aus der Menge aller Anschlusskonflikte \mathbb{AK} beschreibt einen Anschluss A der gefährdet ist, weil die Mindestübergangszeit zwischen dem Zubringer und dem Abbringer nicht mehr gewährleistet ist. Für weitere Informationen zu eines Anschlusskonflikts siehe auch Abschnitt 4.1.1. Ein Anschlusskonflikt wird durch folgendes Tupel beschrieben:

$$A_k = (A, M) \quad (5.1)$$

Parametername	Bedeutung
A	Der gefährdete Anschluss.
M	Die Menge der zur Lösung des Anschlusskonfliktes getroffenen Maßnahmen, repräsentiert durch deren Namen.

5.2 Aufbereitung der Trainingsmenge

Nachdem das Datenmodell definiert wurde, muss das von der Deutschen Bahn zu Verfügung gestellte Datenpaket in das Datenmodell eingelesen werden. Dieses Datenpaket besteht aus einer Sammlung Tabellen im CSV- oder Excelformat sowie einigen PDF-Dateien.

Die CSV- und Exceltabellen enthalten Informationen zu von der Deutschen Bahn angefahren Betriebsstellen, Fahrten und Zusatzzüge sowie einen Fahrplan mit Solldaten der Deutschen Bahn und eine Tabelle mit dazugehörigen Istdaten. Außerdem sind zwei Tabellen mit Anschlussinformationen gegeben. Eine der beiden Tabellen stammt aus ISTP, während die andere Tabelle aus RIS entnommen wurde. Hinzu kommen Tabellen, mit den durchgeführten Gleiswechseln der Fahrten an Betriebsstellen, Zusatz- und Ersatzzügen, Halteausfällen und durchgeführten Zugumleitungen. Aus diesen Tabellen können einige der Maßnahmen für die Lösung von Anschlusskonflikten abgeleitet werden.

Neben den bereits genannten Tabellen sind außerdem eine Tabelle mit den Mindestübergangszeiten für die vorhandenen Betriebsstellen sowie eine Tabelle der Regelwartezeiten für die Abbringer-Zubringer Kombinationen der Anschlüsse in Form einer PDF-Datei der Deutschen Bahn gegeben.

Zuerst wird im Folgenden beschrieben, wie die Informationen für die Beschreibung der Anschlüsse eingelesen werden. Anschließend wird erläutert, wie die Maßnahmen aufbereitet und Anschlusskonflikten zugewiesen werden.

5.2.1 Betriebsstellen

Die Betriebsstellen, die durch Fahrten der Deutschen Bahn angefahren werden, liegen in einer Tabelle vor. Jede Zeile dieser Tabelle enthält Informationen zu einer Betriebsstelle. Zu jeder Betriebsstelle liegen die folgenden Informationen vor:

- ein RIL100-Code für die eindeutige Identifikation der Betriebsstelle
- der Name der Betriebsstelle
- der Kurzname der Betriebsstelle, bestehend aus maximal 16 Zeichen

Zusätzlich dazu sind für jede Betriebsstelle weitere Attribute gegeben, wie der Betriebszustand, der Betriebsstellentyp (zum Beispiel: Haltepunkt, Bahnhof, Abzweigung und Bahnhofsteil), das Inbetriebnahme-Datum sowie gegebenenfalls das Datum der Außerbetriebnahme.

Vorgehen

Das Einlesen der Betriebsstellen erfolgt Zeile um Zeile. Dafür wird für die gerade bearbeitete Betriebsstelle überprüft, ob sie sich in Planung befindet und ob sie derzeit betrieben wird. Eine Betriebsstelle befindet sich in Planung, wenn der Betriebszustand den Wert „Planung“ enthält. Währenddessen ist eine Betriebsstelle in Betrieb, wenn in der Spalte mit dem Datum der Außerbetriebnahme kein Eintrag vorhanden ist. Ist die Betriebsstelle derzeit in Betrieb und nicht in Planung, wird im Datenmodell eine neue Betriebsstelle mit den Informationen aus der Tabellenzeile angelegt.

5.2.2 Fahrten

Auch eine Tabelle mit Fahrten wurde von der Deutschen Bahn zu Verfügung gestellt. Diese Tabelle enthält eine Liste von Fahrten, die an drei Betriebstagen der Deutschen Bahn durchgeführt wurden. Jede Zeile der Fahrtentabelle enthält die Informationen zu einer Fahrt. Für jede Fahrt liegen die folgenden Attribute vor:

- die Zugnummer
- das Fahrtdatum
- die Zuggattung
- der RIL100 des Startbahnhofs
- die Startuhrzeit
- der RIL100 des Zielbahnhofs
- die Zieluhrzeit
- der Liniename

Insgesamt können an einem Betriebstag zu unterschiedlichen Zeiten mehrere Fahrten mit der gleichen Zugnummer vorkommen, allerdings nicht zur selben Zeit. Aus diesem Grund lässt sich in den gegebenen Tabellen jede Fahrt eindeutig durch die Kombination aus deren Zugnummer, Datum, Startbahnhof, Startuhrzeit, Zielbahnhof und Zieluhrzeit erkennen.

Vorgehen

Das Einlesen der Fahrten erfolgt, genau wie das Einlesen der Betriebsstellen, Zeile für Zeile. Jede Zeile der Fahrtentabelle enthält die Informationen zu einer Fahrt. Daher wird für jede Zeile aus der Fahrtentabelle eine neue Fahrt im Datenmodell mit den in der Zeile enthaltenen Informationen angelegt. Eventuell noch nicht bekannte Betriebsstellen, beim Einfügen des Start- und Zielbahnhofs der Fahrt, werden dabei dem Datenmodell hinzugefügt. t_{Start} sowie t_{Ziel} werden beim Einfügen der neuen Fahrt in das Datenmodell, um die Datumsinformationen aus der Tabellenzeile ergänzt, da diese im Datenmodell nicht separat vorgesehen sind. Befindet sich t_{Ziel} zeitlich vor t_{Start} der neuen Fahrt, wird davon ausgegangen, dass die Fahrt erst am nächsten Tag endet und t_{Ziel} wird um einen Tag erhöht, damit t_{Ziel} hinter t_{Start} liegt. Ebenso werden zusätzliche, die Linienrichtung kennzeichnende, Zeichen hinter dem Liniennamen entfernt, da sie in dieser Arbeit nicht verwendet werden. Außerdem wird der Typ der neu angelegten Fahrt auf den Wert „normalen Zug“ gesetzt, da es sich dabei um eine im Fahrplan geplante Fahrt handelt.

5.2.3 Fahrplandaten

Die Fahrplandaten, auch Solldaten genannt, enthalten die Informationen, zu welcher Uhrzeit ein Zug an welcher Betriebsstelle ankommen und abfahren soll. Die Solldaten stehen in Form einer Tabelle zu Verfügung. In dieser Tabelle stellt jede Zeile ein eigenes Fahrplan-Ereignis dar. Ein Fahrplan-Ereignis besteht aus:

- der Zugnummer der durchführenden Fahrt
- der RIL100 der Betriebsstelle für dieses Ereignis
- einem Fahrplan-Ereignistyp
- einer Sollzeit, die angibt wann dieses Ereignis laut Fahrplan eintreten soll
- der Startzeit der durchführenden Fahrt
- der Zuggattung der durchführenden Fahrt

Insgesamt existieren fünf verschiedene Fahrplan-Ereignistypen, die in den Fahrplandaten enthalten sind:

1. die Abfahrt am Startbahnhof eines Zuges
2. die Ankunft eines Zuges an einer Betriebsstelle
3. die Abfahrt eines Zuges an einer Betriebsstelle
4. die Ankunft am Zielbahnhof eines Zuges
5. die Durchfahrt eines Zuges an einer Betriebsstelle

Vorgehen

Aus den Solldaten werden die Halt-Objekte der einzelnen Fahrten im Datenmodell generiert. Dafür werden die Solldaten-Ereignisse aus dem Fahrplan zuerst nach der Zugnummer sowie der Startzeit der durchführenden Fahrt gruppiert. Daraufhin wird für jede dieser Haltegruppen, die jeweils zu einer Fahrt gehören, das Startereignis und das Zielereignis gesucht. Das Startereignis ist entweder der Halt mit dem Fahrplan-Ereignistyp „Abfahrt am Startbahnhof“ oder, falls ein solches Ereignis nicht vorliegt, das Abfahrt-Ereignis mit der frühesten Sollzeit. Analog dazu ist das Zielereignis entweder das Ereignis mit dem Fahrplan-Ereignistyp „Ankunft am Zielbahnhof“ oder, falls ein solches Ereignis nicht vorliegt, das Ankunft-Ereignis mit der spätesten Sollzeit.

Anschließend wird die Fahrt im Datenmodell gesucht, die alle Halte in der Haltegruppe anfährt. Dafür werden zuerst alle Fahrten aus dem Datenmodell selektiert, die dieselbe Zugnummer besitzen, wie die Fahrplan-Ereignisse in der Haltegruppe. Als nächstes wird nach der Fahrt gesucht, deren Startzeit t_{Start} vor der Sollzeit des Startereignisses und deren Zielzeit t_{Ziel} nach der Sollzeit des Zielereignisses liegt. Konnte keine Fahrt im Datenmodell gefunden werden, die diese Bedingungen erfüllt, wird ein neues Fahrtobjekt mit der Zugnummer, dem Startbahnhof und der Startzeit aus dem Startereignis und dem Zielbahnhof und der Zielzeit aus dem Zielereignis erzeugt.

Daraufhin werden die Ereignisse einer Haltegruppe nach der Betriebsstelle der Ereignisse gruppiert, um jeder angefahrenen Betriebsstelle der Fahrt ein oder zwei Fahrplan-Ereignisse zuzuordnen.

Wenn nur ein Ereignis zu einer Betriebsstelle vorliegt, handelt es sich entweder um das Ereignis für die Abfahrt vom Startbahnhof, die Ankunft am Zielbahnhof, eine normale Abfahrt ohne vorherige Ankunft,

eine normale Ankunft ohne anschließende Abfahrt oder eine Durchfahrt. Abhängig vom Ereignistyp werden die folgenden Halte erzeugt:

- handelt es sich um ein „Abfahrt am Startbahnhof“-Ereignis, wird ein neuer Halt mit der Sollzeit des Ereignisses als $t_{\text{soll abfahrt}}$ und dem Halttyp „Starthalt“ für die Fahrt im Datenmodell erzeugt
- handelt es sich um ein normales Abfahrts-Ereignis an einer Betriebsstelle, wird ein neuer Halt mit der Sollzeit des Ereignisses als $t_{\text{soll abfahrt}}$ und dem Halttyp „Zwischenhalt“ erzeugt
- handelt es sich um ein „Ankunft am Zielbahnhof“-Ereignis, wird ein neuer Halt mit der Sollzeit des Ereignisses als $t_{\text{soll ankunft}}$ und dem Halttyp „Zielhalt“ erzeugt
- handelt es sich um ein normales Ankunfts-Ereignis an einer Betriebsstelle, wird ein neuer Halt mit der Sollzeit des Ereignisses als $t_{\text{soll ankunft}}$ und dem Halttyp „Zwischenhalt“ erzeugt
- handelt es sich um ein Durchfahrt-Ereignis wird ein neuer Halt mit dem Wert „Durchfahrt“ für Fahrtryp_{soll} und der Sollzeit des Ereignisses als $t_{\text{soll ankunft}}$ und $t_{\text{soll abfahrt}}$ erzeugt

Liegen zwei Ereignisse zu einer Betriebsstelle vor, beschreiben diese das Ankunftsereignis und das Abfahrtsereignis an einem Zwischenhalt der Fahrt. In diesem Fall wird für beide Ereignisse zusammen ein Halt im Datenmodell, mit der dazugehörigen Ankunfts- und Abfahrtszeit $t_{\text{soll ankunft}}$ und $t_{\text{soll abfahrt}}$ und dem Halttyp „Zwischenhalt“, angelegt.

Gelegentlich sind auch mehr als zwei Ereignisse für eine Betriebsstelle einer Fahrt vorhanden, zum Beispiel, wenn die Betriebsstelle mehrfach von einer Fahrt angefahren wird. In diesem Fall wird für jedes zeitlich nah beieinanderliegende Ereignis-Paar, bestehend aus einem Ankunfts-Ereignis und einem Abfahrts-Ereignis, ein eigener Halt im Datenmodell erzeugt. Eine weitere Ausnahme zu dem oben genannten Vorgehen ist, wenn die Fahrt an der gleichen Betriebsstelle endet, an der sie begonnen hat. Dies ist der Fall, wenn das „Abfahrt am Startbahnhof“-Ereignis der Fahrt die gleiche Betriebsstelle enthält, wie das „Ankunft am Zielbahnhof“-Ereignis. In dieser Situation werden zwei Halt-Objekte im Datenmodell erzeugt, ein Halt mit Halttyp „Starthalt“ sowie ein Halt mit Halttyp „Zielhalt“.

5.2.4 Zusatzzüge (Sonder-, Ersatz- und Zusatzzüge)

Neben den Fahrten aus dem normalen Fahrplan liegt auch eine Tabelle mit Sonder-, Ersatz- und Zusatzzügen (im Folgenden nur noch Zusatzzüge genannt) der insgesamt drei Betriebstage von der Deutschen Bahn vor. Diese Tabelle enthält eine Liste von Ereignissen der eingesetzten Zusatzzüge, die ähnlich zu den Einträgen aus der Fahrplantabelle aufgebaut sind. Jeder Eintrag der Zusatzzug-Tabelle enthält die folgenden Informationen:

- der RIL100 der Betriebsstelle, an der der Halt stattfinden soll
- die Zugnummer der Fahrt, die diesen Halt durchführen soll
- die Zuggattung der Fahrt
- der Ereignistyp, der angibt, um was für ein Ereignis es sich handelt. Die Ereignistypen hier sind identisch zu denen aus der Fahrplantabelle
- einer Sollzeit, die angibt, wann dieses Ereignis eintreten soll
- die Zugart, der Fahrt die diesen Halt durchführen soll. Dabei handelt es sich entweder um einen Ersatzzug, einen Zusatzzug oder einen Sonderzug
- die Eingangszeit des Ereignisses

Vorgehen

Das Vorgehen beim Einlesen der Zusatzzug-Tabelle ist ähnlich zu dem Vorgehen beim Einlesen der Fahrplandaten. Zuerst werden alle Sonderhalt-Einträge ausgefiltert, für die bereits ein Halt im Datenmodell mit der gleichen Zugnummer, der gleichen Sollzeit und der gleichen Betriebsstelle existieren. Diese Filterung wird durchgeführt, um Verwechslungen von Halten in späteren Einleseschritten, wie dem Einlesen der Istdaten, zu verhindern. Anschließend werden die verbleibenden Einträge der Tabelle nach ihrer Zugnummer und Eingangszeit gruppiert. Für jede dieser erzeugten Zuggruppen wird ein neues Fahrt-Objekt im Datenmodell erzeugt, dem die Zugnummer, die Gattung und die Zugart der betrachteten Zuggruppe als Typ zugewiesen werden. Anschließend wird das Start-Ereignis sowie das Ziel-Ereignis der Zuggruppe gesucht, um B_{Start} , B_{Ziel} , t_{Start} , t_{Ziel} und die Richtung der neuen Fahrt zu bestimmen.

Das Start-Ereignis der Zuggruppe ist das Ereignis mit den Ereignistyp „Abfahrt am Startbahnhof“. Analog dazu ist das Ziel-Ereignis, das Ereignis mit dem Ereignistyp „Ankunft am Zielbahnhof“. Aus dem Start-Ereignis und dem Ziel-Ereignis werden anschließend der Startbahnhof B_{Start} , der Zielbahnhof B_{Ziel} sowie die Startzeit t_{Start} und die Zielzeit t_{Ziel} für die neue Fahrt entnommen. Konnte kein Start-Ereignis für einen Zusatzzug gefunden werden, bleiben die Felder B_{Start} und t_{Start} im neu erstellten Fahrtobjekt leer. Genauso bleiben die Felder B_{Ziel} und t_{Ziel} leer, wenn kein Zielereignis für den Zusatzzug gefunden wurde. Die Fahrtrichtung des neuen Fahrtobjekts wird nur eingetragen, wenn sowohl ein Start-Ereignis als auch ein Ziel-Ereignis für den Zusatzzug gefunden wurde.

Das übrige Vorgehen für die Erzeugung der neuen Halt-Objekte im Datenmodell ist analog zu dem Vorgehen beim Einlesen der Fahrplandaten und wird dort genauer erklärt.

5.2.5 Istdaten

Neben den Solldaten sind auch Istdaten zu den Fahrplandaten in einer Tabelle gegeben. Genau wie bei den Solldaten zuvor stellt auch jede Zeile der Istdaten-Tabelle ein Istdaten-Ereignis dar. Ein Istdaten-Ereignis besteht aus:

- der Zugnummer der durchführenden Fahrt
- der RIL100, der Betriebsstelle an der das Ereignis durchgeführt wurde
- einem Istdaten-Ereignistyp
- der Istzeit, wann das Ereignis eingetreten ist
- der Sollzeit, wann das Ereignis laut Fahrplan eintreten soll
- der Eingangszeit, wann das Ereignis eingegangen ist
- der Zuggattung der durchführenden Fahrt

Analog zu den Fahrplandaten-Ereignissen hat jedes Istdaten-Ereignis einen Ereignistyp, der angibt, wie das Ereignis zu interpretieren ist. Die Istdaten-Ereignisse sind analog zu den Solldaten-Ereignissen definiert:

1. die Abfahrt am Startbahnhof eines Zuges
2. die Ankunft eines Zuges an einer Betriebsstelle
3. die Abfahrt eines Zuges an einer Betriebsstelle
4. die Ankunft am Zielbahnhof eines Zuges
5. die Durchfahrt eines Zuges an einer Betriebsstelle

Vorgehen

Die Istdaten-Ereignisse sind als Ergänzung zu den Fahrplandaten-Ereignissen zu betrachten. Daher wird beim Einlesen der Istdaten-Tabelle für jedes Istdaten-Ereignis nach einem korrespondierenden Halt-Objekt im Datenmodell gesucht. Ob ein Halt zu einem Istdaten-Ereignis korrespondiert ist abhängig von dessen Ereignistyp. Handelt es sich bei dem Ereignis um ein Abfahrtsereignis, korrespondiert ein Halt mit dem Ereignis, wenn die im Ereignis angegebene Sollzeit identisch zu der im Halt angegebenen $t_{\text{soll abfahrt}}$ ist. Analog dazu korrespondiert ein Ankunftsereignis mit einem Halt, wenn die im Ereignis angegebene Sollzeit identisch zu der im Halt angegebenen $t_{\text{soll ankunft}}$ ist. Wurde ein korrespondierender Halt zu einem Istdaten-Ereignis gefunden, wird anschließend abhängig von dem Ereignistyp die Istdaten in den gefundenen Halt übernommen. Handelt es sich um ein Abfahrtsereignis, wird die Ereignis-Istzeit als $t_{\text{ist abfahrt}}$ in den gefundenen Halt eingetragen. Handelt es sich hingegen um ein Ankunftsereignis, wird die Ereignis-Istzeit als $t_{\text{ist ankunft}}$ in den gefundenen Halt übernommen.

Ein Durchfahrt-Ereignis korrespondiert mit einem Halt, wenn die im Ereignis angegebene Sollzeit entweder der im Halt angegebenen $t_{\text{soll abfahrt}}$ oder $t_{\text{soll ankunft}}$ entspricht. Wurde ein zu einem Durchfahrt-Ereignis korrespondierender Halt gefunden, wird die Ereignis-Istzeit sowohl als $t_{\text{ist abfahrt}}$ als auch als $t_{\text{ist ankunft}}$ in den gefundenen Halt eingetragen. Außerdem wird für den Wert $\text{Fahrtyp}_{\text{ist}}$ des gefundenen Halts „Durchfahrt“ eingetragen.

Konnte kein korrespondierender Halt im Datenmodell für ein Ereignis gefunden werden, wird dieses verworfen.

5.2.6 Regelwartezeiten und Übergangszeiten

Neben den Betriebsstellen, den Fahrten, den Fahrplandaten sowie den dazugehörigen Istdaten, ist in dem Datenpaket auch jeweils eine Tabelle mit den Regelwartezeiten und Übergangszeiten enthalten. Die Regelwartezeiten der Abbringer auf die jeweiligen Zubringer sowie die Übergangszeiten an den einzelnen Betriebsstellen müssen aus einem PDF-Dokument der DB Netz AG [2] extrahiert werden. Diese PDF-Datei enthält, in mehrere Abschnitte unterteilt, die Übergangszeiten der einzelnen Betriebsstellen, die Regelungen für die Bestimmung der allgemeinen Regelwartezeiten für Anschlüsse sowie eine Liste mit Abweichungen von den Regelwartezeiten.

Um die, in dem PDF-Dokument, enthalten Tabellen am besten verarbeiten zu können, wird das PDF-Dokument im ersten Schritt mithilfe der Software Adobe Acrobat Pro in ein Excel-Dokument konvertiert. Wird das PDF-Dokument nicht zuerst in ein Excel-Dokument konvertiert, kann auf die Zellen der enthaltenen Tabellen nur schwer zugegriffen werden. Der Grund dafür ist, dass beim Einlesen einer Tabelle aus einem PDF-Dokument in der Regel alle Spalten- und Zeilen-Trennlinien entfallen. Dadurch ist es nicht mehr möglich, die Zugehörigkeit von Wörtern, oder sogar ganzen Zeilen, zueinander zu bestimmen. Nach der Konvertierung in ein Excel-Dokument bleiben diese Zusammenhänge erhalten, wodurch sich das Dokument dank guter Programmbibliotheken, einfach einlesen lässt.

Aachen Hbf	6	3	
Aalen	4	3	
Achern	5	3	
Ahrensfelde	4	3	
Albstadt-Ebingen	3	3	
Altenbeken	5	3	
Altenburg	4	3	
Altenkirchen	3	3	
Alzey	4	3	
Andernach	6	4	
Angermünde	3	3	
Ansbach	5	3	
Appenweier	6	3	
Armsheim	4	3	
Arnsdorf (b Dresden)	5	3	
Arnstadt Hbf	4	3	
Aschaffenburg	6	4	
Aschersleben	4	3	
Au (Sieg)	4	3	
Aue (Sachs)	4	3	

Abbildung 5.1: Ausschnitt der Übergangstabelle aus [2].

Einlesen der Übergangszeiten

Die Übergangszeiten sind in einer Tabelle angeordnet. Jede Zeile dieser Tabelle enthält die Übergangszeit-Informationen zu einer Betriebsstelle. Diese Informationen sind:

- der Name der betrachteten Betriebsstelle
- die Übergangszeit für bahnsteiggleiche Übergänge in Minuten
- die Übergangszeit für nicht bahnsteiggleiche Übergänge in Minuten
- ein optionaler Kommentar für die Übergangszeit

Die Abbildung 5.1 zeigt einen Ausschnitt der Übergangszeit-Tabelle.

Vorgehen: Wie auch zuvor die anderen Tabellen, wird jede Zeile der Übergangszeit-Tabelle nacheinander eingelesen. Dazu wird für jede Zeile aus der Übergangszeit-Tabelle zuerst die dazugehörige Betriebsstelle aus dem Datenmodell gesucht. Die zu einer Zeile gehörende Betriebsstelle ist die Betriebsstelle, deren Name gleich dem Namen der Betriebsstelle in der betrachteten Zeile der Übergangszeit-Tabelle ist. Anschließend wird die Übergangszeit sowohl für bahnsteiggleiche Übergänge als auch für nicht bahnsteiggleiche Übergänge aus der Übergangszeit-Zeile, in die Felder $t_{\text{mindestübergangszeit bg}}$ und $t_{\text{mindestübergangszeit}}$ der gefundenen Betriebsstelle übertragen.

Einlesen der Regelwartezeiten

Genau wie die Übergangszeiten sind auch die Regelwartezeiten in Form einer Tabelle in [2] angeordnet. Abbildung 5.2 zeigt einen Ausschnitt der in [2] enthaltenen Regelwartezeit-Tabelle. Diese Tabelle besteht aus insgesamt 6x6 Feldern, wobei die oberste Zeile sowie die erste Spalte der Tabelle Zubringer- und Abbringer-Informationen über die eingetragenen Regelwartezeiten enthalten. Die restlichen 5x5 Felder enthalten die eigentlichen Regelwartezeiten in Minuten. Jede Zeile dieser Tabelle beschreibt die möglichen Regelwartezeiten einer Gruppe von Abbringer-Gattungen, die in der ersten Zelle der Zeile

Es kann warten in Minuten	auf				
	EC, ICE, IC, TGV, Thalys, Railjet	EN, NZ, CNL, AZ, D	DPF, IRE, RE, RB	S-Bahn	DPE, DPN Übrige Züge
EC, ICE, IC, TGV, Thalys, Railjet 1) 2) 3)	3	0	0	0	0
EN, NZ, CNL, AZ, D 4)	5	5	0	0	5
DPF, IRE, RE, RB 5)	0	0	0	0	0
S-Bahn	0	0	0	0	0
DPE, DPN Übrige Züge	5	5	5	0	5

Abbildung 5.2: Tabelle mit den Regelwartezeiten für Anschlussfahrten.

definiert werden. Jede Spalte der Tabelle beschreibt die Regelwartezeiten einer Gruppe von Zubringer-Gattungen, die in der obersten Zelle definiert werden. Um eine Regelwartezeit zu einem Anschluss zu ermitteln, wird nach der Zeile gesucht, die in der ersten Spalte die Zuggattung des Abbringers enthält. Anschließend wird in dieser Zeile nach der Spalte gesucht, die in der ersten Zeile dieser Spalte die Zuggattung des Zubringers enthält. Das so ermittelte Feld enthält die Regelwartezeit des Anschlusses in Minuten.

Vorgehen: Die Regeln der Regelwartezeiten werden per Hand in das Datenmodell übernommen. Der Zugriff auf die Regelwartezeiten erfolgt anschließend in Form einer Lookup-Tabelle, in der die Regelwartezeiten einer Kombination aus einer Zubringer- und einer Abbringer-Gattung zugeordnet werden. Wird nach der Regelwartezeit eines Anschlusses gesucht, wird die Regelwartezeit aus der Lookup-Tabelle zurückgegeben, deren Gattungspaar der Zubringergattung und AbbringerGattung des Anschlusses entspricht. Kann keine solche Gattungs-Kombination für einen Anschluss in der Lookup-Tabelle gefunden werden, wird eine Regelwartezeit, wie in Abbildung 5.2 beschrieben, von 5 Minuten zurückgegeben.

Einlesen der Abweichungen von den Regelwartezeiten

Zusätzlich zu den allgemeinen Regelwartezeiten, existiert eine weitere Tabelle mit Abweichungen von den Regelwartezeiten. Beim Einlesen dieser Tabelle mit Abweichungen kann nicht jede Zeile alleine betrachtet werden, wie bei den vorherigen Tabellen. Der Grund dafür ist, dass eine Zeile der Abweichungen-Tabelle nicht, wie bei den vorherigen Tabellen, alle benötigten Informationen zu einer Regelwartezeit enthält. Die Informationen für eine Regelwartezeit-Abweichung sind über mehrere Zeilen verteilt. Abbildung 5.3 zeigt einen Ausschnitt der Tabelle mit Abweichungen von den Regelwartezeiten mit den Abweichungen für den Hauptbahnhof Darmstadt.

In der Tabelle sind die Abweichungen nach dem jeweils betroffenen Anschlussbahnhof gruppiert, auf den sich die Abweichungen beziehen. Innerhalb einer solchen Gruppe, sind die Abweichungen eines Anschlussbahnhofs wiederum nach der Fahrtrichtung der betroffenen Abbringer gruppiert. Alle Abweichungen, die für beide Fahrtrichtungen gelten, sind in einer separaten Gruppe enthalten. Für jede an einem Anschlussbahnhof angefahrne Fahrtrichtung sind die Abweichungen in Form einer Liste angegeben. Diese Liste enthält für jede Regelwartezeiten-Abweichung Informationen zu den davon betroffenen Abbringern und Zubringern sowie die dafür anzuwendende Regelwartezeit. Die Informationen zu den betroffenen Abbringern und Zubringern bestehen entweder aus einer Liste mit Zugnummern, Zuglinien

Darmstadt Hbf		
▶ Ri Aschaffenburg		
15727	15368	10
15731	15273	3
▶ Ri Bensheim		
15377, 15397	15735, 15737	10
▶ Ri Mannheim		
15378	15737	10
▶ Ri Groß-Umstadt – Wiebelsbach		
25227	38600, 35307	3
25247	15370, 15393, 15371, 35365	10
25249	15374, 15390, 15375, 35373	10
▶ Ri Pfungstadt		
25126	15395, 15734, 35369, 15373, 1976, 15732, 15731, 15777, 1956	10

Abbildung 5.3: Ausschnitt aus der Liste der Abweichungen von den Regelwartezeiten für den Hauptbahnhof Darmstadt.

oder Zuggattungen. Zusätzlich können darin auch zeitliche Einschränkungen angegeben werden, wie Tage oder Uhrzeiten an denen die Regelwartezeit gilt. Für die Angabe der betroffenen Zubringer kann außerdem die Angabe „alle Zubringer“ gemacht werden. Diese Angabe bedeutet, dass der Abbringer auf alle Züge in die zuvor angegebene Richtung wartet. Wurde für eine Regelwartezeiten-Abweichung keine Richtung angegeben, gilt die Regelwartezeit für beide Fahrtrichtungen der in der Regel angegebenen Abbringer.

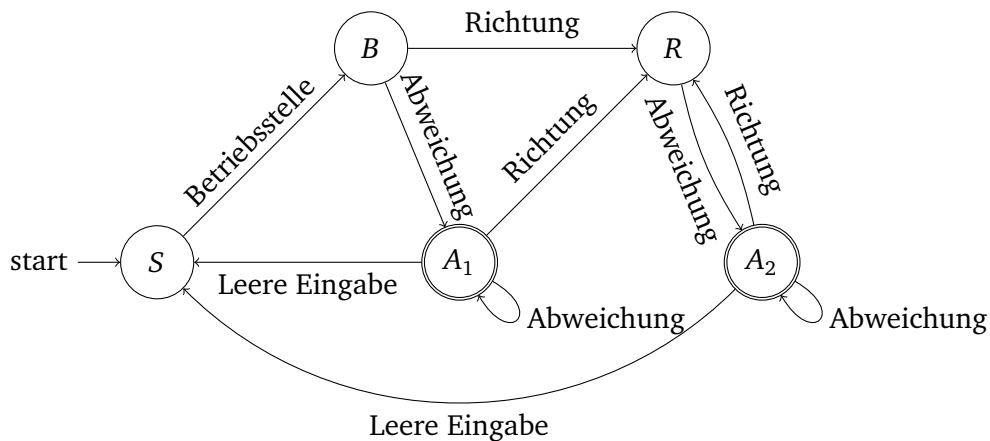


Abbildung 5.4: Zustandsautomat für den Einlesevorgang der Abweichungen von den Regelwartezeiten.

Vorgehen beim Einlesen: Die Abbildung 5.4 zeigt das Vorgehen beim Einlesen der Regelwartezeiten-Abweichungen als Zustandsautomat. Der Zustandsautomat liest dabei nacheinander alle Zeilen der Abweichungen-Tabelle ein und wechselt abhängig vom Inhalt der aktuellen Zeile seinen Zustand.

Der Automat beginnt den Einlesevorgang im Zustand *S*, am Anfang der Tabelle. Anschließend muss die erste Zeile der Tabelle eine Betriebsstelle enthalten, woraufhin der Automat in den Zustand *B* wechselt.

Nach dem Eintritt in Zustand *B* muss der Automat in der nächsten Zeile entweder eine Richtungsangabe oder direkt eine Abweichung lesen. Findet der Automat eine Richtungsangabe, wechselt er in den Zustand *R*. Findet der Automat eine Abweichung, wechselt er in den Zustand *A*₁. Dies bedeutet, dass nach

Eintritt in Zustand A_1 in den nächsten Zeilen die Abweichungen folgen, für die es keinen Unterschied macht, in welche Richtung die Abbringer fahren.

Befindet sich der Automat in Zustand A_1 existieren drei Möglichkeiten für den Inhalt der nächsten Zeile. Zum einen kann in der nächsten Zeile eine weitere Regelwartezeit-Abweichung enthalten sein, wodurch der Automat in Zustand A_1 bleibt. Alternativ kann der nächsten Zeile eine Richtungsangabe enthalten sein. Dies bedeutet, dass die Abbringer der folgenden Abweichungen in die gelesene Richtung fahren. Wird eine Richtungsangabe aus Zustand A_1 eingelesen, wechselt der Automat in Zustand R . Zuletzt kann der Automat auch eine leere Zeile einlesen. Das Einlesen einer leeren Zeile bedeutet, dass die Abweichungen für die gerade betrachtete Betriebsstelle fertig eingelesen wurden, woraufhin der Automat in den Zustand S wechselt.

Befindet sich der Automat in Zustand R muss er in der nächsten gelesenen Zeile eine Regelwartezeit-Abweichung vorfinden. Die in der nächsten Zeile enthaltene Abweichung gilt für die dort genannten Abbringerzüge, die in die zuvor eingelesene Richtung fahren, am zuvor eingelesenen Anschlussbahnhof. Anschließend wechselt der Automat in den Zustand A_2 .

Aus Zustand A_2 kann der Automat entweder im Zustand A_2 durch das Einlesen einer weiteren Regelwartezeit-Abweichung verbleiben, oder durch das Einlesen einer leeren Zeile in den Startzustand S zurückkehren. Findet der Automat in der betrachteten Zeile eine weitere Abweichung, gilt diese ebenfalls für den zuvor eingelesenen Anschlussbahnhof für die in der Zeile genannten Abbringerzüge, die in die zuvor gelesene Richtung fahren. Handelt es sich bei der aktuell betrachteten Zeile hingegen um eine leere Zeile, wurden für den zuvor betrachteten Anschlussbahnhof alle Abweichungen gelesen. Woraufhin die Regelwartezeit-Abweichungen für den nächsten Anschlussbahnhof folgen.

Terminieren kann der Automat entweder in Zustand A_1 oder A_2 , nachdem mindestens eine Regelwartezeit-Abweichung für den aktuell betrachteten Anschlussbahnhof eingelesen wurde.

Finden einer Regelwartezeit zu einem Anschluss: Nachdem die Abweichungen-Tabelle eingelesen wurde, kann in den Regelwartezeit-Abweichungen nach der Regelwartezeit zu Anschlüssen gesucht werden. Die Suche nach der Regelwartezeit zu einem Anschluss A erfolgt nach dem folgenden Schema:

1. Zuerst werden alle zuvor eingelesenen Regelwartezeit-Abweichungen auf die Abweichungen reduziert, deren Betriebsstelle gleich dem Anschlussbahnhof von A ist.
2. Anschließend werden die gefundenen Regelwartezeit-Abweichungen nach der Richtung des Abbringers von A gefiltert.
3. Daraufhin wird in den gefundenen Regelwartezeit-Abweichungen nach jenen gesucht, die den Abbringer von A betreffen.
4. Je nach der Art der angegebenen Informationen des Abbringers kann dies entweder aufgrund der Gattung, der Zugnummer oder der Linie des Abbringers erfolgen.
5. Zuletzt werden die noch vorhandenen Regelwartezeit-Abweichungen auf die Abweichungen, die den Zubringer von A betreffen, reduziert.

Ist nach einem der obigen Schritte keine Regelwartezeit-Abweichung mehr für A möglich, wird in der Lookup-Tabelle der allgemeinen Regelwartezeiten nach der Regelwartezeit für A gesucht.

Nicht verarbeitbare Abweichungen von den Regelwartezeiten: Von allen eingelesenen Regelwartezeit-Abweichungen lassen sich allerdings nur insgesamt circa 80% vollständig verarbeiten. Der Grund dafür ist, dass einige Abweichungen Informationsformate für die Angabe der Abbringer und Zubringer verwenden, die nur selten auftreten. Für solche Regelwartezeit-Abweichungen wurde daher kein Sonderverhalten implementiert um sie verarbeiten zu können. Zu den verarbeitbaren Regelwartezeit-Abweichungen gehören die Abweichungen, deren Abbringer und Zubringer Informationen aus einer oder mehreren Zugnummern, Linien oder Gattungen bestehen. Ebenso können die Regelwartezeit-Abweichungen verarbeitet werden, die für alle Zubringer und/oder Abbringer gelten. Auch Regelwartezeit-Abweichungen, die nur für eine Wochentagauswahl gelten, wie zum Beispiel Wochenendfahrten können verarbeitet werden.

5.2.7 Anschlüsse und Anschlusskonflikte

Die Anschlüsse liegen in zwei Tabellen der Deutschen Bahn AG vor. Die erste Anschluss-Tabelle wurde aus RIS exportiert, während die zweite Anschluss-Tabelle aus ISTP exportiert wurde. Beide Tabellen sind ähnlich zu den vorherigen Tabellen aufgebaut. Dies bedeutet, dass jede Zeile aus einer der Tabellen die Informationen zu einem Anschluss enthält.

Einlesen der RIS-Anschlussstabelle

Die RIS-Anschlussstabelle enthält zu jedem Anschluss die folgenden Informationen:

- der RIL100 des Anschlussbahnhofs
- die Zugnummer des Anschlusszubringers
- die Gattung des Zubringers
- der Ereignistyp, im Fahrplan des Zubringers am Anschlussbahnhof
- die Sollankunftszeit des Zubringers
- die Zugnummer des Anschlussabbringers
- die Gattung des Abbringers
- die Sollabfahrtszeit des Abbringers
- der Ereignistyp, im Fahrplan des Abbringers am Anschlussbahnhof
- die Eingangszeit der Anschlussmeldung
- der Anschlussstatus, ob der Anschluss gehalten wurde, neu ist oder nicht gehalten wurde
- die Angabe ob es sich bei dem gelesenen Anschluss um einen Alternativanschluss handelt

Vorgehen: Beim Einlesen der RIS-Anschlussstabelle wird für jeden enthaltenen Anschluss zuerst die Anschlussbetriebsstelle im Datenmodell gesucht, an der der Anschluss auftritt. Anschließend werden die Abbringer- und Zubringer-Halte am Anschlussbahnhof gesucht. Der Anschlusshalt des Zubringers ist der Halt H aus dem Datenmodell, für den die folgenden Bedingungen gelten:

- die Zugnummer der Fahrt von H ist gleich der Zubringerzugnummer in der Anschlusszeile
- die Betriebsstelle von H ist gleich der Anschlussbetriebsstelle in der Anschlusszeile
- $t_{\text{soll ankunft}}$ von H ist gleich der Sollankunftszeit des Zubringers in der Anschlusszeile

Der Anschlusshalt des Abbringers ist der Halt H aus dem Datenmodell, für den die folgenden Bedingungen gelten:

- die Zugnummer der Fahrt von H ist gleich der Abbringerzugnummer in der Anschlusszeile
- die Betriebsstelle von H ist gleich der Anschlussbetriebsstelle in der Anschlusszeile
- $t_{\text{soll abfahrt}}$ von H ist gleich der Sollabfahrtszeit des Abbringers in der Anschlusszeile

Über die Anschlusshalte der Abbringer- und der Zubringerfahrt kann direkt die dazugehörige Abbringer- und Zubringerfahrt herausgefunden werden. Konnte der Zubringerhalt oder der Abbringerhalt auf diese Weise nicht gefunden werden, wird stattdessen direkt nach der Fahrt gesucht, die den Halt durchführt. Konnte der Zubringerhalt nicht gefunden werden, werden alle Fahrten mit der im Anschluss angegebenen Zubringer-Zugnummer gesucht. Anschließend wird die Fahrt als Zubringerfahrt übernommen, die während der im Anschluss angegebenen Ankunfts-Sollzeit des Zubringers unterwegs ist. Für die Identifikation des fehlenden Zubringerhalts wird in den Halten der gefundenen Fahrt nach einem Halt gesucht, der am Anschlussbahnhof durchgeführt wird. Konnte ein Halt gefunden werden, der am Anschlussbahnhof durchgeführt wird, wird er als Zubringerhalt des Anschlusses übernommen. Das gleiche Vorgehen wird für die Abbringerfahrt durchgeführt, falls zuvor kein Abbringerhalt gefunden werden konnte.

Aus den gefundenen Objekten des Datenmodells wird anschließend ein neues Anschlussobjekt A im Datenmodell erzeugt. Abhängig von dem Anschluss-Status aus der Anschlusszeile und der Angabe, ob es sich bei dem Anschluss in der Anschlusszeile um einen Alternativanschluss handelt, wird der Status von A gesetzt:

- handelt es sich um einen neuen Anschluss und die Angabe zum Alternativanschluss ist vorhanden, wird der Status von A auf „Alternativanschluss“ gesetzt
- handelt es sich um einen gehaltenen Anschluss, wird der Status von A auf „Gehalten“ gesetzt
- handelt es sich um einen nicht gehaltenen Anschluss, wird der Status von A auf „nicht Gehalten“ gesetzt

Zuletzt wird die Eingangszeit der Anschlussmeldung aus der Anschlusszeile als Eingangszeit t_{Eingang} in A übernommen.

Konnte für einen Anschluss kein Zubringer oder Abbringer im Datenmodell gefunden werden, kann dieser Anschluss nicht eingelesen werden und wird verworfen. Durch dieses Vorgehen werden 86% der Anschlusseinträge aus der RIS-Anschlusstabelle in das Datenmodell eingelesen.

Einlesen der ISTP-Anschlusstabelle

Die ISTP-Anschlusstabelle enthält zu jedem Anschluss die folgenden Informationen:

- der RIL100 des Anschlussbahnhofs
- die Zugnummer des Anschlusszubringers
- die Gattung des Zubringers
- die Soll-Ankunftszeit des Zubringers
- die Abweichung der Ist-Ankunftszeit des Zubringers von seiner Soll-Ankunftszeit zur Erstellungszeit der Anschlussmeldung in Minuten
- die Zugnummer des Anschlussabbringers
- die Soll-Abfahrtszeit des Abbringers

-
- das Sollgleis, an dem der Abbringer laut Fahrplan am Anschlussbahnhof halten soll
 - das Istgleis, an dem der Abbringer tatsächlich am Anschlussbahnhof gehalten hat
 - die Erstellungszeit der Anschlussmeldung
 - der Anschlussstatus, ob der Anschluss gehalten wurde, neu ist oder nicht gehalten wurde
 - eine Informationsmeldung zu dem Anschluss, der an die Kunden weitergegeben wurde

Bei der ISTP-Anschlusstabelle ist zu beachten, dass sie auch Anschlussduplikate enthält. Einer der Gründe dafür ist, dass für jede Änderung eines Anschlusses, wie Änderungen an Informationsmeldungen, ein neuer Eintrag in der ISTP-Anschlusstabelle enthalten ist. Weil diese Änderungen nur in dieser Arbeit nicht betrachtete Informationen betreffen, werden beim Einlesen der ISTP-Anschlusstabelle zuerst alle veralteten Anschluss-Einträge ausgefiltert. Damit wird sichergestellt, dass zum einen jeder Anschluss exakt einmal erfasst wird und zum anderen, dass nur der aktuellste Zustand des Anschlusses erfasst wird. Dafür werden die Anschlusseinträge zuerst nach einer Kombination bestehend aus dem Bahnhof, der Zubringerzugnummer, der Abbringerzugnummer, der Ankunftszeit des Zubringers sowie der Abfahrtszeit des Abbringers gruppiert. Anschließend werden die Einträge jeder erstellten Gruppe nach der Erstellungszeit der Meldungen gruppiert. Daraus wird der erste Anschlusseintrag der Gruppe mit der spätesten Erstellungszeit behalten.

Vorgehen: Der Einlesevorgang der ISTP-Anschlusstabelle erfolgt weitestgehend analog zum Einlesevorgang der RIS-Anschlusstabelle. Der Unterschied beim Einlesen der ISTP-Anschlusstabelle im Vergleich zur RIS-Anschlusstabelle ist, dass hier davon ausgegangen wird, dass es sich bei einem Anschluss um einen Alternativanschluss handelt, wenn der aktuelle Eintrag für den Anschlussstatus „neu“ ist. Davon abgesehen ist der Einlesevorgang der Anschlüsse aus der ISTP-Anschlusstabelle analog zu dem Einlesevorgang der RIS-Anschlusstabelle. Beim Einlesen der Anschlüsse aus der ISTP-Anschlusstabelle fallen insgesamt rund 65% der Einträge aus der Tabelle weg, da für sie entweder keine Informationen für die Zubringer- oder Abbringer-Fahrt vorliegen oder weil die Einträge doppelt enthalten sind.

Erkennen der konfliktierenden Anschlüsse

Die Information, ob ein Konflikt für einen Anschluss vorliegt, und damit ob er zu einem Anschlusskonflikt gehört, lässt sich nicht automatisch aus der gegebenen Anschlusstabelle ablesen. Dies ist darin begründet, dass die Daten keine explizite Zuordnung von Maßnahmen zu Anschlüssen enthalten und damit keine eindeutigen Informationen dazu, bei welchem Anschluss ein Konflikt vorliegt. Aus diesem Grund wird hier angenommen, dass für die Anschlüsse ein Anschlusskonflikt vorliegt, für die eine Maßnahme zur Lösung eines Anschlusskonflikts gefunden wurde.

5.2.8 Maßnahmen

Nachdem die für die Anschlüsse relevanten Informationen eingelesen wurden, müssen die Maßnahmen für die aufgetretenen Anschlusskonflikte identifiziert und hinzugefügt werden. Ebenso wie bei der Erkennung der konfliktierenden Anschlüsse liegt auch für die durchgeführten Dispositions-Maßnahmen keine durch die Disponenten gelabelten Daten vor, die angeben, welche Maßnahmen für einen Anschlusskonflikt durchgeführt wurden. Dies führt dazu, dass die Maßnahmen aufgrund der gegebenen Daten in dieser Arbeit abgeleitet werden müssen. Dabei kann allerdings nicht sichergestellt werden, dass Ereignisse, wie Gleiswechsel oder Wartezeiten, tatsächlich aufgrund einer Dispositionsmaßnahme und nicht aus einem anderen Grund durchgeführt wurden. Die Identifikation der Maßnahmen für die Lösung der Anschlusskonflikte erfolgt je nach Maßnahmentyp unterschiedlich. Die Identifikation der

Gleiswechsel-Maßnahmen erfolgt durch das Einlesen der dazugehörigen Tabelle im Datenpaket. Die übrigen Maßnahmen, müssen wiederum aus den zuvor aufbereiteten Daten des Datenmodells abgeleitet werden.

Neben einzelnen Maßnahmen können auch Maßnahmenkombinationen, wie ein Gleiswechsel des Zubringers und das Warten in Regelwartezeit des Abbringers, für einen Anschlusskonflikt durchgeführt werden. Die Maßnahmenkombinationen werden aufgrund der Einzelmaßnahmen erkannt. Dafür werden alle unabhängig voneinander gefundenen Maßnahmen für einen Anschlusskonflikt zu einer Maßnahmenkombination zusammengefügt.

Alternativanschluss-Maßnahmen

Die Maßnahmen „Alternativanschluss“ werden aus den zuvor eingelesenen Anschlüssen extrahiert, weil für die Anschlüsse beim Einlesen angegeben ist, ob es sich dabei um einen Alternativanschluss handelt. Um die Alternativanschluss-Maßnahmen zu extrahieren, werden zuerst alle Anschlüsse aus dem Datenmodell gesucht, bei denen es sich um Alternativanschlüsse handelt. Bei einem Anschluss A handelt es sich um einen Alternativanschluss, wenn der Status von A „Alternativanschluss“ enthält. Weil für die Alternativanschlüsse in den Daten nicht angegeben ist, für welchen anderen Anschluss der Alternativanschluss ein Alternativanschluss ist, muss nach diesen Anschlüssen gesucht werden. Dafür werden für jeden Alternativanschluss $A_{\text{alternativ}}$ im Datenmodell alle nicht Alternativanschlüsse A_i gesucht, für die folgende Bedingungen gelten:

- der Anschlussstatus von A_i ist „nicht gehalten“
- der Zubringer von $A_{\text{alternativ}}$ ist gleich dem Abbringer von A_i
- die Anschlussbahnhöfe von $A_{\text{alternativ}}$ und A_i sind identisch zueinander

Für jeden gefundenen Anschluss A_i , der diese Bedingungen erfüllt, wird, falls noch nicht vorhanden, ein Anschlusskonflikt mit der Maßnahme „Alternativanschluss“ in das Datenmodell eingefügt. Existiert bereits ein Anschlusskonflikt für A_i im Datenmodell, wird diesem die Maßnahme „Alternativanschluss“ hinzugefügt.

Gleiswechsel-Maßnahmen

Die durchgeführten Gleiswechsel sind in einer Gleiswechseltabelle geführt. Aufgrund der Einträge in dieser Gleiswechseltabelle werden die „Gleiswechsel“-Maßnahmen identifiziert. Diese Tabelle enthält in jeder Zeile einen eigenen Gleiswechsel. Für jeden Gleiswechsel sind die folgenden Informationen in der Tabelle enthalten:

- der RIL100 der Betriebsstelle B , an der der Gleiswechsel stattfindet
- die Zugnummer der Fahrt F , die den Gleiswechsel durchführt
- die Zuggattung der Fahrt, die den Gleiswechsel durchführt
- eine Sollzeit die angibt, wann F an B ankommen oder abfahren soll
- die Typinformation, ob es sich bei der Sollzeit um eine Abfahrtszeit oder eine Ankunftszeit handelt
- das Soll-Gleis G_S
- das Ist-Gleis G_I
- die Eingangszeit der Gleiswechselfeldmeldung

Ein Eintrag in dieser Gleiswechseltabelle bedeutet, dass die Fahrt F an der Betriebsstelle B nicht wie geplant an Gleis G_S , sondern an Gleis G_I gehalten hat.

Da in den Daten nicht angegeben ist, welche Gleiswechsel aufgrund einer Dispositionsentscheidung durchgeführt wurden, wird hier angenommen, dass jede Fahrt die an einem Anschlussbahnhof einen Gleiswechsel durchführt, dies aufgrund einer Gleiswechsel-Maßnahme macht. Dafür werden für jeden Gleiswechsel in der Gleiswechseltabelle alle Anschlüsse aus dem Datenmodell gesucht, die an der gleichen Betriebsstelle des Gleiswechsels stattfinden und an denen die gleiche Fahrt des Gleiswechsels, als Zubringer oder als Abbringer, zur beim Gleiswechsel angegebenen Sollzeit teilnimmt. Abhängig davon, ob es sich bei der gefundenen Fahrt des Gleiswechsels um die Zubringer- oder Abbringer-Fahrt eines gefundenen Anschlusses handelt, wird für den Anschlusskonflikt zu dem Anschluss eine Zubringer- oder eine Abbringergleiswechsel-Maßnahme im Datenmodell erzeugt. Dadurch werden Gleiswechsel, die von einem Zubringer durchgeführt wurden als Zubringergleiswechsel-Maßnahmen dem Datenmodell hinzugefügt, während Gleiswechsel, die von einem Abbringer durchgeführt wurden, als Abbringergleiswechsel-Maßnahmen dem Datenmodell hinzugefügt werden. Zusätzlich wird jedem Anschlusskonflikt, für den ein Zubringergleiswechsel oder ein Abbringergleiswechsel durchgeführt wurde, eine allgemeine Gleiswechsel-Maßnahme hinzugefügt.

Warte-Maßnahmen

Die Warte-Maßnahmen „Warten in Regelwartezeit“, „Warten mit Wartezeitüberschreitung“ und „Weiterfahren ohne Warten“ werden anhand der Fahrplan- und Istdaten erkannt.

Warten in Regelwartezeit: Damit für einen Anschlusskonflikt die Maßnahme „Warten in Regelwartezeit“ getroffen wurde, müssen für den dazugehörigen Anschluss die folgenden Bedingungen gelten:

- die Ankunftsverspätung des Zubringers ist größer 0:
 $t_{\text{zubringer ist ankunft}} - t_{\text{zubringer soll ankunft}} > 0$
- die Abfahrtsverspätung des Abbringers ist kleiner oder gleich der Regelwartezeit des Anschlusses und größer 0: $0 < t_{\text{abbringer ist abfahrt}} - t_{\text{abbringer soll abfahrt}} \leq t_{\text{Regelwartezeit}}$
- die Wartezeit des Abbringers ist größer 0: $t_{\text{abbringer ist haltezeit}} - t_{\text{abbringer min linie haltezeit}} > 0$
- der Anschluss wurde gehalten: $t_{\text{abbringer ist abfahrt}} - t_{\text{zubringer ist ankunft}} \geq t_{\text{mindestübergangszeit}}$

Warten mit Wartezeitüberschreitung: Reicht die Regelwartezeit für den Abbringer zur Anschlusssicherung nicht aus, kann der Abbringer auch mit Wartezeitüberschreitung warten. Damit die Maßnahme „Warten mit Wartezeitüberschreitung“ für einen Anschlusskonflikt getroffen wird, müssen die folgenden Bedingungen von dem Anschluss erfüllt werden:

- die Ankunftsverspätung des Zubringers ist größer 0:
 $t_{\text{zubringer ist ankunft}} - t_{\text{zubringer soll ankunft}} > 0$
- die Abfahrtsverspätung des Abbringers ist größer der Regelwartezeit des Anschlusses:
 $t_{\text{abbringer ist abfahrt}} - t_{\text{abbringer soll abfahrt}} > t_{\text{Regelwartezeit}}$
- die Wartezeit des Abbringers ist größer 0: $t_{\text{abbringer ist haltezeit}} - t_{\text{abbringer min linie haltezeit}} > 0$
- der Anschluss wurde gehalten: $t_{\text{abbringer ist abfahrt}} - t_{\text{zubringer ist ankunft}} \geq t_{\text{mindestübergangszeit}}$

Weiterfahren ohne Warten: Falls die benötigte Wartezeit eines Anschluss-Abbringers auf den dazugehörigen Zubringer sehr hoch ausfällt, ist eine mögliche Alternative die Maßnahme „Weiterfahren ohne Warten“. Wird die Maßnahme „Weiterfahren ohne Warten“ getroffen, fährt der Abbringer ohne zu Warten ab. Die Maßnahme „Weiterfahren ohne Warten“ wurde für einen Anschluss getroffen, wenn die folgenden Bedingungen von dem Anschluss erfüllt sind:

- die Ankunftsverspätung des Zubringers ist größer 0:
 $t_{\text{zubringer ist ankunft}} - t_{\text{zubringer soll ankunft}} > 0$
- der Abbringer hat in etwa so lange gehalten, wie es laut Fahrplan geplant war:
 $t_{\text{abbringer ist haltezeit}} - t_{\text{abbringer soll haltezeit}} \leq 1$
- der Anschluss wurde nicht gehalten: $t_{\text{abbringer ist abfahrt}} - t_{\text{zubringer ist ankunft}} < t_{\text{mindestübergangszeit}}$

Zusatzzug-Maßnahmen

Ähnlich zu den „Alternativanschluss“-Maßnahmen werden auch die „Zusatzzug“-Maßnahmen anhand der bereits eingelesenen Anschlüsse erkannt.

Dafür werden zuerst die Anschlüsse identifiziert, deren Abbringerfahrt den Typ „Sonderzug“ oder „Zusatzzug“ besitzen. Diese Typ-Filterung wird durchgeführt, weil die Ersatzfahrten nicht aufgrund von Anschlussdispositionsmaßnahmen durchgeführt werden. Anschließend werden für jeden gefundenen Zusatzzuganschluss A_{zusatz} die Anschlüsse A_i gesucht, für die die folgenden Bedingungen gelten:

- der Zubringer von A_i ist gleich dem Zubringer von A_{zusatz}
- der Anschlussbahnhof von A_{zusatz} ist gleich dem Anschlussbahnhof von A_i
- mindestens 70% der Betriebsstellen der Folgehalte der Abbringer von A_{zusatz} und A_i sind identisch zueinander

Jedem so gefundenen Anschluss A_i wird anschließend die Maßnahme „Zusatzzug“ zugewiesen.

5.3 Konvertierung des Datenmodells in das Attribute-Relation File Format

Nachdem die Anschlusskonflikte und die dazugehörigen Maßnahmen identifiziert und in das Datenmodell eingelesen wurden, muss daraus ein Modell erzeugt werden. Für die Modellerzeugung wird WEKA eingesetzt. Damit WEKA das erzeugte Datenmodell für die Generierung einer Regelmenge verwenden kann, müssen die Daten aus dem Datenmodell zuvor in eine ARFF-Datei exportiert werden. Diese Datei kann anschließend von Weka als Trainingsset geladen werden.

Die Informationen aus dem Datenmodell werden als Javaobjekte in das Attribute-Relation File Format übertragen.

Ist ein Wert für einen Parameter aus dem Datenmodell nicht bekannt, wird dieser Wert durch ein „?“ im ARFF angegeben. Das „?“ gibt dem Regellerner in WEKA an, dass die erwartete Information für den betrachteten Anschlusskonflikt nicht vorliegt.

Export von Maßnahmenkombinationen

Eine Herausforderung beim Exportieren der Anschlusskonflikte ist deren Eigenschaft, dass sich mehrere Maßnahmen zu deren Lösung kombinieren lassen. Ein Beispiel dafür ist die Kombination einer Gleiswechselmaßnahme mit einer Warte-Maßnahme des Abbringers. Dabei kann ein Gleiswechsel einem Abbringer ermöglichen, schneller einen Anschlussbahnhof zu erreichen. Dies wiederum kann dazu führen, dass der Abbringer eine Wartemaßnahme am Anschlussbahnhof durchführen kann, was den Fahrgästen des Zubringers ermöglicht, trotz Verspätung in den Abbringer umzusteigen.

Da eine Problem Instanz bei der Verwendung eines Regellerners einer eindeutigen Klasse zugeordnet sein muss, ist es notwendig, die zur Lösung eines Anschlusskonflikts eingesetzten Maßnahmen zu einer einzigen Maßnahme zu kombinieren. Dadurch kann für jede Problem Instanz eine eindeutige Klasse definiert werden.

Um dies zu erreichen, wird beim Export der Maßnahmenkombinationen, ähnlich wie beim Exportieren von String-Attributen vorgegangen. Dafür wird zuerst eine Menge aller vorkommenden Maßnahmenkombination, unabhängig der Reihenfolge, in der die Maßnahmen für die Anschlusskonflikte getroffen wurden, erzeugt. Im zweiten Schritt werden die Namen, aller enthaltenen Maßnahmen $m \in \mathbb{M}$ einer Maßnahmenkombination \mathbb{M} konkateniert, und als Enum in die ARFF-Datei geschrieben.

5.4 Exportierte Attribute der Anschlussdisposition

Für das Trainieren eines Regellerners werden Attribute beziehungsweise Features von jedem Anschlusskonflikt, als Trainingsdaten benötigt. Dabei wird im Kontext dieser Arbeit, zwischen bereits vorhandenen und abzuleitenden Attributen unterschieden. Die beiden Attributtypen unterscheiden sich darin, dass die gegebenen Attribute direkt aus dem Datenmodell entnommen werden können. Währenddessen die abgeleiteten Attribute zuerst durch eine Verarbeitung anderer Attribute bestimmt werden müssen.

5.4.1 Gegebene Attribute

Tabelle 5.6 zeigt die direkt aus dem Datenmodell ablesbaren Attribute, unterteilt nach den Entitäten im Datenmodell, von denen die jeweiligen Attribute stammen.

Betroffene Entität	Parametername	Bedeutung
Anschlussattribute	$t_{\text{Regelwartezeit}}$	Die Regelwartezeit, die der Abbringer dieses Anschlusses zum Warten auf den Zubringer hat.
Betriebsstellenattribute	RIL100	Eine eindeutige Abkürzung der Betriebsstelle.
	Bahnhofstyp	Der Typ dieser Betriebsstelle, zum Beispiel Hauptbahnhof, Bahnhof, Haltepunkt oder Umspannwerk.
	$t_{\text{mindestübergangszeit}}$	Die Zeit, die ein Fahrgast mindestens benötigt, um von einem Bahnsteig der Betriebsstelle zu einem anderen zu gelangen.
	$t_{\text{mindestübergangszeit bg}}$	Die Zeit, die ein Fahrgast mindestens benötigt, um von einem Zug zu einem anderen am selben Bahnsteig zu gelangen.
Zubringer-/Abbringer-Attribute	Linie	Die Linie dieser Fahrt.
	Richtung	Die Richtung, die diese Fahrt fährt in der Form: (Von - Nach).

Betroffene Entität	Parametername	Bedeutung
	B_{Start}	Der Startbahnhof dieser Fahrt. Falls der Startbahnhof nicht direkt bekannt ist, wird der erste Halt dieser Fahrt aus dem Fahrplan als Startbahnhof verwendet.
	B_{Ziel}	Der Zielbahnhof dieser Fahrt. Falls der Zielbahnhof nicht direkt bekannt ist, wird der letzte Halt dieser Fahrt aus dem Fahrplan als Zielbahnhof verwendet.
Zubringerhalt- /Abbringerhalt- Attribute	$t_{\text{soll ankunft}}$	Die im Fahrplan angegebene Ankunftszeit von F an B für diesen Halt.
	$t_{\text{soll abfahrt}}$	Die im Fahrplan angegebene Abfahrtszeit von F an B für diesen Halt. Ist $\text{Fahrtyp}_{\text{soll}} = \text{Durchfahrt}$, gilt automatisch auch $t_{\text{soll ankunft}} = t_{\text{soll abfahrt}}$ an diesem Halt.
	$t_{\text{ist ankunft}}$	Die tatsächlich eingetretene Ankunftszeit von F an B für diesen Halt.
	$t_{\text{ist abfahrt}}$	Die tatsächlich eingetretene Abfahrtszeit von F an B für diesen Halt. Ist $\text{Fahrtyp}_{\text{ist}} = \text{Durchfahrt}$, gilt automatisch auch $t_{\text{ist ankunft}} = t_{\text{ist abfahrt}}$ an diesem Halt.

Tabelle 5.6: Zeigt die direkt exportierten Attribute der jeweiligen Entitäten aus dem Datenmodell.

5.4.2 Weitere abgeleitete Attribute

Ergänzend zu den direkt aus dem Datenmodell ablesbaren Attributen, werden zusätzliche Attribute für den Regellerner, aus den ablesbaren Attributen abgeleitet. Diese Attribute sind in Tabelle 5.7 dargestellt.

Entität	Parametername	Bedeutung
Anschlussattribute	$t_{A, \text{soll umstieg}}$	Die laut Fahrplan festgelegte Umsteigezeit in Minuten für einen Anschluss A , um von dessen Zubringer in dessen Abbringer zu wechseln.
	$t_{A, \text{ist umstieg}}$	Die tatsächliche Umsteigezeit in Minuten eines Anschlusses A , um von dessen Zubringer in dessen Abbringer zu wechseln.
Zubringer- /Abbringer- Attribute	$t_{F, \text{donefahrtdauer}}$	Die bereits gefahrene Fahrtzeit einer Fahrt F von seinem Startbahnhof bis zur Ankunft am Anschlussbahnhof in Minuten.
	$n_{F, \text{donehalte}}$	Die bereits angefahrenen Halte einer Fahrt F von seinem Startbahnhof bis zur Ankunft am Anschlussbahnhof in Minuten. Dazu gehören auch die durchgeführten Durchfahrten, an denen F nicht gehalten hat.
	$t_{F, \text{restfahrtdauer}}$	Die restliche Fahrtzeit einer Fahrt F nach Abfahrt vom Anschlussbahnhof bis zur Ankunft am Zielbahnhof in Minuten.
	$n_{F, \text{resthalte}}$	Die Anzahl der restlichen Halte einer Fahrt F nach Abfahrt vom Anschlussbahnhof bis zur Ankunft am Zielbahnhof. Dazu gehören auch die durchgeführten Durchfahrten, an denen F nicht gehalten hat.
	$t_{A,F\text{Zubringer}, \text{eingangsverspätung}}$	Die Verspätung des Zubringers zur Eingangszeit des Anschlusses in Minuten.

Entität	Parametername	Bedeutung
	$t_{A,F}$ Abbringer, eingangsverspätung	Die Verspätung des Abbringers zur Eingangszeit des Anschlusses in Minuten.
Zubringerhalt- /Abbringerhalt- Attribute	t_H , soll haltezeit	Die laut Fahrplan festgelegte Haltezeit der Fahrt an Halt H in Minuten.
	t_H , ist haltezeit	Die tatsächliche Haltezeit der Fahrt an Halt H in Minuten.
	t_H , wartezeit	Die Wartezeit an Halt H in Minuten.
	t_H , ankunft verspätung	Die Ankunftsverspätung einer Fahrt F an Halt H in Minuten.
	t_H , abfahrt verspätung	Die Abfahrtsverspätung einer Fahrt F an Halt H in Minuten.
	t_H , soll ankunft uhrzeit	Die Uhrzeit, bestehend aus Stunden und Minuten, zu der die Fahrt F an Halt H laut Fahrplan ankommen soll.
	t_H , soll abfahrt uhrzeit	Die Uhrzeit, bestehend aus Stunden und Minuten, zu der die Fahrt F an Halt H laut Fahrplan abfahren soll.
	t_H , ist ankunft uhrzeit	Die Uhrzeit, bestehend aus Stunden und Minuten, zu der die Fahrt F an Halt H ankommt.
	t_H , ist abfahrt uhrzeit	Die Uhrzeit, bestehend aus Stunden und Minuten, zu der die Fahrt F an Halt H abfährt.
	d_H , soll ankunft wochentag	Der Wochentag, kodiert in einer Zahl zwischen 1 für Montag und 7 für Sonntag, an dem die Fahrt F an Halt H laut Fahrplan ankommen soll.
	d_H , soll abfahrt wochentag	Der Wochentag, kodiert in einer Zahl zwischen 1 für Montag und 7 für Sonntag, an dem die Fahrt F an Halt H laut Fahrplan abfahren soll.
	d_H , ist ankunft wochentag	Der Wochentag, kodiert in einer Zahl zwischen 1 für Montag und 7 für Sonntag, an dem die Fahrt F an Halt H ankommt.
	d_H , ist abfahrt wochentag	Der Wochentag, kodiert in einer Zahl zwischen 1 für Montag und 7 für Sonntag, an dem die Fahrt F an Halt H abfährt.

Tabelle 5.7: Zeigt die zu berechnenden Attribute für die Entitäten aus dem Datenmodell.

6 Gewähltes Lernproblem

In Kapitel 4.2 wurde erläutert, dass für den Einsatz von Verfahren aus dem Maschinellen Lernen ein Lernproblem benötigt wird. In diesem Abschnitt wird das in dieser Arbeit betrachtete Lernproblem definiert, das anschließend in Kapitel 7 (Ergebnisse) evaluiert wird.

Lernaufgabe

Die Aufgabe des Regellers ist es, eine Menge von Regeln zu lernen, mit deren Hilfe Maßnahmen für die Sicherung eines gegebenen Anschlusskonflikts gefunden werden können. Die gefundene Maßnahme kann dabei entweder eine einzelne Maßnahme oder eine Kombination mehrerer Maßnahmen sein.

Betrachtet werden dafür die folgenden Maßnahmen zur Sicherung von Anschlusskonflikten: „Alternativanschluss“, „Warten in Regelwartezeit“, „Warten mit Wartezeitüberschreitung“, „Weiterfahren ohne Warten“, „Zubringer-Gleiswechsel“, „Abbringer-Gleiswechsel“ sowie der Einsatz eines „Zusatzzugs“. Diese Maßnahmen werden betrachtet, weil sie, wie in Kapitel 5 beschrieben, aus den zu Verfügung gestellten Daten extrahiert werden können.

Vorhandene Datenmenge

Als Wissensbasis stehen dem Regeller Betriebsdaten der Deutschen Bahn zur Verfügung, die zuvor in Kapitel 5 aufbereitet wurden und aus den folgenden Tabellen bestehen:

- eine Tabelle mit Betriebsstellen
- eine Tabelle mit Fahrten
- eine Tabelle mit Soll-Ankunfts- und Abfahrtszeiten der Fahrten
- eine Tabelle mit Istdaten, passend zu den Soll-Ankunfts- und Abfahrtszeiten
- eine Tabelle mit Übergangszeiten
- eine Tabelle mit Regelwartezeiten, sowie eine Tabelle mit Regelwartezeit-Abweichungen
- eine Tabelle mit Gleiswechseln
- eine Tabelle mit Zusatzzügen und deren Soll-Halten
- zwei Tabellen mit Anschlüssen

Bewertungskriterium

Bewertet werden die gelernten Regeln zum einen durch den Autor und durch einen Experten¹, der die Regeln hinsichtlich ihrer Plausibilität überprüft. Zusätzlich zu einer Bewertung durch Experten wird eine Kreuz-Validierung durchgeführt, um herauszufinden, wie gut das gelernte Modell auf ungesesehenen Testdaten funktioniert. Bei der Kreuz-Validierung wird die Accuracy verwendet, um die Qualität des gelernten Modells im Vergleich zu einem Baseline-Klassifizierer zu bestimmen.

¹ Anselmo Stelzer

7 Ergebnisse

In diesem Kapitel wird eine Evaluation der Einsetzbarkeit eines Regellers für die Anschlussdisposition durchgeführt.

Dafür wird zuerst erläutert, welche Versuchsaufbauten evaluiert werden, und wie diese aufgebaut sind. Anschließend wird darauf eingegangen, anhand welcher Methode die Verwendbarkeit eines Regellers für die Anschlussdisposition bewertet wird. Außerdem wird auf die für die Evaluation verwendete Baseline eingegangen. Zuletzt werden die eigentlichen Ergebnisse der Evaluation geschildert. Dabei wird auf die Gesamterkenntnisse der Evaluation aller Versuchsaufbauten eingegangen.

7.1 Versuchsaufbauten

In dieser Arbeit werden insgesamt 16 Versuchsaufbauten untersucht.

In einem Versuchsaufbau wird eine Klassenzusammenstellung in Kombination mit einer Featuremenge evaluiert. Eine Klassenzusammenstellung beschreibt eine Menge von betrachteten Maßnahmen und Maßnahmenkombinationen. Mit den insgesamt 16 Versuchsaufbauten werden damit verschiedene Kombinationen von Klassenzusammenstellungen und Featuremengen betrachtet.

7.1.1 Klassenzusammenstellungen

Insgesamt werden in dieser Arbeit vier verschiedene Klassenzusammenstellungen betrachtet. Jede Klassenzusammenstellung unterscheidet sich durch die betrachteten Klassen beziehungsweise Maßnahmen. Eine Klasse beschreibt eine Maßnahme oder eine Kombination von Maßnahmen, die zur Lösung eines Anschlusskonfliktes eingesetzt wurden. Die verwendeten Klassenzusammenstellungen sind:

Klassenzusammenstellung	Beschreibung
\mathbb{K}_1	In \mathbb{K}_1 werden alle Maßnahmen und deren Kombinationen betrachtet, mit Ausnahme der allgemeinen Gleiswechsel-Maßnahme. Tabelle 7.2 zeigt für die Klassenzusammenstellung \mathbb{K}_1 , für wie viele Anschlusskonflikte jede Maßnahme angewendet wurde. Insgesamt sind in \mathbb{K}_1 1.512 Anschlusskonflikte enthalten, für die insgesamt 19 unterschiedliche Maßnahmen und Maßnahmenkombinationen durchgeführt wurden. Dabei fällt auf, dass für keinen Anschlusskonflikt eine Zusatzzug-Maßnahme getroffen wurde. Der Grund dafür ist, dass beim Einlesen des Datenpakets der Deutschen Bahn kein Anschluss gefunden werden konnte, dessen Zubringer oder Abbringer ein Zusatzzug ist.

Klassenzusammenstellung	Beschreibung
\mathbb{K}_2	<p>Bei der Betrachtung der Maßnahmen in Tabelle 7.2 fällt auf, dass 4 Maßnahmenkombinationen, nur für jeweils einen Anschlusskonflikt angewendet wurden. Diese vier Maßnahmenkombinationen lassen sich nicht für eine Evaluation nutzen, weil die Anzahl der vorhandenen Anschlusskonflikte nicht ausreichen, um sie in eine Test- und Trainingsmenge aufzuteilen. Für 4 weitere Maßnahmenkombinationen gilt, dass sie bei weniger als 10 Anschlusskonflikten eingesetzt werden. Auch diese 4 Maßnahmenkombinationen eignen sich nicht für eine Evaluation, da die Anzahl der vorhandenen Anschlusskonflikte nicht ausreichend sind, um sie auf die Test- und Trainingsmengen der Kreuz-Validierung aufzuteilen. In den Daten lassen sich mehrere sehr ähnliche Maßnahmenkombinationen finden, die sich jeweils nur durch die Verwendung einer anderen Gleiswechsel-Maßnahme voneinander unterscheiden. Aus diesem Grund wird eine Klassenzusammenstellung \mathbb{K}_2 betrachtet, in der sowohl die Abbringer- und Zubringergleiswechsel als eine Maßnahme, der „Gleiswechsel“-Maßnahme, zusammengefasst werden.</p> <p>Insgesamt befinden sich in \mathbb{K}_2 die Maßnahmen: „Weiterfahren ohne Warten“, „Warten in Regelwartezeit“, „Warten mit Wartezeitüberschreitung“, „Gleiswechsel“, „Alternativanschluss“ sowie deren Kombinationen, die auf insgesamt 1.512 Anschlusskonflikten angewendet wurden.</p>
\mathbb{K}_3	<p>Weil allerdings auch nach dem Zusammenfassen der Gleiswechsel, immer noch zwei Maßnahmenkombinationen bei weniger als 10 Anschlusskonflikten angewendet werden, werden in einer dritten Klassenzusammenstellung \mathbb{K}_3, nur die sechs Einzelmaßnahmen betrachtet. Zu diesen sechs Einzelmaßnahmen gehören, wie bereits in \mathbb{K}_2 keine Abbringer- und Zubringergleiswechsel, dafür werden die zusammengefassten Gleiswechsel, berücksichtigt. Damit kann bewertet werden, wie gut sich automatisch gelernte Regeln dafür eignen, einzelne Maßnahmen zu erkennen, ohne dass Maßnahmenkombinationen die gelernten Regeln verfälschen.</p> <p>Insgesamt werden in \mathbb{K}_3 die Maßnahmen: „Weiterfahren ohne Warten“, „Warten in Regelwartezeit“, „Warten mit Wartezeitüberschreitung“, „Gleiswechsel“ und „Alternativanschluss“ betrachtet, die bei 1.410 Anschlusskonflikte durchgeführt wurden.</p>
\mathbb{K}_4	<p>In einem letzten Versuchsaufbau wird abschließend untersucht, wie gut sich automatisch gelernte Regeln dazu eignen, zwischen den am häufigsten getroffenen Wartemaßnahmen zu unterscheiden. Diese Maßnahmen sind: „Warten mit Wartezeitüberschreitung“, „Warten in Regelwartezeit“ und „Weiterfahren ohne zu Warten“. Auch für diese Klassenzusammenstellung werden keine Maßnahmenkombinationen miteinbezogen, sondern lediglich Anschlusskonflikte betrachtet, für die eine der oben genannten Maßnahmen getroffen wurden. Damit soll verhindert werden, dass das Ergebnis durch die Einbeziehung von Maßnahmenkombinationen beeinflusst wird. Insgesamt sind in \mathbb{K}_4 925 Anschlusskonflikte enthalten.</p>

Tabelle 7.4: Definition der evaluierten Klassenzusammenstellungen.

Die Tabellen 7.2 und 7.3 zeigen die in den vier Klassenzusammenstellungen enthaltenen Maßnahmen sowie die Anzahl der dazugehörigen Anschlusskonflikte. Dabei enthält Tabelle 7.2 die Maßnahmen für \mathbb{K}_1 und Tabelle 7.3 die Maßnahmen für \mathbb{K}_2 , \mathbb{K}_3 und \mathbb{K}_4 .

Die Werte aus den beiden Tabellen basieren auf den in Kapitel 5.2 beschriebenen Daten, die wie dort beschrieben aufbereitet wurden.

Getroffene Maßnahmen	Anzahl Anschlusskonflikte in \mathbb{K}_1
Alle Anschlüsse	1.512
Warten in Regelwartezeit, Zubringergleiswechsel, Abbringergleiswechsel	1
Warten mit Wartezeitüberschreitung, Zubringergleiswechsel, Abbringergleiswechsel	1
Zubringergleiswechsel, Abbringergleiswechsel, Weiterfahren ohne Warten	1
Abbringergleiswechsel, Alternativanschluss	3
Abbringergleiswechsel, Weiterfahren ohne Warten	5
Alternativanschluss, Weiterfahren ohne Warten	12
Warten in Regelwartezeit, Zubringergleiswechsel	9
Warten mit Wartezeitüberschreitung, Abbringergleiswechsel	16
Warten mit Wartezeitüberschreitung, Alternativanschluss	2
Warten mit Wartezeitüberschreitung, Zubringergleiswechsel	28
Zubringergleiswechsel, Abbringergleiswechsel	33
Zubringergleiswechsel, Alternativanschluss	1
Zubringergleiswechsel, Weiterfahren ohne Warten	23
Abbringergleiswechsel	112
Alternativanschluss	94
Warten in Regelwartezeit	139
Warten mit Wartezeitüberschreitung	462
Weiterfahren ohne Warten	324
Zubringergleiswechsel	246
Zusatzzug	0

Tabelle 7.2: Verteilung der vorkommenden Maßnahmen auf die Anschlusskonflikte in \mathbb{K}_1 .

Getroffene Maßnahmen	Anzahl Anschlusskonflikte in \mathbb{K}_2	Anzahl Anschlusskonflikte in \mathbb{K}_3	Anzahl Anschlusskonflikte in \mathbb{K}_4
Alle Anschlüsse	1.512	1.410	925
Alternativanschluss, Weiterfahren ohne Warten	12	0	0
Gleiswechsel, Alternativanschluss	4	0	0
Gleiswechsel, Weiterfahren ohne Warten	29	0	0
Warten in Regelwartezeit, Gleiswechsel	10	0	0
Warten mit Wartezeitüberschreitung, Alternativanschluss	2	0	0
Warten mit Wartezeitüberschreitung, Gleiswechsel	45	0	0
Alternativanschluss	94	94	0
Gleiswechsel	391	391	0
Warten in Regelwartezeit	139	139	139
Warten mit Wartezeitüberschreitung	462	462	462
Weiterfahren ohne Warten	324	324	324
Zusatzzug	0	0	0

Tabelle 7.3: Verteilung der vorkommenden Maßnahmen auf die Anschlusskonflikte in \mathbb{K}_2 , \mathbb{K}_3 und \mathbb{K}_4 .

7.1.2 Featuremengen

Neben der Unterteilung in Klassenzusammenstellungen werden für jede Klassenzusammenstellung vier Experimente mit unterschiedlichen Features durchgeführt. Für jedes dieser vier Experimente wird eine andere Featuremenge eingesetzt.

Featuremenge	Beschreibung
\mathbb{F}_1	Diese Featuremenge enthält alle in Kapitel 5.4 exportierten Features. Dazu gehören sowohl die im Datenmodell enthaltenen Features als auch die in 5.4.2 abgeleiteten Features. Diese Featuremenge dient dazu, eine erste Einschätzung der Funktion eines Regellerners auf den aufbereiteten Daten zu erhalten. Dies ist insbesondere wichtig, weil für zwei gleiche Anschlüsse in der Trainingsmenge verschiedene Maßnahmen getroffen worden sein können, die die gelernten Regeln verfälschen.
\mathbb{F}_2	Das Datenmodell enthält allerdings auch Informationen, die dem Disponenten, zu der Zeit, in der die Maßnahme getroffen wurde, noch nicht zur Verfügung standen. Ein Beispiel dafür ist die tatsächliche Abfahrtszeit des Abbringers oder die tatsächliche Umsteigezeit, die sich beide erst durch die getroffene Maßnahme des Disponenten ergeben. Aus diesem Grund sind in dieser Featuremenge nur die Attribute enthalten, die dem Disponenten tatsächlich zur Dispositionszeit zur Verfügung gestanden haben. Die aus \mathbb{F}_1 entfernten Attribute sind: $t_{A, \text{ist umstieg}}$, $t_{H_{B,F} \text{Abbringer}, \text{abfahrt verspaetung}}$, $t_{H_{B,F} \text{Abbringer}, \text{ist abfahrt}}$, $t_{H_{B,F} \text{Abbringer}, \text{ist abfahrt uhrzeit}}$, $d_{H_{B,F} \text{Abbringer}, \text{ist abfahrt wochentag}}$, $t_{H_{B,F} \text{Abbringer}, \text{ist haltezeit}}$, $t_{A, F \text{Abbringer}, \text{eingangsverspätung}}$ und $t_{H_{B,F} \text{Abbringer}, \text{wartezeit}}$.
\mathbb{F}_3	Weiterhin enthalten \mathbb{F}_1 und \mathbb{F}_2 für die Darstellung von Zeiten sowohl Timestamps, als auch Uhrzeit und Wochentags Informationen als Attribute. Diese Zeit-Attribute eignen sich dafür Regeln mit Zeitintervallen zu bilden, die nur in einem definierten Zeitfenster gelten. Damit könnten zum Beispiel Regeln gebildet werden, die nur zu Stoßzeiten gelten. Um zu überprüfen, wie die Verwendung von Uhrzeit- und Wochentags-Attributen, im Gegensatz zu Timestamp-Attributen für Zeitintervalle sich auf die erzeugten Regeln auswirken, wurden in dieser Featuremenge alle Timestamp-Attribute aus \mathbb{F}_2 entfernt. Die aus \mathbb{F}_2 entfernten Attribute sind: $t_{H_{B,F} \text{Zubringer}, \text{soll ankunft}}$, $t_{H_{B,F} \text{Zubringer}, \text{ist ankunft}}$ und $t_{H_{B,F} \text{Abbringer}, \text{soll abfahrt}}$.
\mathbb{F}_4	In der letzten Featuremenge \mathbb{F}_4 werden die Uhrzeit- und Wochentags-Attribute aus \mathbb{F}_2 entfernt. Damit lässt sich testen, ob in den Uhrzeit- und Wochentags-Attributen oder in den Timestamp-Attributen wichtigere Informationen, für die Erkennung der Maßnahmen für einen Anschlusskonflikt enthalten sind. Die aus \mathbb{F}_2 entfernten Attribute sind: $t_{H_{B,F} \text{Zubringer}, \text{soll ankunft uhrzeit}}$, $d_{H_{B,F} \text{Zubringer}, \text{soll ankunft wochentag}}$, $t_{H_{B,F} \text{Zubringer}, \text{ist ankunft uhrzeit}}$, $d_{H_{B,F} \text{Zubringer}, \text{ist ankunft wochentag}}$, $t_{H_{B,F} \text{Abbringer}, \text{soll abfahrt uhrzeit}}$ und $d_{H_{B,F} \text{Abbringer}, \text{soll abfahrt wochentag}}$.

Tabelle 7.5: Definition der evaluierten Featuremengen.

Die in den Featuremengen verwendeten Features werden zur Verdeutlichung in der nachfolgenden Tabelle 7.6 abgebildet.

	Featurename	Beispiel	F ₁	F ₂	F ₃	F ₄
Anschlussfeatures	t_A , soll umstieg	16	✓	✓	✓	✓
	t_A , ist umstieg	4	✓			
	t_A , Regelwartezeit	0	✓	✓	✓	✓
Bahnhofsfeatures	RIL100	KK	✓	✓	✓	✓
	Bahnhofstyp	Bf	✓	✓	✓	✓
	$t_{\text{mindestübergangszeit}}$	7	✓	✓	✓	✓
	$t_{\text{bahnsteiggleiche übergangszeit}}$	3	✓	✓	✓	✓
Zubringerfeatures	$B_{\text{Zubringer Start}}$	TS	✓	✓	✓	✓
	$B_{\text{Zubringer Ziel}}$	AA	✓	✓	✓	✓
	$\text{Linie}_{\text{Zubringer}}$	30.0	✓	✓	✓	✓
	$\text{Richtung}_{\text{Zubringer}}$	Stuttgart_Hbf - _Hamburg-Altona	✓	✓	✓	✓
	$t_{H_B,F_{\text{Zubringer}}}$, ankunft verspaetung	13	✓	✓	✓	✓
	$t_{H_B,F_{\text{Zubringer}}}$, soll ankunft	'2015-06-15 15:05'	✓	✓		✓
	$t_{H_B,F_{\text{Zubringer}}}$, soll ankunft uhrzeit	'15:05'	✓	✓	✓	
	$d_{H_B,F_{\text{Zubringer}}}$, soll ankunft wochentag	1	✓	✓	✓	
	$t_{H_B,F_{\text{Zubringer}}}$, ist ankunft	'2015-06-15 15:18'	✓	✓		✓
	$t_{H_B,F_{\text{Zubringer}}}$, ist ankunft uhrzeit	'15:18'	✓	✓	✓	
	$d_{H_B,F_{\text{Zubringer}}}$, ist ankunft wochentag	1	✓	✓	✓	
	$t_{H_B,F_{\text{Zubringer}}}$, soll haltezeit	5	✓	✓	✓	✓
	$t_{H_B,F_{\text{Zubringer}}}$, ist haltezeit	4	✓	✓	✓	✓
	$t_{F_{\text{Zubringer}}}$, donehalte	111	✓	✓	✓	✓
	$t_{F_{\text{Zubringer}}}$, donefahrtdauer	208	✓	✓	✓	✓
	$t_{F_{\text{Zubringer}}}$, resthalte	98	✓	✓	✓	✓
	$t_{F_{\text{Zubringer}}}$, restfahrtdauer	258	✓	✓	✓	✓
$t_{A,F_{\text{Zubringer}}}$, eingangverspätung	10	✓	✓	✓	✓	
$t_{H_B,F_{\text{Zubringer}}}$, wartezeit	2	✓	✓	✓	✓	
Abbringerfeatures	$B_{\text{Abbringer Start}}$	KKR	✓	✓	✓	✓
	$B_{\text{Abbringer Ziel}}$	EMST	✓	✓	✓	✓
	$\text{Linie}_{\text{Abbringer}}$	D.7	✓	✓	✓	✓
	$\text{Richtung}_{\text{Abbringer}}$	Krefeld_Hb - _Münster_(Westf)_Hbf	✓	✓	✓	✓
	$t_{H_B,F_{\text{Abbringer}}}$, abfahrt verspaetung	1	✓			
	$t_{H_B,F_{\text{Abbringer}}}$, soll abfahrt	'2015-06-15 15:21'	✓	✓		✓
	$t_{H_B,F_{\text{Abbringer}}}$, soll abfahrt uhrzeit	'15:21'	✓	✓	✓	
	$d_{H_B,F_{\text{Abbringer}}}$, soll abfahrt wochentag	1	✓	✓	✓	
	$t_{H_B,F_{\text{Abbringer}}}$, ist abfahrt	'2015-06-15 15:22'	✓			
	$t_{H_B,F_{\text{Abbringer}}}$, ist abfahrt uhrzeit	'15:22'	✓			
$d_{H_B,F_{\text{Abbringer}}}$, ist abfahrt wochentag	1	✓				

	Featurename	Beispiel	\mathbb{F}_1	\mathbb{F}_2	\mathbb{F}_3	\mathbb{F}_4
	$t_{H_{B,F}Abbringer}$, soll haltezeit	3	✓	✓	✓	✓
	$t_{H_{B,F}Abbringer}$, ist haltezeit	2	✓			
	$t_{F_{Abbringer}}$, donehalte	18	✓	✓	✓	✓
	$t_{F_{Abbringer}}$, donefahrtdauer	43	✓	✓	✓	✓
	$t_{F_{Abbringer}}$, resthalte	35	✓	✓	✓	✓
	$t_{F_{Abbringer}}$, restfahrtdauer	121	✓	✓	✓	✓
	$t_{A,F_{Abbringer}}$, eingangsverspätung	0	✓			
	$t_{H_{B,F}Abbringer}$, wartezeit	2	✓			

Tabelle 7.6: Vergleich der in den Featuremengen $\mathbb{F}_1, \mathbb{F}_2, \mathbb{F}_3$ und \mathbb{F}_4 enthaltenen Attribute, gruppiert nach ihrer Objektzugehörigkeit im Datenmodell.

7.1.3 Abkürzungen der Maßnahmen

Aus Gründen der besseren Lesbarkeit der erzeugten Regeln durch den Regellerner, werden im folgenden, bei der Erwähnung von Regeln, die nachfolgenden Abkürzungen für die Maßnahmen verwendet:

Maßnahme	Abkürzung
Weiterfahren ohne Warten	W_w
Warten in Regelwartezeit	W_r
Warten mit Wartezeitüberschreitung	$W_{\ddot{u}}$
Zusatzzug	Z
Gleiswechsel	G
Zubringergleiswechsel	G_z
Abbringergleiswechsel	G_a
Alternativanschluss	A

Tabelle 7.7: Abkürzungen der Maßnahmen in den Regelmengen.

7.2 Evaluations-Baseline

Für die Evaluierung der Versuchsaufbauten wird eine Baseline benötigt, an der gemessen werden kann, ab wann die gelernten Regeln gut, beziehungsweise schlecht sind. Als Baseline wird in dieser Arbeit ein Klassifikator eingesetzt, der für jede Testinstanz die am häufigsten in der Trainingsmenge vorkommende Klasse vorhersagt.

7.3 Evaluationsmethode

Die Evaluation wird mit dem in Kapitel 4.2.3 beschriebenen Kreuz-Validierungsverfahren mit den Parametern $k = 1$ und $n = 10$ durchgeführt. Für die Durchführung einer 10-fachen Kreuz-Validierung wird die Trainingsmenge \mathbb{T} in 10 relativ gleich große Untermengen $\{T_i | i \in [1, 10]\}$ aufgeteilt. Anschließend wird mit jeder Untertestmenge $T_i \subset \mathbb{T}$, eine Klassifikation anhand der übrigen Untermengen $\{T_j | j \neq i\}$ durchgeführt.

Maßnahmengruppe	\mathbb{F}_1	\mathbb{F}_2	\mathbb{F}_3	\mathbb{F}_4
\mathbb{K}_1	32	32	32	28
\mathbb{K}_2	24	22	19	27
\mathbb{K}_3	22	16	17	17
\mathbb{K}_4	7	14	13	6

Tabelle 7.8: Vergleich der Anzahl erzeugter Regeln für die verschiedenen Klassenzusammenstellungen in Abhängigkeit der einzelnen Featuregruppen.

Maßnahmengruppe	\mathbb{F}_1	\mathbb{F}_2	\mathbb{F}_3	\mathbb{F}_4
\mathbb{K}_1	11	8	10	8
\mathbb{K}_2	5	5	5	10
\mathbb{K}_3	6	4	2	4
\mathbb{K}_4	1	1	2	0

Tabelle 7.9: Zeigt die Anzahl gelernter Regeln die mindestens eine Bedingung mit einer Linie, einer Fahrtrichtung oder einem Bahnhof enthalten.

7.4 Evaluation

Hier werden die Ergebnisse der Evaluation der 16 zuvor aufgestellten Versuchsaufbauten präsentiert. Dabei zeigt die Tabelle 7.10 die erreichten Accuracy-Ergebnisse für die einzelnen Versuchsaufbauten und für die jeweils dazugehörige Baseline. In Abbildung 7.1 werden die Accuracy-Ergebnisse aus Tabelle 7.10 in Form eines Koordinatensystems dargestellt. Die x -Achse enthält die evaluierten Featuremengen, wohingegen auf der y -Achse die Accuracy im Intervall $[0, 1]$ dargestellt wird. Die evaluierte Klassenzusammenstellung wird durch die Farbe sowie das Symbol der Koordinaten dargestellt. Dabei hat jede der vier Klassenzusammenstellungen die in der Legende wiedergegebenen Symbole und Farben.

Die Tabelle 7.8 enthält die Anzahl der für die einzelnen Versuchsaufbauten gelernten Regeln, während die Tabelle 7.9 die Anzahl der Regeln enthält, die mindestens eine Bedingung mit einer Fahrtrichtung oder einer Betriebsstelle enthalten. Zur besseren Lesbarkeit dieser Arbeit befinden sich die für alle Versuchsaufbauten gelernten Regelmengen in Anhang A.

Die Ergebnisse in Abbildung 7.1 zeigen, dass alle Versuchsaufbauten zu besseren Ergebnissen führen, als die dazugehörige Baseline. Für alle Versuchsaufbauten ist die Differenz der erreichten Accuracy und der dazugehörigen Baseline größer als 12%. Dies bedeutet, dass für mindestens 12% Prozent der Anschlusskonflikte jeder Klassenzusammenstellung, die dazugehörigen Maßnahmen anhand der gelernten Regelmengen, korrekt identifiziert werden konnten. Das beste Accuracy-Ergebnis wurde für \mathbb{K}_4 unter Verwendung von \mathbb{F}_1 mit 98% erzielt. Für \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 wurden die besten Accuracy-Ergebnisse bei Betrachtung der Anschlusskonflikte aus \mathbb{K}_3 erreicht, diese sind je nach Featuremenge zwischen 1% und 4%

Maßnahmengruppe	\mathbb{F}_1	\mathbb{F}_2	\mathbb{F}_3	\mathbb{F}_4	Baseline
\mathbb{K}_1	76,98	56,42	55,69	57,61	30,56
\mathbb{K}_2	83,2	61,04	61,71	61,38	30,56
\mathbb{K}_3	89,86	67,94	66,6	66,88	32,77
\mathbb{K}_4	98,16	65,62	62,92	65,73	49,95

Tabelle 7.10: Vergleich der Accuracy-Ergebnisse der verschiedenen Versuchsaufbauten im Vergleich zur Baseline.

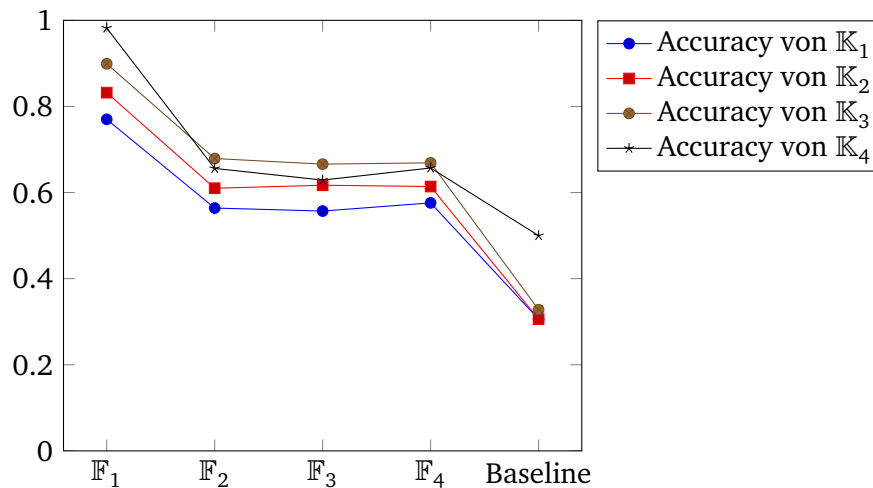


Abbildung 7.1: Zeigt die Accuracy der Regelmengen aller Klassenzusammenstellungen unter Verwendung aller Featuregruppen.

höher als die zweitbesten erreichten Accuracy-Ergebnisse für \mathbb{K}_4 . Das bessere Ergebnis gegenüber \mathbb{K}_1 und \mathbb{K}_2 ist darin begründet, dass \mathbb{K}_3 keine Anschlusskonflikte mit Maßnahmenkombinationen enthält, und genauso wie in \mathbb{K}_2 die Abbringer- und Zubringergleiswechsel in einer Maßnahme kombiniert sind.

Bei Betrachtung der Featuremengen wird über alle Klassenzusammenstellungen für \mathbb{F}_1 die beste Accuracy erreicht. Die Ergebnisse für \mathbb{F}_1 zeigen, dass mit den richtigen Attributen, Accuracy-Ergebnisse zwischen 76% und 98% erreichbar sind. Allerdings zeigen die vergleichsweise schlechteren Ergebnisse für die übrigen Featuremengen, dass die Attribute aus \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 nicht dafür ausreichen, die fehlenden Attribute aus \mathbb{F}_1 zu kompensieren, beziehungsweise diese Attribute näherungsweise zu ersetzen. Bei der Betrachtung der zweitbesten Accuracy-Ergebnisse kann keine Featuremenge erkannt werden, die zu besonders guten Accuracy-Ergebnissen führt, da abhängig von den betrachteten Klassenzusammenstellungen jeweils eine andere Featuremenge zur höchsten Accuracy führt.

Eine Erkenntnis aus Tabelle 7.8 ist, dass die Anzahl der gelernten Regeln, generell mit Abnahme der möglichen Maßnahmen, sinkt. Dies liegt daran, dass für jede Maßnahme eine eigene Regelmenge gelernt wird, womit die Gesamtzahl der gelernten Regeln ebenfalls abnimmt, wenn die Anzahl möglicher Maßnahmen abnimmt. Die meisten Regeln werden mit 32 Regeln für die Versuchsaufbauten für \mathbb{K}_1 in Kombination mit \mathbb{F}_1 , \mathbb{F}_2 und \mathbb{F}_3 gelernt. Die wenigsten Regeln, mit insgesamt 6 Regeln, für den Versuchsaufbau bestehend aus \mathbb{K}_4 und \mathbb{F}_4 gelernt. Durchschnittlich werden für einen Versuchsaufbau zwischen 20 und 21 Regeln gelernt. Außerdem fällt bei Betrachtung der Tabelle 7.8 auf, dass die Anzahl gelernter Regeln keine Beziehung zu der verwendeten Featuremenge besitzt, da die maximale sowie die minimale Anzahl gelernter Regeln, für die verschiedenen Klassenzusammenstellungen durch Verwendung unterschiedlicher Featuremengen gelernt wurden.

Die Werte in Tabelle 7.9 zeigen, dass einige der erzeugten Regelmengen einen deutlich erkennbaren Anteil Regeln mit sehr anschlusspezifischen Informationen, wie einer Betriebsstelle, einer Linienangabe oder einer Fahrtrichtung, enthalten. Dabei fällt auf, dass für die Klassenzusammenstellungen \mathbb{K}_1 und \mathbb{K}_2 solche Regeln vermehrt gelernt wurden. Sowohl \mathbb{K}_1 als auch \mathbb{K}_2 haben gemeinsam, dass sie Maßnahmenkombinationen enthalten, für die im Vergleich zu den einzelnen Maßnahmen sehr wenige Trainingsbeispiele vorliegen. Für \mathbb{K}_3 wurden weniger solche Regeln gelernt. Die für \mathbb{K}_3 gelernten anschlusspezifischen Regeln beziehen sich großteils auf die Identifikation von Alternativanschluss-Maßnahmen. Für \mathbb{K}_4 wurden währenddessen so gut wie keine Regeln gelernt, die sehr anschlusspezifische Informationen enthalten. Dies weist darauf hin, dass die Regeln für die Erkennung von Maßnahmenkombinationen sowie einige Regeln für die Erkennung von Alternativanschluss-Maßnahmen zu sehr an die vorhandenen

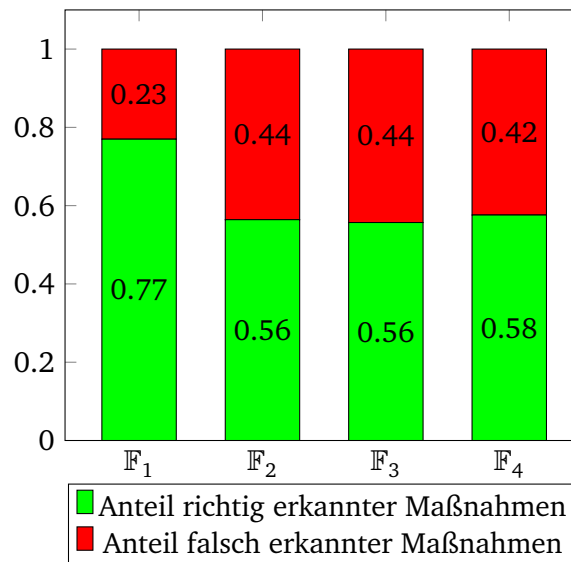


Abbildung 7.2: Vergleich der richtig und falsch erkannten Maßnahmen in \mathbb{K}_1 unter Verwendung aller Featuregruppen.

Daten angepasst sind. Diese Beobachtung wird auch dadurch bestätigt, dass lediglich Trainingsdaten für drei Fahrplantage der Deutschen Bahn vorliegen.

Eine weitere Beobachtung der Regelmengen ist, dass einige der Regeln eine zeitlich offene Bedingung der Form $t_{\text{feature}} \geq \text{timestamp}$ oder $t_{\text{feature}} \leq \text{timestamp}$ enthalten. Solche Bedingungen können allerdings in der Praxis zu Problemen führen, da sie ein offenes Zeitintervall abdecken. In der Praxis macht dies vor allem Probleme, da die Betriebsdaten zeitlich nach den Trainingsdaten, auf denen die Regeln gelernt wurden, erfasst werden, womit eine Bedingung mit einem Zeitintervall entweder alle getesteten Anschlussinstanzen akzeptiert oder keine.

Bei Betrachtung der detaillierten Recall-Ergebnisse der einzelnen Versuchsaufbauten mit F_2 , F_3 und F_4 fällt auf, dass bei allen Versuchsaufbauten ca. 51% bis 60% der „Warten in Regelwartezeit“-Maßnahmen und ca. 36% bis 58% der „Weiterfahren ohne Warten“-Maßnahmen falsch erkannt wurden. Diese Maßnahmen werden meistens fälschlicherweise der „Warten mit Wartezeitüberschreitung“-Maßnahme zugeordnet. Dies deutet darauf hin, dass bei diesen Featuremengen Informationen, die für die Unterscheidung zwischen den drei Warte-Maßnahmen notwendig sind, fehlen.

Über alle Versuchsaufbauten, für die Klassenzusammenstellungen \mathbb{K}_1 und \mathbb{K}_2 , hinweg wurde beobachtet, dass nur sehr wenige Anschlusskonflikte mit Maßnahmenkombinationen korrekt erkannt werden. Häufig werden die Anschlusskonflikte mit Maßnahmenkombinationen nur einer der enthaltenen Maßnahmen zugeordnet. Enthält die Maßnahmenkombination eine Warte-Maßnahme, werden die dazugehörigen Anschlusskonflikte unter Verwendung von F_2 , F_3 und F_4 häufig fälschlicherweise einer anderen Warte-Maßnahme zugeordnet.

7.4.1 Klassenzusammenstellung \mathbb{K}_1

Die Tabelle 7.11 zeigt die Precision-Ergebnisse für die einzelnen Maßnahmen aus K_1 in Kombination mit allen Featuremengen, die durch die Verwendung eines Regellerners berechnet wurden. Ebenso zeigt Tabelle 7.12 die Recall-Ergebnisse, für die einzelnen Maßnahmen in K_1 in Kombination mit allen Featuremengen. Der Anteil richtig und falsch erkannter Anschlusskonflikte wird in Abbildung 7.2 dargestellt.

Maßnahmen	F ₁	F ₂	F ₃	F ₄
Warten in Regelwartezeit, Zubringergleiswechsel	0	0	0	0
Warten mit Wartezeitüberschreitung, Abbringergleiswechsel	0	0	0	0
Zubringergleiswechsel, Weiterfahren ohne Warten	25	0	0	0
Alternativanschluss, Weiterfahren ohne Warten	25	0	0	0
Abbringergleiswechsel, Weiterfahren ohne Warten	0	0	0	0
Warten mit Wartezeitüberschreitung, Zubringergleiswechsel, Abbringergleiswechsel	0	0	0	0
Zubringergleiswechsel, Abbringergleiswechsel, Weiterfahren ohne Warten	0	0	0	0
Weiterfahren ohne Warten	82,87	57,47	57,81	55,95
Zubringergleiswechsel, Alternativanschluss	0	0	0	0
Zubringergleiswechsel, Abbringergleiswechsel	40	62,5	25	0
Warten in Regelwartezeit, Zubringergleiswechsel, Abbringergleiswechsel	0	0	0	0
Warten in Regelwartezeit	86,39	51,15	49,22	54,14
Warten mit Wartezeitüberschreitung, Alternativanschluss	0	0	0	0
Warten mit Wartezeitüberschreitung	82,04	51,96	50,79	54,63
Abbringergleiswechsel	58,82	56,96	55,88	61,97
Zubringergleiswechsel	64,85	65,38	65,82	65,17
Warten mit Wartezeitüberschreitung, Zubringergleiswechsel	14,29	25	20	0
Abbringergleiswechsel, Alternativanschluss	0	0	0	0
Alternativanschluss	76	71,05	79,71	69,14

Tabelle 7.11: Precision-Ergebnisse in Prozent für \mathbb{K}_1

Maßnahmen	F ₁	F ₂	F ₃	F ₄
Warten in Regelwartezeit, Zubringergleiswechsel	0	0	0	0
Warten mit Wartezeitüberschreitung, Abbringergleiswechsel	0	0	0	0
Zubringergleiswechsel, Weiterfahren ohne Warten	4,35	0	0	0
Alternativanschluss, Weiterfahren ohne Warten	8,33	0	0	0
Abbringergleiswechsel, Weiterfahren ohne Warten	0	0	0	0
Warten mit Wartezeitüberschreitung, Zubringergleiswechsel, Abbringergleiswechsel	0	0	0	0
Zubringergleiswechsel, Abbringergleiswechsel, Weiterfahren ohne Warten	0	0	0	0
Weiterfahren ohne Warten	92,59	46,3	45,68	53,7
Zubringergleiswechsel, Alternativanschluss	0	0	0	0
Zubringergleiswechsel, Abbringergleiswechsel	12,12	15,15	9,09	0
Warten in Regelwartezeit, Zubringergleiswechsel, Abbringergleiswechsel	0	0	0	0
Warten in Regelwartezeit	91,37	48,2	45,32	51,8
Warten mit Wartezeitüberschreitung, Alternativanschluss	0	0	0	0
Warten mit Wartezeitüberschreitung	95,89	74,46	76,41	72,73
Abbringergleiswechsel	35,71	40,18	33,93	39,29
Zubringergleiswechsel	77,24	76,02	73,58	76,83
Warten mit Wartezeitüberschreitung, Zubringergleiswechsel	3,57	3,57	3,57	0
Abbringergleiswechsel, Alternativanschluss	0	0	0	0
Alternativanschluss	60,64	57,45	58,51	59,57

Tabelle 7.12: Recall-Ergebnisse in Prozent für \mathbb{K}_1

Für \mathbb{K}_1 werden, je nach verwendeter Featuremenge, zwischen 28 und 32 Regeln gelernt. Die Accuracy für die gelernten Regeln liegt abhängig von den verwendeten Features zwischen 55% und 76%, während die Baseline-Accuracy für \mathbb{K}_1 bei 30,56% liegt. Dies zeigt, dass durch die Verwendung von gelernten Regeln eine Verbesserung der Accuracy von mindestens 25% im Vergleich zur Baseline erreicht werden kann. Die beste Accuracy wird für \mathbb{K}_1 unter Verwendung von \mathbb{F}_1 erzielt, was darin begründet ist, dass \mathbb{F}_1 Attribute enthält, die zur Zeit der Disposition noch nicht bekannt sind, wie zum Beispiel die tatsächliche Umsteigezeit.

Eine weitere Beobachtung der Regelmengen für \mathbb{K}_1 ist, dass die Anzahl der gelernten Regeln unter Betrachtung von \mathbb{F}_4 mit 28 gelernten Regeln am geringsten ist. Zudem werden für \mathbb{K}_1 unter Betrachtung von \mathbb{F}_1 , \mathbb{F}_2 und \mathbb{F}_3 jeweils 32 Regel gelernt. Außerdem erreicht unter den Featuremengen \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 die Featuremenge \mathbb{F}_4 die höchste Accuracy nach \mathbb{F}_1 . Dies deutet darauf hin, dass für \mathbb{K}_1 die Verwendung von Timestamp-Attributen zu weniger gelernten Regeln und einer höheren Accuracy führt, als die Verwendung von Wochentag- und Uhrzeit-Attributen mit oder ohne Timestamp-Attribute.

Bei der Betrachtung der Regeln für die Erkennung der Maßnahmenkombinationen fällt auf, dass für \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 fast alle gelernten Regeln eine Betriebsstelle, eine Linie oder eine Fahrtrichtung enthalten. Dies deutet auf ein Overfitting der Regeln für die Erkennung der Maßnahmenkombinationen hin.

Bei der Betrachtung der Precision-Ergebnisse für die Alternativanschluss-Maßnahmen in Tabelle 7.11 fällt auf, dass die Regeln für die Erkennung der Alternativanschluss-Maßnahmen mit einer Precision von mindestens 69% vergleichsweise selten Anschlusskonflikte mit anderen Maßnahmen abdecken. Allerdings zeigen die Recall-Ergebnisse für die Alternativanschluss-Maßnahme, dass 40% - 42% der Anschlusskonflikte mit Alternativanschluss-Maßnahmen nicht als solche erkannt werden. Dies lässt darauf schließen, dass die Regeln für die Erkennung der Alternativanschluss-Maßnahmen stark an ausgewählte Anschlusskonflikte mit einer Alternativanschluss-Maßnahme angepasst sind und deshalb in Summe nur sehr wenige Alternativanschluss-Maßnahmen erkennen und dafür sehr wenige andere Maßnahmen abdecken.

Die Regeln für die Erkennung der „Warten in Regelwartezeit“-Maßnahmen haben alle bis auf eine Regel gemeinsam, dass sie eine Bedingung enthalten, welche die Regelwartezeit des Anschlusses überprüft. Dies ist interessant, weil die Regelwartezeit für die Entscheidung welche Wartemaßnahme für einen Anschlusskonflikt getroffen wird von zentraler Bedeutung ist, wie in Abschnitt 4.1.2 und 5.2.8 diskutiert wurde. Zwei der drei Regeln, die unter Verwendung der Features aus \mathbb{F}_1 für die Identifikation der „Warten in Regelwartezeit“-Maßnahme gelernt wurden und die meisten „Warten in Regelwartezeit“-Maßnahmen erkennen enthalten die beiden Bedingungen:

$$t_{A, \text{Regelwartezeit}} \geq x \wedge t_{A, \text{ist umstieg}} \geq y \wedge t_{H_B, F_{\text{Abbringer}}, \text{abfahrt verspaetung}} \leq z$$

In den Featuremengen \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 sind die Features $t_{A, \text{ist umstieg}}$ und $t_{H_B, F_{\text{Abbringer}}, \text{abfahrt verspaetung}}$ nicht enthalten. Anstelle $t_{A, \text{ist umstieg}}$ enthalten die unter Verwendung von \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 gelernten Regeln für die „Warten in Regelwartezeit“-Maßnahmen häufig die Bedingungen:

$$t_{A, \text{Regelwartezeit}} \geq x \wedge t_{H_B, F_{\text{Zubringer}}, \text{ankunft verspaetung}} \leq y \wedge t_{A, \text{soll umstieg}} \geq z \rightarrow \text{Maßnahme} = W_r$$

Interessant daran ist, dass sich das Attribut $t_{A, \text{ist umstieg}}$ tatsächlich durch eine Kombination der beiden Werte $t_{H_B, F_{\text{Zubringer}}, \text{ankunft verspaetung}}$ und $t_{A, \text{soll umstieg}}$ annähern lässt:

$$t_{A, \text{ist umstieg}} \approx t_{A, \text{soll umstieg}} - t_{H_B, F_{\text{Zubringer}}, \text{ankunft verspaetung}}$$

Maßnahmen	\mathbb{F}_1	\mathbb{F}_2	\mathbb{F}_3	\mathbb{F}_4
Alternativanschluss, Weiterfahren ohne Warten	25	0	0	0
Warten mit Wartezeitüberschreitung, Gleiswechsel	11, 11	0	0	0
Gleiswechsel, Alternativanschluss	0	0	0	0
Gleiswechsel, Weiterfahren ohne Warten	0	0	0	0
Weiterfahren ohne Warten	81, 84	53, 73	57, 68	56, 61
Gleiswechsel	91, 16	91, 59	90, 6	90, 91
Warten in Regelwartezeit, Gleiswechsel	0	0	0	0
Warten in Regelwartezeit	84, 77	45, 31	45, 6	52, 83
Warten mit Wartezeitüberschreitung, Alternativanschluss	0	0	0	0
Warten mit Wartezeitüberschreitung	82, 21	50, 94	51, 11	50, 14
Alternativanschluss	77, 65	72, 62	75, 32	70, 24

Tabelle 7.13: Precision-Ergebnisse in Prozent für \mathbb{K}_2

Dennoch werden unter Verwendung von \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 zwischen 49% und 55% der Anschlusskonflikte mit der Maßnahme „Warten in Regelwartezeit“ nicht als solche erkannt, wie die Ergebnisse in Tabelle 7.12 zeigen. Ein Großteil der nicht erkannten Maßnahmen wird anschließend von der Defaultregel abgefangen und damit der Maßnahme „Warten mit Wartezeitüberschreitung“ zugeordnet. Im Vergleich dazu wurden unter Verwendung von \mathbb{F}_1 nur knapp 9% der „Warten in Regelwartezeit“-Maßnahmen nicht als solche erkannt. Die schlechten Ergebnisse für \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 sind darin begründet, dass das Attribut $t_{H_B, F_{\text{Abbringer}}}$, abfahrt verspaetung nicht mit den vorliegenden Attributen durch den Regellerner ersetzt werden kann.

Bei der Betrachtung der Regeln für die Weiterfahren-Maßnahmen fällt auf, dass sie sehr ähnlich zu den „Warten in Regelwartezeit“-Regeln aufgebaut sind. Der Hauptunterschied zwischen den beiden Regelmengen ist, dass die „Warten in Regelwartezeit“-Regeln jeweils eine Regelwartezeit-Bedingung enthalten. Ebenso wird die Substitution des $t_{A, \text{ist umstieg}}$ Attributs durch eine Kombination der $t_{H_B, F_{\text{Zubringer}}}$, ankunft verspaetung und $t_{A, \text{soll umstieg}}$ Attribute sowohl bei den „Warten in Regelwartezeit“-Regeln, als auch bei einem Großteil der Weiterfahren-Regeln, für \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 angewendet. Ebenso werden zwischen 47% bis 55% der „Weiterfahren ohne Warten“-Maßnahmen unter Verwendung von \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 nicht als solche erkannt und letztendlich von der Defaultregel abgefangen und der Maßnahme „Warten mit Wartezeitüberschreitung“ zugeordnet.

Alle vier für \mathbb{K}_1 gelernten Regelmengen enthalten mindestens eine Regel für die Identifikation von Zubringergleiswechsel-Maßnahmen mit der Bedingung: $t_{F_{\text{Abbringer}}}$, donehalte ≤ 0 . Außerdem enthalten die Regelmengen für \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 eine weitere Regel für die Erkennung von Zubringergleiswechsel-Maßnahmen, welche die Soll-Abfahrtszeit des Abbringers testet. Für \mathbb{F}_2 und \mathbb{F}_4 enthält diese Regel das Attribut $t_{H_B, F_{\text{Abbringer}}}$, soll abfahrt. Für \mathbb{F}_3 enthält die Regel hingegen zwei Bedingungen mit den Attributen $d_{H_B, F_{\text{Zubringer}}}$, soll ankunft wochentag und $t_{H_B, F_{\text{Abbringer}}}$, soll abfahrt uhrzeit. Alle Regeln mit einer Bedingung die einen Timestamp enthalten haben allerdings den Nachteil, dass sie nur für abgeschlossene Szenarien, die einen frühesten und einen spätesten Zeitpunkt enthalten, sinnvoll sind. In einem offenen Szenario können hingegen neue Anschlusskonflikte hinzukommen, für die die Regel nicht mehr gilt, da sie aufgrund einer anderen Sachlage erstellt wurde. In der Praxis handelt es sich um ein offenes Szenario, in dem es nach dem Erzeugen der Regelmenge einen neuen Betriebstag gibt, dessen Fahrplan und Anschlusskonflikte vorher nicht bekannt sind.

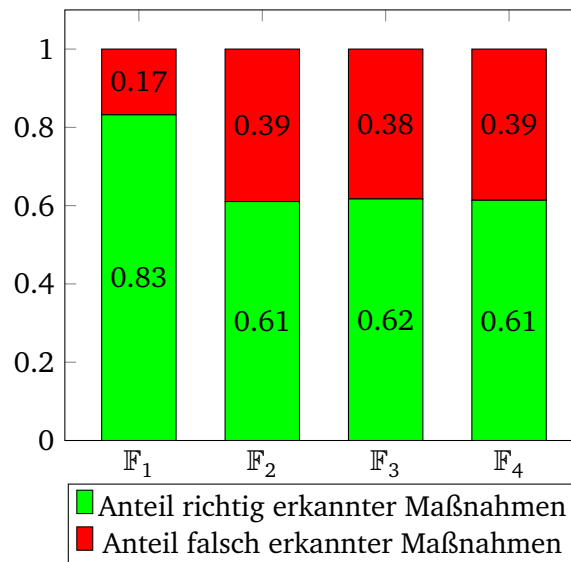


Abbildung 7.3: Vergleich der richtig und falsch erkannten Maßnahmen in \mathbb{K}_2 unter Verwendung aller Featuregruppen.

Maßnahmen	F ₁	F ₂	F ₃	F ₄
Alternativanschluss, Weiterfahren ohne Warten	8,33	0	0	0
Warten mit Wartezeitüberschreitung, Gleiswechsel	2,22	0	0	0
Gleiswechsel, Alternativanschluss	0	0	0	0
Gleiswechsel, Weiterfahren ohne Warten	0	0	0	0
Weiterfahren ohne Warten	90,43	42,28	47,53	42,28
Gleiswechsel	84,4	80,82	81,33	81,84
Warten in Regelwartezeit, Gleiswechsel	0	0	0	0
Warten in Regelwartezeit	92,09	41,73	41,01	40,29
Warten mit Wartezeitüberschreitung, Alternativanschluss	0	0	0	0
Warten mit Wartezeitüberschreitung	95,02	75,97	74,89	77,06
Alternativanschluss	70,21	64,89	61,7	62,77

Tabelle 7.14: Recall-Ergebnisse in Prozent für \mathbb{K}_2

7.4.2 Klassenzusammenstellung \mathbb{K}_2

Die Tabelle 7.13 zeigt die Precision-Ergebnisse für die verschiedenen Maßnahmen aus K_2 in Kombination mit den betrachteten Featuremengen. Tabelle 7.14 enthält die Recall-Ergebnisse, für die Maßnahmen aus K_2 in Kombination mit den Featuremengen. Der Anteil der für \mathbb{K}_2 richtig und falsch erkannter Anschlusskonflikte wird in Abbildung 7.3 in Abhängigkeit der verwendeten Features dargestellt.

Für \mathbb{K}_2 werden, in Abhängigkeit von der verwendeten Featuremenge, zwischen 19 und 27 Regeln gelernt. Dies zeigt, dass ein Zusammenlegen der Gleiswechsel-Maßnahmen aus \mathbb{K}_1 zu einer Reduktion der gelernten Regeln führt. Davon abgesehen steigt die erreichte Accuracy durch das Zusammenlegen der Gleiswechsel-Maßnahmen, um 3% bis 6% für alle betrachteten Featuremengen an. Die Ergebnisse in den Tabellen 7.13 und 7.14 zeigen, dass sich die Accuracy für \mathbb{K}_2 hauptsächlich aufgrund der Zusammenführung der Gleiswechsel-Maßnahmen gegenüber \mathbb{K}_1 verbessert. Während für \mathbb{K}_1 für die Abbringergleiswechsel-Maßnahmen ein Recall zwischen 33% und 40% und für die Zubringergleiswechsel-Maßnahmen zwischen 73% und 77% erreicht wird, kann für die zusammen-

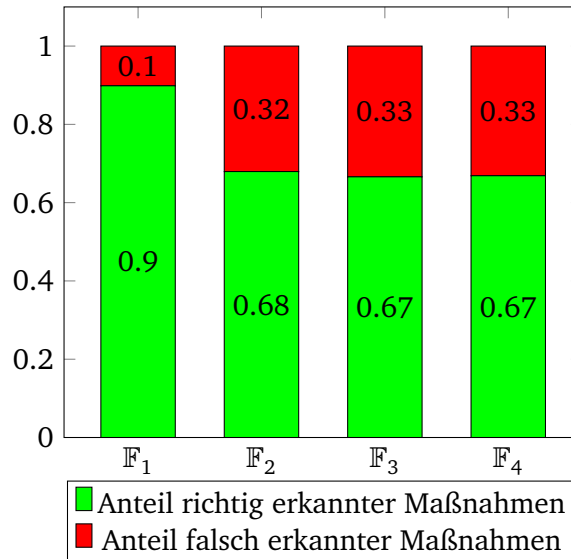


Abbildung 7.4: Vergleich der richtig und falsch erkannten Maßnahmen in \mathbb{K}_3 unter Verwendung aller Featuregruppen.

Maßnahmen	\mathbb{F}_1	\mathbb{F}_2	\mathbb{F}_3	\mathbb{F}_4
Weiterfahren ohne Warten	92,1	63,61	59,35	59,93
Gleiswechsel	89,65	90,83	90,78	91,45
Warten in Regelwartezeit	92,31	51,35	51,82	54,74
Warten mit Wartezeitüberschreitung	91,91	59,51	57,81	58,44
Alternativanschluss	67,42	71,08	74,36	67,86

Tabelle 7.15: Precision-Ergebnisse in Prozent für \mathbb{K}_3

gefassten Gleiswechsel-Maßnahmen aus \mathbb{K}_2 eine Accuracy von 80% bis 84% erreicht werden. Insgesamt werden für \mathbb{K}_2 die meisten Regeln unter Verwendung von \mathbb{F}_4 mit 27 Regeln gelernt, während die wenigsten Regeln unter Verwendung von \mathbb{F}_3 gelernt wurden.

Bei Betrachtung der Ergebnisse in Abbildung 7.3 fällt auf, dass die Accuracy unter Verwendung der Featuremengen \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 mit Abweichungen von unter einem Prozent sehr nah beieinander liegen, während die Accuracy für \mathbb{F}_1 mit Abstand am größten ist.

Genauso wie für \mathbb{K}_1 enthalten fast alle Regeln für die Erkennung von Maßnahmenkombinationen in \mathbb{K}_2 mit \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 sehr spezifische Bedingungen, die nur von wenigen Anschlusskonflikten abgedeckt werden. Diese Bedingungen enthalten Betriebsstellen, Fahrtrichtungen und Fahrtrlinien.

Bei der Betrachtung der gelernten Regeln für Gleiswechsel-Maßnahmen fällt auf, dass sie sehr stark den Zubringergleiswechsel-Regeln für \mathbb{K}_1 ähneln, zum Beispiel dadurch, dass sie alle eine Regel mit der Bedingung $t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0$ und eine Regel mit der einer Soll-Abfahrtszeit des Abbringers enthalten. Dies ist zu erwarten, da die Gleiswechsel-Regeln für \mathbb{K}_1 auch die Zubringergleiswechsel-Maßnahmen aus \mathbb{K}_1 abdecken sollen.

Für die übrigen Maßnahmen in \mathbb{K}_2 werden großteils ähnliche Ergebnisse, mit Abweichungen von bis zu 6% in der Precision und von bis zu 11% im Recall, wie für \mathbb{K}_1 erreicht.

Maßnahmen	\mathbb{F}_1	\mathbb{F}_2	\mathbb{F}_3	\mathbb{F}_4
Weiterfahren ohne Warten	93,52	57,72	56,79	53,09
Gleiswechsel	84,14	81,07	80,56	79,28
Warten in Regelwartezeit	94,96	54,68	51,08	53,96
Warten mit Wartezeitüberschreitung	95,89	69,05	67,32	71,21
Alternativanschluss	63,83	62,77	61,7	60,64

Tabelle 7.16: Recall-Ergebnisse in Prozent für \mathbb{K}_3

7.4.3 Klassenzusammenstellung \mathbb{K}_3

Die Tabelle 7.15 zeigt die Precision-Ergebnisse für die verschiedenen Maßnahmen aus \mathbb{K}_3 , in Kombination mit den betrachteten Featuremengen. Tabelle 7.16 zeigt die Recall-Ergebnisse, für die Maßnahmen aus \mathbb{K}_3 in Kombination mit den Featuremengen. Der Anteil der für \mathbb{K}_3 richtig sowie falsch erkannter Anschlusskonflikte wird in Abhängigkeit der verwendeten Features in Abbildung 7.4 dargestellt.

Insgesamt werden für \mathbb{K}_3 , in Abhängigkeit von der betrachteten Featuremenge, zwischen 16 und 22 Regeln gelernt. Der Unterschied zwischen \mathbb{K}_2 und \mathbb{K}_3 ist, dass in \mathbb{K}_3 alle Anschlusskonflikte aus \mathbb{K}_2 , für die mehrere Maßnahmen getroffen wurden, entfernt wurden. Dies wirkt sich positiv auf die Accuracy der Versuchsaufbauten mit \mathbb{K}_3 im Vergleich zu den Versuchsaufbauten mit \mathbb{K}_2 aus und führt dazu, dass bei Betrachtung aller Featuremengen, ein Anstieg der Accuracy zwischen 5% und 6%, im Vergleich zu den Ergebnissen für \mathbb{K}_2 , erreicht wird. Für \mathbb{K}_3 wurden unter Einsatz der Featuremenge \mathbb{F}_1 mit 22 Regeln die meisten Regeln gelernt, wohingegen unter Verwendung von \mathbb{F}_2 die wenigsten Regeln gelernt wurden.

Die Zahlen in Tabelle 7.9 zeigen, dass die Anzahl der Regeln für \mathbb{K}_3 , die im Verdacht stehen zu sehr an die Trainingsdaten angepasst zu sein, im Vergleich zu den Zahlen für \mathbb{K}_1 und \mathbb{K}_2 , mit Ausnahme von \mathbb{F}_1 , abgenommen haben. Dies zeigt, dass die Maßnahmenkombinationen ein Grund für die Anpassung der gelernten Regeln an die Trainingsdaten sind. Außerdem ist bei Betrachtung der gelernten Regeln in Anhang A erkennbar, dass ein Großteil der übrigen an die Trainingsdaten angepassten Regeln der Erkennung von Alternativanschluss-Maßnahmen dienen.

Genauso wie bei den Accuracy-Ergebnissen für \mathbb{K}_2 fällt bei Betrachtung der Ergebnisse für \mathbb{K}_3 in Abbildung 7.4 auf, dass die Accuracy unter Verwendung der Featuremengen \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 mit Abweichungen von bis zu 1% sehr nah beieinander liegen, während die Accuracy für \mathbb{F}_1 mit Abstand am größten ist. Der Grund dafür, dass für \mathbb{F}_1 die höchste Accuracy erzielt wurde, ist, dass \mathbb{F}_1 auch Informationen enthält, die erst durch das Treffen einer Dispositionsmaßnahme festgelegt werden. Die Accuracy-Ergebnisse für \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 zeigen, dass es in Hinblick auf die Anzahl korrekt erkannter Maßnahmen kaum einen Unterschied macht, welche Art zeitlicher Attribute verwendet werden.

Wie auch für \mathbb{K}_1 und \mathbb{K}_2 verwenden alle „Warten in Regelwartezeit“-Regeln für die Featuremengen \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 eine Substitution für das Attribut $t_{A, \text{ist umstieg}}$, das in fast allen „Warten in Regelwartezeit“-Regeln mit \mathbb{F}_1 eingesetzt wird. Dabei wird, wie auch bei den anderen Klassenzusammenstellungen, die Bedingung mit dem Attribut $t_{A, \text{ist umstieg}}$ durch zwei Bedingungen mit den Attributen $t_{A, \text{soll umstieg}}$ und $t_{H_B, F_{\text{Zubringer}}, \text{ankunft verspätung}}$ ersetzt. Genau wie für die anderen Klassenzusammenstellungen werden auch für \mathbb{K}_3 unter Verwendung von \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 bis zu 23% schlechtere Accuracy-Ergebnisse erzielt als für \mathbb{F}_1 . Auch für \mathbb{K}_3 ist dies darin begründet, dass das Attribut $t_{H_B, F_{\text{Abbringer}}, \text{abfahrt verspätung}}$ nicht mit den in \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 vorliegenden Attributen durch den Regellerner ersetzt werden kann.

Alle Regelmengen für \mathbb{K}_3 haben gemeinsam, dass ein Großteil der enthaltenen Regeln für die Maßnahmen „Weiterfahren ohne Warten“- und Alternativanschluss gelernt wurden.

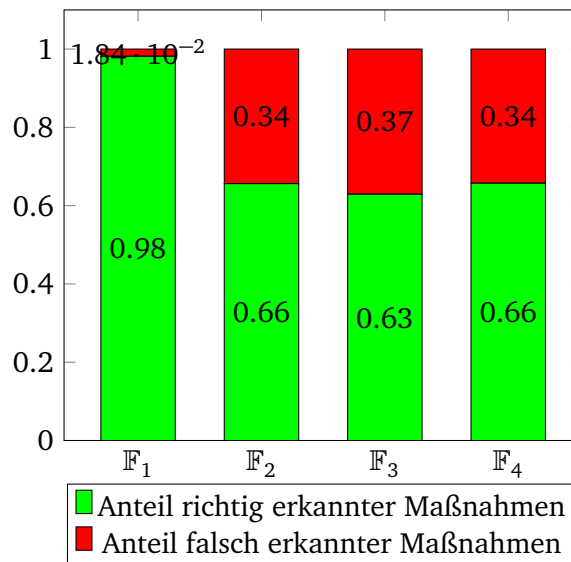


Abbildung 7.5: Vergleich der richtig und falsch erkannten Maßnahmen in \mathbb{K}_4 unter Verwendung aller Featuregruppen.

Maßnahmen	F_1	F_2	F_3	F_4
Weiterfahren ohne Warten	99,37	70,36	64,95	72,22
Warten in Regelwartezeit	95,71	59,2	55,88	53,79
Warten mit Wartezeitüberschreitung	98,08	64,62	63,6	65,15

Tabelle 7.17: Precision-Ergebnisse in Prozent für \mathbb{K}_4

7.4.4 Klassenzusammenstellung \mathbb{K}_4

Die Tabelle 7.17 zeigt die Precision-Ergebnisse für die verschiedenen Maßnahmen aus \mathbb{K}_4 , in Kombination mit den betrachteten Featuremengen. Tabelle 7.18 enthält die Recall-Ergebnisse, für die Maßnahmen aus \mathbb{K}_4 in Kombination mit den Featuremengen. Der Anteil der für \mathbb{K}_4 richtig und falsch erkannter Anschlusskonflikte wird in Abbildung 7.5 in Abhängigkeit der verwendeten Features dargestellt.

Für \mathbb{K}_4 , wurden abhängig von der betrachteten Featuremenge, zwischen 6 und 14 Regeln gelernt. \mathbb{K}_4 enthält alle Anschlusskonflikte, für die eine der drei Warte-Maßnahmen „Weiterfahren ohne Warten“, „Warten in Regelwartezeit“ und „Warten mit Wartezeitüberschreitung“ getroffen wurden. Bei der Evaluation von \mathbb{K}_4 wurde mit Verwendung der Featuremenge F_1 , die auch Attribute enthält, die dem Disponenten zur Dispositionszeit nicht bekannt sind, die höchste Accuracy von 98% erreicht. Während für die restlichen Featuremengen F_2 , F_3 und F_4 eine schlechtere Accuracy von 62% bis 65% erzielt wurde. Wie Tabelle 7.9 zeigt, wurden insgesamt, bei allen für \mathbb{K}_4 evaluierten Featuremengen, die wenigsten Regeln gelernt, die im Verdacht stehen zu stark an die Trainingsdaten angepasst zu sein.

Maßnahmen	F_1	F_2	F_3	F_4
Weiterfahren ohne Warten	96,91	60,8	62,35	64,2
Warten in Regelwartezeit	96,4	53,24	54,68	51,08
Warten mit Wartezeitüberschreitung	99,57	72,73	65,8	71,21

Tabelle 7.18: Recall-Ergebnisse in Prozent für \mathbb{K}_4

Bei der Betrachtung der Accuracy-Ergebnisse von \mathbb{K}_4 , im Vergleich zu den Ergebnissen von \mathbb{K}_3 fällt auf, dass \mathbb{K}_4 unter Verwendung der Featuremengen \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 eine um 1% bis 3% schlechtere Accuracy erreicht als für \mathbb{K}_3 unter Betrachtung der gleichen Featuremengen. Die einzige Ausnahme dazu stellt die Featuremenge \mathbb{F}_1 dar, bei der mit 98% Accuracy das beste Ergebnis aller Versuchsaufbauten erzielt wird. Bei der Betrachtung der erzeugten Regeln für den Versuchsaufbau mit \mathbb{K}_4 in Kombination mit \mathbb{F}_1 fällt auch auf, dass die meisten darin enthaltenen Regeln einfach verständlich sind. Ein Beispiel dafür ist die Regel:

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_B, F_{\text{Abbringer}}, \text{ abfahrt verspaetung}} \leq 3 \wedge t_{A, \text{ ist umstieg}} \geq 6 \rightarrow \text{Maßnahme} = W_r$$

Bei dieser Regel fällt auf, dass sie die notwendige Bedingung, dass ein Warten in Regelwartezeit nur durchführbar ist, wenn die Abfahrtsverspätung des Abbringers die Regelwartezeit für den Anschluss nicht übersteigt, enthält. Weiterhin folgen alle für \mathbb{F}_1 gelernten Regeln zur Erkennung der Maßnahme „Weiterfahren ohne Warten“ dem Schema:

$$t_{A, \text{ ist umstieg}} \leq x \wedge t_{\text{mindestübergangszeit}} \geq x + 1 \rightarrow \text{Maßnahme} = W_w$$

Dieses Schema stellt ebenfalls eine notwendige Bedingung für den Einsatz der Maßnahme „Weiterfahren ohne Warten“ dar, es beschreibt, dass ein Anschluss nicht mehr gehalten werden kann, wenn die tatsächliche Umsteigezeit des Anschlusses kürzer ist, als die benötigte Mindestübergangszeit am Anschlussbahnhof.

Auch für die Featuremengen \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 werden für die Maßnahme „Weiterfahren ohne Warten“ einige Regeln gelernt, die Wissen über die Mindestübergangszeit am Anschlussbahnhof verwenden. Diese Regeln verwenden ebenfalls die vorgestellte Substitution für das Attribut $t_{A, \text{ ist umstieg}}$ mit zwei Bedingungen der Form:

$$t_{A, \text{ soll umstieg}} \leq x \wedge t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq y$$

Dabei muss die angenäherte $t_{A, \text{ ist umstieg}}$, als Differenz zwischen x und y , in der Regel kleiner oder gleich 3 sein. Dies entspricht bis auf zwei Ausnahmen der minimalen bahnsteiggleichen Übergangszeit aus [2]. Damit beachten die gelernten Regeln für die Erkennung der Maßnahme „Weiterfahren ohne Warten“ implizit, dass ein Anschluss nicht mehr gehalten wird, wenn $t_{A, \text{ ist umstieg}}$ kleiner der Übergangszeit von 3 Minuten ist.

Ebenso wie bei den anderen Klassenzusammenstellungen fällt bei den für die Maßnahme „Warten in Regelwartezeit“ für \mathbb{K}_4 gelernten Regeln auf, dass unter Betrachtung der Featuremengen \mathbb{F}_2 , \mathbb{F}_3 und \mathbb{F}_4 die aus \mathbb{F}_1 weggefallene Umsteigezeit $t_{A, \text{ ist umstieg}}$ häufig durch eine Kombination der zwei Attribute $t_{A, \text{ soll umstieg}}$ und $t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}}$ ersetzt wird. Allerdings fehlt auch hier, wie in den vorherigen Klassenzusammenstellungen, das Attribut $t_{H_B, F_{\text{Abbringer}}, \text{ abfahrt verspaetung}}$ aus \mathbb{F}_1 , um mehr als 54% der „Warten in Regelwartezeit“-Maßnahmen korrekt zu erkennen.

8 Fazit

In dieser Arbeit wurde der Frage nachgegangen, ob sich Verfahren aus dem Gebiet des Maschinellen Lernens dafür eignen, in der Anschlussdisposition eingesetzt zu werden. Detaillierter wurde in dieser Arbeit ein Regellerner dafür eingesetzt, um für den Menschen verständliche Regeln für Maßnahmen der Anschlussdisposition, aufgrund von Maßnahmen-Entscheidungen menschlicher Experten (sogenannter Disponenten), zu lernen. Dafür wurde ein geeignetes Datenmodell mit Anschlüssen, Anschlusskonflikte, Betriebsstellen, Fahrten, Halten und Maßnahmen aufgestellt. Dieses Datenmodell wurde anschließend mit von der Deutschen Bahn zur Verfügung gestellten Betriebsdaten und daraus abgeleiteten Anschlusskonflikten und Maßnahmen befüllt. Dabei wurden die folgenden Erkenntnisse gewonnen:

Beim Einlesen der Betriebsdaten der Deutschen Bahn ist aufgefallen, dass die Daten keine gelabelten Anschlusskonflikte und Maßnahmen dazu enthalten, sondern mehrere Tabellen mit Ereignismeldungen, die Rückschlüsse auf Maßnahmen zu Anschlusskonflikten erlauben. Aus diesem Grund wurden die Maßnahmen und die dazugehörigen Anschlusskonflikte aufgrund der Ereignismeldungen abgeleitet. Allerdings ist für die Ereignismeldungen nicht angegeben, aus welchem Grund die Ereignisse durchgeführt wurden, weshalb nicht mit absoluter Sicherheit eine Aussage darüber getroffen werden kann, ob ein solches Ereignis als Dispositionsmaßnahme oder aus einem anderen Grund durchgeführt wurde.

Weiterhin ist beim Einlesen der Daten der Deutschen Bahn aufgefallen, dass verschiedene Dateiformate und Bezeichnungen für die Daten verwendet wurden. So wurden zum Beispiel die Betriebsstellen an einigen Stellen mit ihrer RIL100 und an anderer Stelle mit ihrem Namen angegeben.

Bei der Betrachtung der Klassenzusammenstellungen \mathbb{K}_1 und \mathbb{K}_2 fällt auf, dass darin Maßnahmenkombinationen enthalten sind, die in der Praxis aus fachlichen Gründen nicht durchgeführt werden können. Dazu gehören alle Maßnahmenkombinationen mit der Maßnahme „Weiterfahren ohne Warten“, mit Ausnahme der Maßnahmenkombination „Weiterfahren ohne Warten“ und „Alternativanschluss“. Ebenso gehören dazu alle Maßnahmenkombinationen mit der Maßnahme „Alternativanschluss“ mit der oben genannten Ausnahme. Dies ist darin begründet, dass wenn der Abbringer nicht auf den Zubringer wartet, der Anschluss nicht gehalten wird, womit eine weitere Maßnahme unnötig wird. Ebenso wird ein Alternativanschluss in der Praxis durchgeführt, wenn ein Anschluss nicht mehr gehalten werden kann, womit ebenfalls weitere Maßnahmen unnötig werden. Solche Maßnahmenkombinationen sind in \mathbb{K}_1 und \mathbb{K}_2 enthalten, weil die von der Deutschen Bahn zu Verfügung gestellten Daten, wie in Abschnitt 5.2.8 beschrieben, keine gelabelten Maßnahmen enthalten.

Insgesamt stellt sich der Einsatz eines Regellerners für die Erkennung von Maßnahmen zur Lösung von gegebenen Anschlusskonflikten als sehr vielversprechend heraus. Dies zeigen die durchweg guten Ergebnisse in Tabelle 7.10, wo erkennbar ist, dass für alle Klassenzusammenstellungen, in Kombination mit allen Featuremengen, eine höhere Accuracy als die verwendete Baseline erreicht wurde. Durch den Einsatz eines Regellerners wurden im Vergleich zur Baseline mindestens 12% mehr Maßnahmen korrekt erkannt.

Wie die Ergebnisse in Tabelle 7.10 für \mathbb{K}_2 im Vergleich \mathbb{K}_1 zeigen, führt ein Zusammenlegen der verschiedenen Gleiswechsel-Maßnahmen zu besseren Accuracy-Ergebnissen, als bei Verwendung der einzelnen Abbringer- und Zubringergleiswechsel-Maßnahmen. Der Grund dafür ist, dass die Vorbedingungen für einen Abbringer- und Zubringergleiswechsel einander sehr ähnlich sind, außerdem wird durch das Zusammenlegen, der Gleiswechselmaßnahmen, die Menge der möglichen Maßnahmenkombinationen reduziert. Dadurch wiederum erhöht sich die Anzahl der Trainingsbeispiele für die Maßnahmen und Maßnahmenkombinationen.

Bei Betrachtung der Ergebnisse für \mathbb{K}_1 und \mathbb{K}_2 , in den Tabellen 7.12 und 7.14, hat sich herausgestellt, dass die Maßnahmen für fast alle Anschlusskonflikte mit Maßnahmenkombinationen nicht korrekt erkannt wurden. Häufig wird für die meisten Anschlusskonflikte mit einer Maßnahmenkombination lediglich eine der enthaltenen Maßnahmen erkannt, und nicht die vollständige Maßnahmenkombination. Enthält die Maßnahmenkombination eine Warte-Maßnahme, wird häufig statt der korrekten Warte-Maßnahme eine andere Warte-Maßnahme erkannt. Daher führt der Ansatz, nur Anschlusskonflikte mit Einzelmaßnahmen zu betrachten, zu besseren Ergebnissen sowohl in der Anzahl gelernter Regeln als auch in der Anzahl korrekt erkannter Maßnahmen, wie die Ergebnisse in den Tabellen 7.10, 7.8, 7.16 und 7.18 zeigen. Ein möglicher Grund dafür ist, dass für die Maßnahmenkombinationen vergleichsweise sehr wenige Trainingsbeispiele vorliegen.

Eine weitere Erkenntnis aus der durchgeführten Evaluation ist, dass eine größere Anzahl möglicher Einzelmaßnahmen zu besseren Ergebnissen führt, als eine geringere Anzahl, wie die Ergebnisse für \mathbb{K}_3 und \mathbb{K}_4 in Tabelle 7.10 zeigen, wo durchweg für \mathbb{K}_3 , mit Ausnahme für die Featuremenge \mathbb{F}_1 , höhere Erkennungsraten erzielt werden konnten, als für \mathbb{K}_4 . Dies ist interessant, da die Alternativanschluss- und Gleiswechsel-Maßnahmen, die in \mathbb{K}_3 und nicht in \mathbb{K}_4 enthalten sind, in ihren Eigenschaften anders als die Warte-Maßnahmen sind, die sowohl in \mathbb{K}_3 und \mathbb{K}_4 enthalten sind.

Zudem ist während der Evaluation in Tabelle 7.9 aufgefallen, dass für die Klassenzusammenstellungen \mathbb{K}_1 und \mathbb{K}_2 je nach betrachteter Featuremenge, bis zu etwa einem Drittel aller Regeln eine oder mehrere Bedingungen enthalten, die darauf hinweisen zu sehr an die Trainingsdaten angepasst zu sein. Ein Grund für die Generierung dieser Regeln ist, dass für die Maßnahmenkombinationen, im Vergleich zu den einzelnen Maßnahmen, nur sehr wenige Anschlusskonflikte vorliegen.

Bei der Unterscheidung in Timestamp-Attribute und Kombinationen aus Uhrzeit- und Wochentag-Attributen kann, bei Betrachtung der Ergebnisse für die verschiedenen Klassenzusammenstellungen, kein Einfluss auf die erreichten Accuracy-Ergebnisse festgestellt werden, da für jede Klassenkombination der Einsatz einer anderen Featuremenge zu der höchsten Accuracy führt, wie die Ergebnisse in Tabelle 7.10 zeigen. Ebenso wirkt sich die Auswahl, ob Timestamp-Attribute oder Kombinationen aus Uhrzeit- und Wochentag-Attributen eingesetzt werden, nicht aussagekräftig auf die Anzahl potentiell zu stark an die Daten angepasster Regeln aus, wie die Ergebnisse in Tabelle 7.9 zeigen, da der Anteil zu stark angepasster Regeln, je nach Klassenkombination und Featuremenge, wechselt.

Des Weiteren wurde in der Evaluation festgestellt, dass nach Entfernen der Attribute aus \mathbb{F}_1 , die erst nach Festlegung der Dispositionsentscheidung bekannt sind, ein starker Abfall von bis zu 35% in der Accuracy eintritt. Zu diesen Attributen, die dem Disponenten zur Zeit seiner Maßnahmenentscheidung noch nicht bekannt sind, gehören unter anderem die Abfahrtszeit des Abbringers, die tatsächliche Umsteigezeit der Fahrgäste und die tatsächliche Warte- und Haltezeit des Abbringers, da diese erst durch die Entscheidung der Maßnahme festgelegt werden. Die Ergebnisse aller Klassenzusammensetzungen für \mathbb{F}_1 zeigen, dass Accuracy-Ergebnisse zwischen 76% und 98% erreicht werden können, wenn die Attribute, die in \mathbb{F}_1 verwendet wurden und dem Disponenten zur Zeit der Maßnahmenentscheidung nicht vorliegen, approximiert werden könnten.

Zusammenfassend ist der Einsatz eines Regellerners in Kombination mit den für diese Arbeit erhaltenen Daten in einem produktiven Umfeld nicht realistisch. Dies ist darin begründet, dass die erhaltenen Daten zum einen keine gelabelten Anschlusskonflikte und Maßnahmen und zum anderen keine Prognoseinformationen zur Zeit der Dispositionsentscheidung, als Annäherung an die Istdaten, die dem Disponenten während der Dispositionsentscheidung nicht bekannt sind, enthalten. Ebenso fehlen in dem hier aufgestellten Modell Informationen über die Anzahl Umsteiger, Zusteiger und Aussteiger an den Anschlüssen, die dabei helfen können, bessere Regeln für die Auswahl der Maßnahmen für einen Anschlusskonflikt zu lernen.

9 Zukünftige Ansätze für die Automatisierte Anschlussdisposition

In den vorherigen Kapiteln wurde erläutert, wie ein Regellerner aus dem Bereich des Maschinellen Lernens für die Erkennung von Maßnahmen für einen gegebenen Anschlusskonflikt eingesetzt werden kann. Dabei wurden verschiedene Annahmen getroffen, mit denen die vorgestellten Ergebnisse erzielt wurden. Neben diesen Annahmen existieren einige weitere alternative Ansätze und Ideen, die während der Bearbeitung dieser Arbeit entstanden sind, denen in zukünftigen Arbeiten nachgegangen werden kann. Sie haben das Potential, die in dieser Arbeit erreichten Ergebnisse zu verbessern. Diese Ansätze und Ideen beziehen sich auf das zugrunde liegende Datenmodell, das in Abschnitt 5.1 eingeführt wurde, als auch auf die Auswahl der Features, die während der Maßnahmen-Klassifikation eingesetzt werden, und auf die Auswahl des gewählten Lernverfahrens.

Nach der Realisation der im Folgenden vorgestellten Ansätze und Ideen sollte die Anwendbarkeit von durch einen Regellerner gelernten Regeln in einem produktiven Umfeld realistischer sein.

9.1 Betrachtungsebenen für Anschlusskonflikte

Eine Möglichkeit die Forschung im Bereich der automatisierten Anschlussdisposition weiterzuführen ist die Betrachtung der Anschlusskonflikte zu wechseln. Dabei existieren mehrere Möglichkeiten, Anschlusskonflikte zu betrachten. Diese reichen von der Betrachtung einzelner Anschlusskonflikte, bestehend aus einem Zubringer, einem Abbringer, einem Anschlussbahnhof und einigen anderen Parametern, bis hin zu Verkettungen von Anschlusskonflikten, die als Einheit betrachtet, und im Folgenden erläutert werden:

Einzelne Anschlusskonflikte

Die einfachste Art Anschlusskonflikte zu betrachten ist, Anschlusskonflikte individuell, ohne Auswirkungen auf andere Anschlusskonflikte, zu betrachten. Dabei müssen lediglich die Parameter eines Anschlusskonflikts betrachtet werden, ohne dabei andere Anschlusskonflikte die als Folgekonflikte auftreten zu betrachten. Genauso müssen für die Wahl der Maßnahme nur die Informationen des Anschlusses betrachtet werden, wie die Umsteigezahl des Zubringers in den Abbringer. In dieser Arbeit wurde diese Betrachtung der Anschlusskonflikte gewählt.

Anschlusskonflikte eines Zugs an einem Halt

Eine weitere Möglichkeit Anschlusskonflikte zu betrachten ist, alle Anschlusskonflikte eines Zugs an einem Halt zusammen zu betrachten. Dabei kann zwischen der Betrachtung des Zubringerzugs und des Abbringerzugs unterschieden werden. Wird der Abbringer eines Anschlusskonfliktes betrachtet, müssen alle anderen Anschlusskonflikte an dem Halt, deren Zubringer diesen Zug als Abbringer besitzen, als Parameter für die Entscheidung der Maßnahme mit einfließen. Wird hingegen der Zubringer eines Anschlusskonfliktes betrachtet, müssen alle anderen Anschlusskonflikte an dem Halt, deren Abbringer diesen Zug als Zubringer verwenden, als Parameter für die Entscheidung der Maßnahme mit einfließen.

Alle Anschlusskonflikte eines Zugs an allen Halten

Eine weitere Art auf Anschlusskonflikte zu blicken ist, alle Anschlusskonflikte eines Zugs zu betrachten. Im Vergleich zu der Betrachtungsart „Anschlusskonflikte eines Zubringers an einem Halt“ bedeutet dies, dass nicht nur die Anschlusskonflikte eines Zugs an einem Bahnhof, sondern an allen Bahnhöfen, die von dem Zug angefahren werden, in die Auswahl der Maßnahme für den Anschlusskonflikt einfließen. Dazu gehört auch ein Augenmerk auf die Folgen einer getroffenen Maßnahme für einen Anschlusskonflikt auf die darauf folgenden Anschlüsse, an später angefahrenen Bahnhöfen, zu legen. Beispielsweise führt das Wartenlassen eines Abbringers dazu, dass der Abbringer an den darauf folgenden Bahnhöfen automatisch eine Verspätung hat, was weitere Anschlusskonflikte auslösen kann.

9.2 Einschränkungen bei der Verwendung von Maschinellen Lernverfahren bei der Anschlussdisposition

Einige Eigenschaften der Anschlussdisposition und deren Maßnahmen lassen sich nur schwierig für das Maschinelle Lernen modellieren. Dazu gehören zum Beispiel die Kombinierbarkeit von Dispositionsmaßnahmen und die Möglichkeit Maßnahmen unterschiedlich zu konfigurieren. Aus diesem Grund bilden sich auch hier mehrere mögliche Ansätze heraus, mit denen an das Problem der Klassifikation von Maßnahmen herangegangen werden kann.

Eine Eigenschaft vieler Maßnahmen ist es, dass sie in der Regel verschiedene Parameter besitzen, die sich auf den Effekt der Maßnahme auswirken. Ein Beispiel hierfür ist die eigentliche Wartezeit der Maßnahmen „Warten in Regelwartezeit“ und „Warten mit Wartezeitüberschreitung“. Um diese Parametrisierung als Klassen für das Maschinelle Lernen zu unterstützen, muss für jede Parameterbelegung jeder Maßnahme eine eigene Klasse erzeugt werden. Da allerdings potenziell unendlich viele Parameterbelegungen existieren können, müssen auch unendlich viele Klassen gebildet werden.

9.3 Gelabelte Anschlusskonflikte und Maßnahmen

Die für diese Arbeit von der Deutschen Bahn zu Verfügung gestellten Daten enthalten keine expliziten Angaben über die konfliktierenden Anschlüsse und durchgeführten Maßnahmen. Aus diesem Grund wurden die Anschlusskonflikte und die dazugehörigen Maßnahmen in dieser Arbeit aufgrund von Zusammenhängen aus den Daten abgeleitet. Weil die Daten allerdings auch keine Informationen zu den Gründen für die Durchführung der Maßnahmen enthalten, kann nicht ausgeschlossen werden, dass Maßnahmen fälschlicherweise als solche erkannt wurden, siehe auch Kapitel 5.2.8. Um dieses Problem auszuschließen, sollte eine weitere Evaluation mit gelabelten Daten, in denen explizit enthalten ist, für welchen Anschluss ein Anschlusskonflikt vorliegt und welche Maßnahmen dafür getroffen wurden, durchgeführt werden. Dadurch kann auch das in Kapitel 8 angesprochene Problem, der Identifikation von Maßnahmenkombinationen, die nicht in der Praxis auftreten, gelöst werden.

9.4 Einführung weiterer Attribute

Weiterhin könnte in Zukunft erforscht werden, wie sich die Einführung neuer Attribute auf das Lernen von Dispositionsregeln auswirkt. Diese Attribute können entweder durch weitere Eingabedaten direkt hinzugefügt, durch Kombinieren, oder anderes Aufbereiten bereits vorhandener Attribute gebildet werden. Eine interessante Idee für eine andere Aufbereitung eines Attributs ist die Diskretisierung der Ankunft- und Abfahrt-Uhrzeiten zum Beispiel in Stoßzeiten und Nicht-Stoßzeiten. Dabei sollte auch

versucht werden, ob die in dieser Arbeit evaluierten Features aus \mathbb{F}_1 , die dem Disponenten zur Dispositionszeit nicht vorliegen, anderweitig approximiert werden können. Zum Beispiel könnte die Abfahrtszeit des Abbringers durch Verwendung von richtigen Prognosedaten, die dem Disponenten zur Zeit der Entscheidungsfindung vorliegen, approximiert werden.

9.5 Weitere und umfangreichere Trainingsdaten

Für diese Arbeit wurden Trainingsdaten mit Informationen, mit einem Umfang von drei Arbeitstagen, von der Deutschen Bahn zu Verfügung gestellt. In den Trainingsdaten konnten insgesamt 1.472 Anschlusskonflikte erkannt werden. Dabei liegen einige Maßnahmenkombinationen nur für sehr wenige Anschlusskonflikte vor, was den Prozess des Regellerns erschwert.

Aus diesem Grund ist eine weitere Evaluierung mit Trainingsdaten über einen längeren Zeitraum sinnvoll, um sich ein besseres Bild von der Verwendbarkeit von Techniken aus dem Maschinellen Lernen für die Anschlussdisposition machen zu können. Ebenso wäre eine Evaluation mit besser gepflegten Daten sinnvoll, die in einem einheitlichen Datenformat vorliegen und sich dadurch einfacher aufarbeiten lassen. Dabei könnten zum Beispiel auch Informationen über die Anzahl von umsteigenden, zusteigenden und aussteigenden Fahrgästen betrachtet werden.

Glossary

Istdaten Die Istdaten geben an, wann und wo ein Zug tatsächlich gehalten hat. Die Istdaten fallen erst an, wenn das Ereignis, das sie beschreiben, bereits eingetreten ist. Damit ersetzen sie nachdem Eintreten eines Ereignisses die Prognosedaten. Zum Beispiel steht die Istankunftszeit für einen Zug erst fest, nachdem er in einem Bahnhof angekommen ist. . 37, 41, 45, 60

Solldaten Zu den Solldaten gehören alle Daten, die im Fahrplan enthalten sind. Dazu gehören: wo ein Zug hält, wann er ankommen soll und wann er abfahren soll. Zum Beispiel ist die Sollankunftszeit eines Zuges die Ankunftszeit, nach der ein Zug laut dem Fahrplan an einem Bahnhof ankommt. . 41, 43

Acronyms

ÜZ Übergangszeit. 18, 19, 21, 37, 46, 47, 60, 79

ARFF Attribute-Relation File Format. 3, 34–36, 38, 56, 57

DS100 Druckschrift 100. 17

ILP Integer Linear Programming. 11

ISTP Informationssystem Transportleitung Personenverkehr. 21, 22, 37, 41, 45, 51–53

ITS Intelligent Transportation Systems. 14, 15

RIL100 Richtlinie 100. 17, 41–45, 51, 52, 54, 80

RIS Reisenden Informations System. 21, 22, 37, 41, 51–53

RMC Ready Mixed Concrete. 14

RWZ Regelwartezeit. 5, 18–20, 37, 46–51, 60, 73, 74

WEKA Waikato Environment for Knowledge Analysis. 34, 35, 56

WZÜ Wartezeitüberschreitung. 18, 20

Literatur

- [1] C. Conte und A. Schoebel. „Identifying dependencies among delays“. In: *In Proc. IAROR 2007*. 2007 (siehe S. 12).
- [2] Deutsche Bahn. Anschlüsse im Fahrplanabschnitt Winter 2014-2015. In: *Modul 420.0401 der Ril 420*. 10. Dez. 2014, S. 141 (siehe S. 13, 18, 23, 46, 47, 79).
- [3] Deutsche Bahn. *Benutzerhandbuch (BHB-ISTP). Modul 01 – Allgemeines, Teilmodul 01.01 – Gesamtanwendung ISTP*. Informationssystem Transportleitung Personenverkehr. Deutsch. Benutzerhandbuch. 29. Juli 2007. 56 S. (siehe S. 21).
- [4] Deutsche Bahn. „Fahrgastinformation als Baustein der Mobilitätskette. Das ReisendenInformationssystem (RIS) der Deutschen Bahn“. Deutsch. 15. Jan. 2015 (siehe S. 22).
- [5] Deutsche Bahn. *Kennzahlen der Deutschen Bahn des Jahres 2014*. 2014. http://www.deutschebahn.com/de/konzern/konzernprofil/zahlen_fakten/kennzahlen2014.html (besucht am 23.07.2015) (siehe S. 9).
- [6] Deutsche Bahn. Modul 02 – Anschlüsse disponieren, Teilmodul 02.02 - Anschlüsse bearbeiten & kommunizieren. In: *Benutzerhandbuch (BHB-ISTP)*. 29. Mai 2007, S. 51 (siehe S. 22).
- [7] Deutsche Bahn. *Ril 61501 „Transportleitung Personenverkehr“*. Deutsche Bahn, S. 200 (siehe S. 19).
- [8] Cai Bai-gen und Wang Ju-Zhen. „A study of the expert system for train dispatching“. In: *TENCON '93. Proceedings. Computer, Communication, Control and Power Engineering. 1993 IEEE Region 10 Conference on*. Bd. 2. Okt. 1993, 730–733 vol.2. DOI: 10.1109/TENCON.1993.320083 (siehe S. 10).
- [9] Annabell Berger u. a. „Stochastic Delay Prediction in Large Train Networks“. In: *11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*. Hrsg. von Alberto Caprara und Spyros Kontogiannis. Bd. 20. OpenAccess Series in Informatics (OASICS). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2011, S. 100–111. ISBN: 978-3-939897-33-0. DOI: <http://dx.doi.org/10.4230/OASICS.ATMOS.2011.100>. <http://drops.dagstuhl.de/opus/volltexte/2011/3270> (siehe S. 12).
- [10] Claus Biederbick und Leena Suhl. „Decision Support Tools for Customer-Oriented Dispatching“. English. In: *Algorithmic Methods for Railway Optimization*. Hrsg. von Frank Geraets u. a. Bd. 4359. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2007, S. 171–183. ISBN: 978-3-540-74245-6. DOI: 10.1007/978-3-540-74247-0_8. http://dx.doi.org/10.1007/978-3-540-74247-0_8 (siehe S. 9, 13).
- [11] Bradley, Hax und Magnanti. „Applied Mathematical Programming“. In: 1977. Kap. 9, S. 751. ISBN: 978-0201004649. <http://web.mit.edu/15.053/www/> (besucht am 17.09.2015) (siehe S. 11).
- [12] William W. Cohen. „Fast Effective Rule Induction“. In: *In Proceedings of the Twelfth International Conference on Machine Learning*. Morgan Kaufmann, 1995, S. 115–123 (siehe S. 31).
- [13] A. D’Ariano. „Innovative Decision Support System for Railway Traffic Control“. In: *Intelligent Transportation Systems Magazine, IEEE* 1.4 (Winter 2009), S. 8–16. ISSN: 1939-1390. DOI: 10.1109/MITS.2010.935910 (siehe S. 11).
- [14] Maged Dessouky, Ali Nowroozi und Karen Mourikas. „Bus dispatching at timed transfer transit stations using bus tracking technology“. In: *Transportation Research Part C* (1999), S. 187–208 (siehe S. 14, 15).
- [15] M. Dotoli u. a. „A Decision Support System for real-time rescheduling of railways“. In: *Control Conference (ECC), 2014 European*. Juni 2014, S. 696–701. DOI: 10.1109/ECC.2014.6862177 (siehe S. 9, 11).

-
- [16] Selim Dündar und İsmail Şahin. „Train re-scheduling with genetic algorithms and artificial neural networks for single-track railways“. In: *Transportation Research Part C: Emerging Technologies* 27 (2013). Selected papers from the Seventh Triennial Symposium on Transportation Analysis (TRISTAN VII), S. 1–15. ISSN: 0968-090X. DOI: <http://dx.doi.org/10.1016/j.trc.2012.11.001>. <http://www.sciencedirect.com/science/article/pii/S0968090X12001349> (siehe S. 11).
- [17] Eibe Frank u. a. *Weka 3 - Data Mining with Open Source Machine Learning Software in Java*. <http://www.cs.waikato.ac.nz/ml/weka/> (besucht am 21.09.2015) (siehe S. 34).
- [18] Johannes Fürnkranz. „Separate-and-Conquer Rule Learning“. In: *Artificial Intelligence Review* 13.1 (Feb. 1999), S. 3–54. <http://www.ofai.at/cgi-bin/tr-online?number+96-25> (siehe S. 26, 28).
- [19] Johannes Fürnkranz und Gerhard Widmer. „Incremental Reduced Error Pruning“. In: *Proceedings of the 11th International Conference on Machine Learning (ML-94)*. Hrsg. von William W. Cohen und H. Hirsh. New Brunswick, NJ: Morgan Kaufmann, 1994, S. 70–77. <http://www.ke.informatik.tu-darmstadt.de/~juffi/publications/ml-94.ps.gz> (siehe S. 30).
- [20] Prof. Dr. J. Fürnkranz. „Data Mining and Machine Learning. Learning Individual Rules and Subgroup Discovery“. 1. Okt. 2014. https://www.ke.tu-darmstadt.de/lehre/ws-14-15/mldm/subgroup_discovery.pdf (besucht am 17.06.2015) (siehe S. 26).
- [21] Peter Jackson. *Introduction to Expert Systems*. 3rd. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1998. ISBN: 0201876868 (siehe S. 10).
- [22] Frederik Janssen. „Heuristic Rule Learning“. dissertation. TU Darmstadt, Knowledge Engineering Group, Okt. 2012. <http://tuprints.ulb.tu-darmstadt.de/3135/> (siehe S. 28, 31).
- [23] M. Klemenz und T. Siefer. „Method for an optimized passenger orientated connection management during the planning and the operation process“. In: *8th World Congress on Railway Research (WCRR 2008 Seoul)*. 2008, S. 14 (siehe S. 9).
- [24] Martin Lemnian u. a. „Timing of Train Disposition: Towards Early Passenger Rerouting in Case of Delays“. In: *14th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*. Hrsg. von Stefan Funke und Matúš Mihalák. Bd. 42. OpenAccess Series in Informatics (OASiCs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2014, S. 122–137. ISBN: 978-3-939897-75-0. DOI: <http://dx.doi.org/10.4230/OASiCs.ATMOS.2014.122>. <http://drops.dagstuhl.de/opus/volltexte/2014/4757> (siehe S. 10).
- [25] Hongfei Li u. a. „Improving rail network velocity: A machine learning approach to predictive maintenance“. In: *Transportation Research Part C: Emerging Technologies* 45 (2014). Advances in Computing and Communications and their Impact on Transportation Science and Technologies, S. 17–26. ISSN: 0968-090X. DOI: <http://dx.doi.org/10.1016/j.trc.2014.04.013>. <http://www.sciencedirect.com/science/article/pii/S0968090X14001107> (siehe S. 15).
- [26] Mojtaba Maghrebi, Claude Sammut und Travis Waller. „Reconstruction of an Expert’s Decision Making Expertise in Concrete Dispatching by Machine Learning“. In: *Journal of Civil Engineering and Architecture*. Bd. 7. 12. 2013, S. 1540–1547 (siehe S. 14).
- [27] Christopher D. Manning, Prabhakar Raghavan und Hinrich Schütze. *Introduction to Information Retrieval*. New York, NY, USA: Cambridge University Press, 2008. ISBN: 0521865719, 9780521865715 (siehe S. 12, 15).
- [28] Thomas M. Mitchell. *Machine Learning*. 1. Aufl. New York, NY, USA: McGraw-Hill, Inc., 1997. ISBN: 0070428077, 9780070428072 (siehe S. 14, 15, 25).
- [29] Gordon Paynter, Len Trigg und Eibe Frank. *Attribute-Relation File Format (ARFF)*. 2008. <http://www.cs.waikato.ac.nz/ml/weka/arff.html> (besucht am 15.09.2015) (siehe S. 34).

-
- [30] J.R. Quinlan. „Learning Logical Definitions from Relations“. In: *Machine Learning* 5.3 (), S. 239–266. ISSN: 1573-0565. DOI: [10.1023/A:1022699322624](http://dx.doi.org/10.1023/A:1022699322624). <http://dx.doi.org/10.1023/A:1022699322624> (siehe S. 29).
- [31] J. Rissanen. „Modeling by shortest data description“. In: *Automatica* 14.5 (1978), S. 465–471. ISSN: 0005-1098. DOI: [http://dx.doi.org/10.1016/0005-1098\(78\)90005-5](http://dx.doi.org/10.1016/0005-1098(78)90005-5). <http://www.sciencedirect.com/science/article/pii/0005109878900055> (siehe S. 31).
- [32] H. Schaefer. „Computer-aided train dispatching with expert systems“. In: *Electric Railways in a United Europe, 1995., International Conference on*. März 1995, S. 28–32. DOI: [10.1049/cp:19950175](http://dx.doi.org/10.1049/cp:19950175) (siehe S. 10).
- [33] Anita Schöbel. „A Model for the Delay Management Problem based on Mixed-Integer-Programming“. In: *Electronic Notes in Theoretical Computer Science* 50.1 (2001). {ATMOS} 2001, Algorithmic MeThods and Models for Optimization of RailwayS (Satellite Workshop of {ICALP} 2001), S. 1–10. ISSN: 1571-0661. DOI: [http://dx.doi.org/10.1016/S1571-0661\(04\)00160-4](http://dx.doi.org/10.1016/S1571-0661(04)00160-4). <http://www.sciencedirect.com/science/article/pii/S1571066104001604> (siehe S. 11).
- [34] Rob M. P. Sparing Daniel and Goverde. „Identifying effective guaranteed connections in a multi-modal public transport network“. In: *Public Transport* 5.1 (2013), S. 79–94. ISSN: 1613-7159. DOI: [10.1007/s12469-013-0068-6](http://dx.doi.org/10.1007/s12469-013-0068-6). <http://dx.doi.org/10.1007/s12469-013-0068-6> (siehe S. 11).
- [35] Anselmo Stelzer. *Abschlussbericht Projekt Anschlussdisposition*. Abschlussbericht. Fachgebiet Bahnsysteme und Bahntechnik Technische Universität Darmstadt, 19. Juli 2012. 70 S. (siehe S. 18, 19, 23).
- [36] Bryan S. Todd. *AN INTRODUCTION TO EXPERT SYSTEMS*. 1992, S. 101. ISBN: 0-902928-73-2. <https://www.cs.ox.ac.uk/files/3425/PRG95.pdf> (besucht am 22.07.2015) (siehe S. 10).
- [37] Johanna Törnquist. „Computer-based decision support for railway traffic scheduling and dispatching: A review of models and algorithms“. In: *5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'05)*. Hrsg. von Leo G. Kroon und Rolf H. Möhring. Bd. 2. OpenAccess Series in Informatics (OASICS). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2006. ISBN: 978-3-939897-00-2. DOI: <http://dx.doi.org/10.4230/OASICS.2005.659>. <http://drops.dagstuhl.de/opus/volltexte/2006/659> (siehe S. 11).
- [38] Stefan Wegele und Eckehard Schnieder. „Dispatching of train operations using genetic algorithms“. In: *Proceedings of the 9th International Conference on Computer-Aided Scheduling of Public Transport (CASPT), San Diego, USA*. 2004 (siehe S. 11).
- [39] Stefan Wegele, Roman Slovák und Eckehard Schnieder. „Automatic dispatching of train operations using a hybrid optimisation method“. In: *CD-ROM Proceedings of the 8th World Congress on Railway Research, Seoul, Korea*. 2008 (siehe S. 11).
- [40] Oliver Wolf. „Teilautomatisierung der Anschlussdisposition im Schienenpersonennahverkehr mittels fallbasierten Schließens“. Bachelorarbeit. Fernuniversität in Hagen, 2011 (siehe S. 12).
- [41] Xin Xu und Eibe Frank. *Class JRip*. <http://weka.sourceforge.net/doc/stable/weka/classifiers/rules/JRip.html> (besucht am 06.11.2015) (siehe S. 34).

A Regelmengen

Die Abdeckung der Regeln ist nach der Konklusion im Format (Gesamtanzahl Anschlusskonflikte welche die Regel abdecken, Anzahl Anschlusskonflikte die durch die Regel abgedeckt werden aber eine andere Maßnahme enthalten) angegeben.

A.1 Klassenzusammenstellung \mathbb{K}_1

A.1.1 Featuremenge \mathbb{F}_1

$$\text{Richtung}_{\text{Abbringer}} = \text{Essen Hbf} - \text{Würzburg Hbf} \rightarrow \text{Maßnahme} = A, G_a(3.0, 1.0) \quad (\text{A.1})$$

$$t_{A, \text{ist umstieg}} \leq -1 \wedge t_{F_{\text{Abbringer}}, \text{donefahrtdauer}} \geq 204 \wedge \\ t_{H_{B,F_{\text{Abbringer}}}, \text{soll haltezeit}} \geq 10 \rightarrow \text{Maßnahme} = A, W_w(4.0, 0.0) \quad (\text{A.2})$$

$$\text{Richtung}_{\text{Abbringer}} = \text{Dortmund Hbf} - \text{München Hbf} \wedge B_{\text{Zubringer Start}} = \text{AA} \rightarrow \text{Maßnahme} = A, W_w(2.0, 0.0) \quad (\text{A.3})$$

$$B_{\text{Zubringer Ziel}} = \text{MFSN} \rightarrow \text{Maßnahme} = W_{\ddot{u}}, G_a(2.0, 0.0) \quad (\text{A.4})$$

$$t_{A, \text{ist umstieg}} \leq 0 \wedge t_{H_{B,F_{\text{Abbringer}}}, \text{ist haltezeit}} \geq 14 \rightarrow \text{Maßnahme} = W_w, G_z(6.0, 2.0) \quad (\text{A.5})$$

$$t_{A, \text{ist umstieg}} \leq 4 \wedge t_{A,F_{\text{Zubringer}}, \text{eingangsverspätung}} \geq 24 \wedge \\ t_{F_{\text{Abbringer}}, \text{donefahrtdauer}} \geq 191 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ist haltezeit}} \geq 2 \rightarrow \text{Maßnahme} = W_w, G_z(7.0, 1.0) \quad (\text{A.6})$$

$$B_{\text{Zubringer Start}} = \text{XDKH} \rightarrow \text{Maßnahme} = W_{\ddot{u}}, G_z(3.0, 0.0) \quad (\text{A.7})$$

$$\text{RIL100} = \text{LF} \rightarrow \text{Maßnahme} = W_{\ddot{u}}, G_z(2.0, 0.0) \quad (\text{A.8})$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{soll abfahrt}} \geq 2015-06-16 14:35 \wedge \text{RIL100} = \text{MMF} \wedge \\ t_{F_{\text{Zubringer}}, \text{donehalte}} \geq 10 \rightarrow \text{Maßnahme} = G_a, G_z(11.0, 3.0) \quad (\text{A.9})$$

$$t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{donehalte}} \leq 8 \rightarrow \text{Maßnahme} = A(52.0, 6.0) \quad (\text{A.10})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{restfahrtdauer}} \geq 105 \wedge \\ t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} \geq 12 \wedge t_{A, \text{soll umstieg}} \leq 17 \rightarrow \text{Maßnahme} = A(9.0, 1.0) \quad (\text{A.11})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge \text{Linie}_{\text{Zubringer}} = 28.2 \rightarrow \text{Maßnahme} = A(9.0, 2.0) \quad (\text{A.12})$$

$$\begin{aligned} t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft}} &\geq 2015-06-16 \ 11:09 \wedge \\ &B_{\text{Zubringer Ziel}} = \text{MH} \wedge \\ t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft}} &\leq 2015-06-16 \ 15:23 \rightarrow \text{Maßnahme} = A(7.0, 0.0) \end{aligned} \quad (\text{A.13})$$

$$\begin{aligned} t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft}} &\geq 2015-06-16 \ 09:23 \wedge \\ B_{\text{Zubringer Start}} &= \text{LL} \wedge t_{A, \text{soll umstieg}} \leq 13 \wedge \\ t_{F_{\text{Zubringer}}, \text{donehalte}} &\leq 5 \rightarrow \text{Maßnahme} = A(9.0, 1.0) \end{aligned} \quad (\text{A.14})$$

$$t_{F_{\text{Abbringer}}, \text{donefahrtdauer}} \leq 5 \wedge t_{F_{\text{Zubringer}}, \text{donefahrtdauer}} \leq 50 \rightarrow \text{Maßnahme} = G_a(32.0, 7.0) \quad (\text{A.15})$$

$$\begin{aligned} t_{H_{B,F_{\text{Abbringer}}}, \text{soll abfahrt}} &\geq 2015-06-16 \ 11:03 \wedge \\ t_{F_{\text{Zubringer}}, \text{donefahrtdauer}} &\leq 337 \wedge t_{F_{\text{Zubringer}}, \text{donehalte}} \geq 16 \rightarrow \text{Maßnahme} = G_a(10.0, 2.0) \end{aligned} \quad (\text{A.16})$$

$$\begin{aligned} t_{H_{B,F_{\text{Abbringer}}}, \text{soll abfahrt}} &\geq 2015-06-16 \ 10:21 \wedge \\ t_{F_{\text{Zubringer}}, \text{donehalte}} &\leq 3 \wedge \\ t_{H_{B,F_{\text{Abbringer}}}, \text{soll abfahrt}} &\geq 2015-06-16 \ 14:40 \rightarrow \text{Maßnahme} = G_a(11.0, 0.0) \end{aligned} \quad (\text{A.17})$$

$$\begin{aligned} t_{H_{B,F_{\text{Abbringer}}}, \text{soll abfahrt}} &\geq 2015-06-16 \ 10:21 \wedge \\ t_{H_{B,F_{\text{Abbringer}}}, \text{soll abfahrt}} &\geq 2015-06-16 \ 20:29 \wedge \\ t_{F_{\text{Zubringer}}, \text{restfahrtdauer}} &\geq 51 \wedge t_{A, \text{soll umstieg}} \leq 68 \rightarrow \text{Maßnahme} = G_a(9.0, 0.0) \end{aligned} \quad (\text{A.18})$$

$$\begin{aligned} t_{A, \text{Regelwartezeit}} &\geq 3 \wedge t_{H_{B,F_{\text{Abbringer}}}, \text{abfahrt verspaetung}} \leq 3 \wedge \\ t_{A, \text{ist umstieg}} &\geq 6 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} \leq 11 \rightarrow \text{Maßnahme} = W_r(102.0, 1.0) \end{aligned} \quad (\text{A.19})$$

$$\begin{aligned} t_{A, \text{Regelwartezeit}} &\geq 3 \wedge t_{H_{B,F_{\text{Abbringer}}}, \text{abfahrt verspaetung}} \leq 3 \wedge \\ t_{A, \text{ist umstieg}} &\geq 5 \rightarrow \text{Maßnahme} = W_r(57.0, 23.0) \end{aligned} \quad (\text{A.20})$$

$$B_{\text{Abbringer Start}} = \text{NJS} \rightarrow \text{Maßnahme} = W_r(3.0, 1.0) \quad (\text{A.21})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge t_{A, \text{soll umstieg}} \geq 24 \rightarrow \text{Maßnahme} = G_z(47.0, 4.0) \quad (\text{A.22})$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft}} \geq 2015-06-16 \ 09:50 \rightarrow \text{Maßnahme} = G_z(141.0, 41.0) \quad (\text{A.23})$$

$$\begin{aligned} t_{F_{\text{Abbringer}}, \text{donehalte}} &\leq 0 \wedge t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 10 \wedge \\ t_{F_{\text{Zubringer}}, \text{donehalte}} &\geq 11 \rightarrow \text{Maßnahme} = G_z(22.0, 2.0) \end{aligned} \quad (\text{A.24})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{donefahrtdauer}} \geq 345 \rightarrow \text{Maßnahme} = G_z(11.0, 1.0) \quad (\text{A.25})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{donefahrtdauer}} \leq 198 \wedge t_{F_{\text{Zubringer}}, \text{donehalte}} \geq 4 \wedge t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 18 \rightarrow \text{Maßnahme} = G_z(30.0, 7.0) \quad (\text{A.26})$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{wartezeit}} \geq 4 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} \geq 27 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{soll haltezeit}} \leq 2 \rightarrow \text{Maßnahme} = G_z(12.0, 4.0) \quad (\text{A.27})$$

$$\text{Richtung}_{\text{Zubringer}} = \text{Frankfurt (Main) Flugh Fernbf - Dresden Hbf} \rightarrow \text{Maßnahme} = G_z(2.0, 0.0) \quad (\text{A.28})$$

$$t_{A, \text{ist umstieg}} \leq 4 \wedge t_{\text{mindestübergangszeit}} \geq 5 \rightarrow \text{Maßnahme} = W_w(288.0, 29.0) \quad (\text{A.29})$$

$$t_{A, \text{ist umstieg}} \leq 3 \rightarrow \text{Maßnahme} = W_w(55.0, 14.0) \quad (\text{A.30})$$

$$t_{A, \text{ist umstieg}} \leq 5 \wedge t_{\text{mindestübergangszeit}} \geq 6 \rightarrow \text{Maßnahme} = W_w(15.0, 5.0) \quad (\text{A.31})$$

$$\rightarrow \text{Maßnahme} = W_{\dot{u}}(539.0, 83.0) \quad (\text{A.32})$$

A.1.2 Featuremenge \mathbb{F}_2

$$B_{\text{Zubringer Ziel}} = \text{MFSN} \rightarrow \text{Maßnahme} = W_{\dot{u}}, G_a(2.0, 0.0) \quad (\text{A.1})$$

$$\text{Linie}_{\text{Abbringer}} = 78.1 \rightarrow \text{Maßnahme} = W_{\dot{u}}, G_a(3.0, 1.0) \quad (\text{A.2})$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} \geq 15 \wedge \text{Linie}_{\text{Abbringer}} = 28.2 \wedge t_{A, \text{soll umstieg}} \leq 18 \rightarrow \text{Maßnahme} = W_w, G_z(4.0, 0.0) \quad (\text{A.3})$$

$$t_{A,F_{\text{Zubringer}}, \text{eingangsverspätung}} \geq 25 \wedge t_{A, \text{soll umstieg}} \geq 45 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ist ankunft uhrzeit}} \leq 11:34 \wedge t_{F_{\text{Zubringer}}, \text{restfahrtdauer}} \geq 53 \rightarrow \text{Maßnahme} = W_{\dot{u}}, G_z(8.0, 2.0) \quad (\text{A.4})$$

$$t_{A,F_{\text{Zubringer}}, \text{eingangsverspätung}} \geq 13 \wedge B_{\text{Zubringer Start}} = \text{XDKH} \rightarrow \text{Maßnahme} = W_{\dot{u}}, G_z(3.0, 0.0) \quad (\text{A.5})$$

$$\text{RIL100} = \text{BWP} \rightarrow \text{Maßnahme} = W_{\dot{u}}, G_z(3.0, 1.0) \quad (\text{A.6})$$

$$t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{donefahrtdauer}} \leq 70 \rightarrow \text{Maßnahme} = A(37.0, 1.0) \quad (\text{A.7})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 3 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{soll haltezeit}} \geq 5 \rightarrow \text{Maßnahme} = A(10.0, 2.0) \quad (\text{A.8})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{restfahrtdauer}} \geq 78 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} \geq 13 \wedge t_{A,F_{\text{Zubringer}}, \text{eingangsverspätung}} \leq 13 \rightarrow \text{Maßnahme} = A(8.0, 0.0) \quad (\text{A.9})$$

$$t_{H_B, F_{Zubringer}, \text{ soll ankunft}} \geq 2015-06-16 \text{ 09:23} \wedge$$

$$t_{F_{Zubringer}, \text{ restfahrdauer}} \geq 272 \wedge t_{F_{Zubringer}, \text{ restfahrdauer}} \leq 284 \rightarrow \text{Maßnahme} = A(12.0, 1.0) \quad (\text{A.10})$$

$$t_{F_{Abbringer}, \text{ donehalte}} \leq 0 \wedge t_{\text{bahnsteiggleiche übergangszeit}} \geq 5 \wedge$$

$$t_{H_B, F_{Zubringer}, \text{ soll ankunft uhrzeit}} \leq 14:59 \wedge$$

$$t_{H_B, F_{Zubringer}, \text{ soll ankunft}} \geq 2015-06-15 \text{ 12:48} \wedge$$

$$t_{F_{Zubringer}, \text{ resthalte}} \leq 4 \rightarrow \text{Maßnahme} = A(11.0, 2.0) \quad (\text{A.11})$$

$$t_{F_{Abbringer}, \text{ donefahrdauer}} \leq 5 \wedge t_{F_{Zubringer}, \text{ donefahrdauer}} \leq 50 \rightarrow \text{Maßnahme} = G_a(32.0, 7.0) \quad (\text{A.12})$$

$$t_{H_B, F_{Abbringer}, \text{ soll abfahrt}} \geq 2015-06-16 \text{ 10:31} \wedge$$

$$t_{F_{Zubringer}, \text{ donehalte}} \leq 3 \wedge$$

$$t_{H_B, F_{Abbringer}, \text{ soll abfahrt}} \geq 2015-06-16 \text{ 14:40} \rightarrow \text{Maßnahme} = G_a(11.0, 0.0) \quad (\text{A.13})$$

$$t_{H_B, F_{Abbringer}, \text{ soll abfahrt}} \geq 2015-06-16 \text{ 10:21} \wedge$$

$$t_{H_B, F_{Abbringer}, \text{ soll abfahrt}} \leq 2015-06-16 \text{ 16:34} \wedge$$

$$t_{F_{Zubringer}, \text{ donehalte}} \geq 16 \rightarrow \text{Maßnahme} = G_a(12.0, 2.0) \quad (\text{A.14})$$

$$t_{F_{Abbringer}, \text{ donefahrdauer}} \leq 5 \wedge$$

$$t_{H_B, F_{Abbringer}, \text{ soll abfahrt uhrzeit}} \geq 18:35 \wedge$$

$$t_{H_B, F_{Abbringer}, \text{ soll abfahrt}} \geq 2015-06-16 \text{ 10:21} \wedge$$

$$t_{F_{Zubringer}, \text{ donefahrdauer}} \leq 225 \wedge t_{\text{mindestübergangszeit}} \leq 6 \rightarrow \text{Maßnahme} = G_a(24.0, 9.0) \quad (\text{A.15})$$

$$t_{H_B, F_{Abbringer}, \text{ soll abfahrt}} \geq 2015-06-16 \text{ 10:21} \wedge$$

$$t_{F_{Zubringer}, \text{ donehalte}} \leq 9 \wedge t_{F_{Zubringer}, \text{ donehalte}} \geq 6 \wedge$$

$$t_{H_B, F_{Zubringer}, \text{ soll haltezeit}} \leq 5 \wedge t_{\text{mindestübergangszeit}} \geq 5 \rightarrow \text{Maßnahme} = G_a(14.0, 3.0) \quad (\text{A.16})$$

$$\text{Richtung}_{\text{Abbringer}} = \text{Hamburg-Altona - Dresden Hbf} \rightarrow \text{Maßnahme} = G_a(2.0, 0.0) \quad (\text{A.17})$$

$$t_{H_B, F_{Zubringer}, \text{ soll ankunft}} \geq 2015-06-16 \text{ 19:40} \wedge$$

$$t_{H_B, F_{Zubringer}, \text{ soll ankunft}} \leq 2015-06-16 \text{ 20:23} \rightarrow \text{Maßnahme} = G_a(3.0, 0.0) \quad (\text{A.18})$$

$$\text{Richtung}_{\text{Abbringer}} = \text{Nürnberg Hbf - Dortmund Hbf} \rightarrow \text{Maßnahme} = G_a(2.0, 0.0) \quad (\text{A.19})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{A, \text{ soll umstieg}} \geq 13 \wedge$$

$$t_{H_B, F_{Zubringer}, \text{ ankunft verspätung}} \leq 11 \wedge$$

$$t_{F_{Abbringer}, \text{ donefahrdauer}} \leq 111 \rightarrow \text{Maßnahme} = W_r(45.0, 12.0) \quad (\text{A.20})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_B, F_{Zubringer}, \text{ ankunft verspätung}} \leq 11 \wedge$$

$$t_{A, \text{ soll umstieg}} \geq 14 \wedge t_{F_{Zubringer}, \text{ donehalte}} \geq 5 \wedge$$

$$t_{H_B, F_{Zubringer}, \text{ ist haltezeit}} \leq 3 \rightarrow \text{Maßnahme} = W_r(26.0, 5.0) \quad (\text{A.21})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_B, F_{Zubringer}, \text{ ankunft verspätung}} \leq 4 \wedge$$

$$t_{F_{Zubringer}, \text{ donefahrdauer}} \geq 151 \wedge t_{F_{Zubringer}, \text{ donehalte}} \leq 7 \wedge$$

$$t_{F_{Zubringer}, \text{ restfahrdauer}} \leq 263 \wedge t_{A, \text{ soll umstieg}} \geq 6 \rightarrow \text{Maßnahme} = W_r(17.0, 2.0) \quad (\text{A.22})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{A, \text{ soll umstieg}} \geq 18 \wedge \text{RIL100} = \text{HH} \wedge$$

$$t_{F_{Abbringer}, \text{ donefahrdauer}} \geq 216 \rightarrow \text{Maßnahme} = W_r(10.0, 1.0) \quad (\text{A.23})$$

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{A, \text{soll umstieg}} \geq 65 \wedge t_{A, \text{soll umstieg}} \leq 67 \rightarrow \text{Maßnahme} = W_r(10.0, 2.0) \quad (\text{A.24})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \rightarrow \text{Maßnahme} = G_z(183.0, 45.0) \quad (\text{A.25})$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{soll abfahrt}} \geq 2015-06-16 \ 10:00 \rightarrow \text{Maßnahme} = G_z(85.0, 21.0) \quad (\text{A.26})$$

$$t_{A, \text{soll umstieg}} \leq 10 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} \geq 7 \rightarrow \text{Maßnahme} = W_w(213.0, 85.0) \quad (\text{A.27})$$

$$t_{A, \text{soll umstieg}} \leq 33 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} \geq 29 \rightarrow \text{Maßnahme} = W_w(28.0, 3.0) \quad (\text{A.28})$$

$$t_{A, \text{soll umstieg}} \leq 12 \wedge t_{F_{\text{Abbringer}}, \text{resthalte}} \geq 12 \wedge$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} \geq 4 \wedge \quad (\text{A.29})$$

$$t_{F_{\text{Abbringer}}, \text{restfahrdauer}} \leq 186 \rightarrow \text{Maßnahme} = W_w(29.0, 5.0)$$

$$t_{A, \text{soll umstieg}} \leq 7 \wedge t_{F_{\text{Abbringer}}, \text{restfahrdauer}} \leq 43 \rightarrow \text{Maßnahme} = W_w(58.0, 23.0) \quad (\text{A.30})$$

$$t_{A, \text{soll umstieg}} \leq 16 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} \geq 14 \rightarrow \text{Maßnahme} = W_w(36.0, 13.0) \quad (\text{A.31})$$

$$\rightarrow \text{Maßnahme} = W_{ii}(571.0, 213.0) \quad (\text{A.32})$$

A.1.3 Featuremenge \mathbb{F}_3

$$\text{Richtung}_{\text{Abbringer}} = \text{Essen Hbf} - \text{Würzburg Hbf} \rightarrow \text{Maßnahme} = A, G_a(3.0, 1.0) \quad (\text{A.1})$$

$$B_{\text{Zubringer Ziel}} = \text{MFSN} \rightarrow \text{Maßnahme} = W_{ii}, G_a(2.0, 0.0) \quad (\text{A.2})$$

$$\text{Linie}_{\text{Abbringer}} = 78.1 \rightarrow \text{Maßnahme} = W_{ii}, G_a(3.0, 1.0) \quad (\text{A.3})$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} \geq 15 \wedge \text{Linie}_{\text{Abbringer}} = 28.2 \wedge \quad (\text{A.4})$$

$$t_{A, \text{soll umstieg}} \leq 18 \rightarrow \text{Maßnahme} = W_w, G_z(4.0, 0.0)$$

$$B_{\text{Zubringer Start}} = \text{XDKH} \rightarrow \text{Maßnahme} = W_{ii}, G_z(3.0, 0.0) \quad (\text{A.5})$$

$$\text{RIL100} = \text{BWP} \rightarrow \text{Maßnahme} = W_{ii}, G_z(3.0, 1.0) \quad (\text{A.6})$$

$$t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{donefahrdauer}} \leq 98 \rightarrow \text{Maßnahme} = A(41.0, 2.0) \quad (\text{A.7})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 3 \wedge$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{soll haltezeit}} \geq 5 \rightarrow \text{Maßnahme} = A(10.0, 2.0) \quad (\text{A.8})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{restfahrdauer}} \geq 105 \wedge$$

$$t_{A,F_{\text{Zubringer}}, \text{eingangsverspaetung}} \geq 7 \wedge \quad (\text{A.9})$$

$$t_{A,F_{\text{Zubringer}}, \text{eingangsverspaetung}} \leq 13 \wedge t_{\text{mindestuebergangszeit}} \leq 7 \wedge$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft uhrzeit}} \leq 15:05 \rightarrow \text{Maßnahme} = A(15.0, 1.0)$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge \text{RIL100} = \text{FF} \wedge t_{A, \text{soll umstieg}} \leq 14 \wedge \quad (\text{A.10})$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft uhrzeit}} \geq 19:08 \rightarrow \text{Maßnahme} = A(7.0, 0.0)$$

$$B_{\text{Zubringer Start}} = \text{FG} \rightarrow \text{Maßnahme} = A(2.0, 0.0) \quad (\text{A.11})$$

$$d_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft wochentag}} \geq 2 \wedge B_{\text{Zubringer Ziel}} = \text{MH} \wedge$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft uhrzeit}} \geq 10:41 \wedge \quad (\text{A.12})$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft uhrzeit}} \leq 15:23 \rightarrow \text{Maßnahme} = A(7.0, 0.0)$$

$$t_{F_{\text{Abbringer}}, \text{donefahrdauer}} \leq 5 \wedge t_{F_{\text{Zubringer}}, \text{donefahrdauer}} \leq 50 \rightarrow \text{Maßnahme} = G_a(32.0, 7.0) \quad (\text{A.13})$$

$$t_{F_{\text{Abbringer}}, \text{ donefahrtdauer}} \leq 5 \wedge t_{A, \text{ soll umstieg}} \leq 20 \wedge$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{ soll abfahrt uhrzeit}} \geq 18:35 \wedge t_{A, \text{ soll umstieg}} \leq 13 \rightarrow \text{Maßnahme} = G_a(26.0, 10.0) \quad (\text{A.14})$$

$$d_{H_{B,F_{\text{Abbringer}}}, \text{ soll abfahrt wochentag}} \geq 2 \wedge$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{ soll abfahrt uhrzeit}} \geq 10:21 \wedge$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{ soll abfahrt uhrzeit}} \leq 16:29 \wedge$$

$$t_{F_{\text{Zubringer}}, \text{ donehalte}} \geq 16 \rightarrow \text{Maßnahme} = G_a(12.0, 2.0) \quad (\text{A.15})$$

$$d_{H_{B,F_{\text{Abbringer}}}, \text{ soll abfahrt wochentag}} \geq 2 \wedge$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{ soll abfahrt uhrzeit}} \geq 10:21 \wedge$$

$$t_{F_{\text{Zubringer}}, \text{ donehalte}} \leq 3 \rightarrow \text{Maßnahme} = G_a(16.0, 4.0) \quad (\text{A.16})$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{ soll abfahrt uhrzeit}} \geq 21:10 \wedge$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{ soll abfahrt uhrzeit}} \leq 21:17 \rightarrow \text{Maßnahme} = G_a(8.0, 2.0) \quad (\text{A.17})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \leq 10 \wedge$$

$$t_{A, \text{ soll umstieg}} \geq 9 \wedge t_{F_{\text{Abbringer}}, \text{ restfahrtdauer}} \geq 279 \rightarrow \text{Maßnahme} = W_r(36.0, 8.0) \quad (\text{A.18})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{A, \text{ soll umstieg}} \geq 13 \wedge$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \leq 11 \wedge$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \leq 3 \rightarrow \text{Maßnahme} = W_r(25.0, 7.0) \quad (\text{A.19})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{A, \text{ soll umstieg}} \geq 27 \wedge$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \leq 25 \wedge$$

$$t_{F_{\text{Zubringer}}, \text{ restfahrtdauer}} \leq 132 \wedge t_{F_{\text{Zubringer}}, \text{ restfahrtdauer}} \geq 93 \wedge$$

$$d_{H_{B,F_{\text{Zubringer}}}, \text{ soll ankunft wochentag}} \leq 1 \rightarrow \text{Maßnahme} = W_r(16.0, 1.0) \quad (\text{A.20})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \leq 23 \wedge$$

$$t_{A, \text{ soll umstieg}} \geq 14 \wedge t_{A,F_{\text{Zubringer}}, \text{ eingangsverspätung}} \geq 7 \wedge$$

$$t_{A,F_{\text{Zubringer}}, \text{ eingangsverspätung}} \leq 10 \wedge$$

$$t_{F_{\text{Zubringer}}, \text{ restfahrtdauer}} \geq 151 \rightarrow \text{Maßnahme} = W_r(15.0, 3.0) \quad (\text{A.21})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_{B,F_{\text{Abbringer}}}, \text{ soll haltezeit}} \geq 9 \wedge$$

$$t_{A, \text{ soll umstieg}} \geq 17 \wedge t_{F_{\text{Abbringer}}, \text{ resthalte}} \geq 12 \rightarrow \text{Maßnahme} = W_r(7.0, 0.0) \quad (\text{A.22})$$

$$t_{F_{\text{Abbringer}}, \text{ donehalte}} \leq 0 \rightarrow \text{Maßnahme} = G_z(184.0, 38.0) \quad (\text{A.23})$$

$$d_{H_{B,F_{\text{Zubringer}}}, \text{ soll ankunft wochentag}} \geq 2 \wedge$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{ soll abfahrt uhrzeit}} \geq 10:00 \rightarrow \text{Maßnahme} = G_z(94.0, 31.0) \quad (\text{A.24})$$

$$t_{A, \text{ soll umstieg}} \leq 7 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \geq 8 \wedge$$

$$t_{F_{\text{Abbringer}}, \text{ donefahrtdauer}} \leq 101 \rightarrow \text{Maßnahme} = W_w(60.0, 7.0) \quad (\text{A.25})$$

$$t_{A, \text{ soll umstieg}} \leq 19 \wedge t_{H_{B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \geq 10 \wedge$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \geq 17 \rightarrow \text{Maßnahme} = W_w(60.0, 16.0) \quad (\text{A.26})$$

$$t_{A, \text{ soll umstieg}} \leq 7 \wedge t_{H_{B,F} \text{Zubringer}, \text{ ist haltezeit}} \leq 3 \wedge \quad (\text{A.27})$$

$$t_{F \text{Zubringer}, \text{ restfahrdauer}} \leq 78 \rightarrow \text{Maßnahme} = W_w(51.0, 17.0)$$

$$t_{A, \text{ soll umstieg}} \leq 12 \wedge t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} \geq 8 \wedge \quad (\text{A.28})$$

$$\text{RIL100} = \text{HH} \rightarrow \text{Maßnahme} = W_w(31.0, 9.0)$$

$$t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} \geq 15 \wedge t_{A, \text{ soll umstieg}} \leq 35 \wedge \quad (\text{A.29})$$

$$t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} \geq 29 \rightarrow \text{Maßnahme} = W_w(21.0, 5.0)$$

$$t_{A, \text{ soll umstieg}} \leq 6 \wedge t_{F \text{Abbringer}, \text{ restfahrdauer}} \geq 147 \wedge \quad (\text{A.30})$$

$$t_{F \text{Zubringer}, \text{ donehalte}} \leq 11 \rightarrow \text{Maßnahme} = W_w(29.0, 9.0)$$

$$t_{A, \text{ soll umstieg}} \leq 12 \wedge t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} \geq 4 \wedge \quad (\text{A.31})$$

$$t_{F \text{Abbringer}, \text{ resthalte}} \geq 16 \rightarrow \text{Maßnahme} = W_w(29.0, 10.0)$$

$$\rightarrow \text{Maßnahme} = W_{ii}(660.0, 258.0) \quad (\text{A.32})$$

A.1.4 Featuremenge \mathbb{F}_4

$$\text{Linie}_{\text{Abbringer}} = 78.1 \rightarrow \text{Maßnahme} = W_{ii}, G_a(3.0, 1.0) \quad (\text{A.1})$$

$$B_{\text{Zubringer Ziel}} = \text{MFSN} \rightarrow \text{Maßnahme} = W_{ii}, G_a(2.0, 0.0) \quad (\text{A.2})$$

$$t_{A, F \text{Zubringer}, \text{ eingangverspaetung}} \geq 13 \wedge B_{\text{Zubringer Start}} = \text{XDKH} \rightarrow \text{Maßnahme} = W_{ii}, G_z(3.0, 0.0) \quad (\text{A.3})$$

$$\text{RIL100} = \text{MMF} \wedge t_{H_{B,F} \text{Abbringer}, \text{ soll abfahrt}} \geq 2015-06-16 \text{ 08:34} \wedge \quad (\text{A.4})$$

$$t_{F \text{Zubringer}, \text{ donehalte}} \geq 7 \wedge t_{F \text{Zubringer}, \text{ resthalte}} \leq 5 \rightarrow \text{Maßnahme} = G_a, G_z(13.0, 3.0)$$

$$t_{F \text{Abbringer}, \text{ resthalte}} \leq 0 \wedge t_{F \text{Zubringer}, \text{ donefahrdauer}} \leq 98 \rightarrow \text{Maßnahme} = A(41.0, 2.0) \quad (\text{A.5})$$

$$t_{F \text{Abbringer}, \text{ donehalte}} \leq 0 \wedge t_{F \text{Zubringer}, \text{ restfahrdauer}} \geq 105 \wedge$$

$$t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} \geq 11 \wedge \quad (\text{A.6})$$

$$t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} \leq 20 \rightarrow \text{Maßnahme} = A(16.0, 3.0)$$

$$t_{H_{B,F} \text{Zubringer}, \text{ soll ankunft}} \geq 2015-06-16 \text{ 09:23} \wedge \quad (\text{A.7})$$

$$t_{F \text{Zubringer}, \text{ restfahrdauer}} \geq 272 \wedge t_{F \text{Zubringer}, \text{ restfahrdauer}} \leq 284 \rightarrow \text{Maßnahme} = A(12.0, 1.0)$$

$$t_{F \text{Abbringer}, \text{ donefahrdauer}} \leq 3 \wedge B_{\text{Zubringer Start}} = \text{LL} \wedge \quad (\text{A.8})$$

$$t_{F \text{Abbringer}, \text{ restfahrdauer}} \leq 90 \rightarrow \text{Maßnahme} = A(9.0, 1.0)$$

$$t_{H_{B,F} \text{Zubringer}, \text{ soll ankunft}} \geq 2015-06-16 \text{ 09:23} \wedge \quad (\text{A.9})$$

$$B_{\text{Abbringer Start}} = \text{BHF} \wedge t_{F \text{Zubringer}, \text{ donehalte}} \leq 5 \rightarrow \text{Maßnahme} = A(4.0, 0.0)$$

$$B_{\text{Zubringer Start}} = \text{FG} \rightarrow \text{Maßnahme} = A(2.0, 0.0) \quad (\text{A.10})$$

$$t_{F \text{Abbringer}, \text{ donefahrdauer}} \leq 5 \wedge t_{F \text{Zubringer}, \text{ donefahrdauer}} \leq 49 \rightarrow \text{Maßnahme} = G_a(31.0, 7.0) \quad (\text{A.11})$$

$$t_{H_{B,F} \text{Abbringer}, \text{ soll abfahrt}} \geq 2015-06-16 \text{ 10:27} \wedge \quad (\text{A.12})$$

$$t_{F \text{Zubringer}, \text{ donehalte}} \leq 3 \rightarrow \text{Maßnahme} = G_a(17.0, 4.0)$$

$$t_{F \text{Abbringer}, \text{ donefahrdauer}} \leq 13 \wedge B_{\text{Abbringer Ziel}} = \text{FF} \wedge \quad (\text{A.13})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \rightarrow \text{Maßnahme} = G_a(8.0, 2.0)$$

$$\begin{aligned}
t_{H_B,F_{\text{Abbringer}}}, \text{ soll abfahrt} &\geq 2015-06-16 \ 10:21 \wedge \\
t_{H_B,F_{\text{Abbringer}}}, \text{ soll abfahrt} &\leq 2015-06-16 \ 16:34 \wedge
\end{aligned} \tag{A.14}$$

$$\begin{aligned}
&t_{F_{\text{Zubringer}}}, \text{ donehalte} \geq 18 \rightarrow \text{Maßnahme} = G_a(8.0, 2.0) \\
t_{H_B,F_{\text{Abbringer}}}, \text{ soll abfahrt} &\geq 2015-06-16 \ 10:21 \wedge \\
t_{H_B,F_{\text{Zubringer}}}, \text{ soll ankunft} &\geq 2015-06-16 \ 19:40 \wedge
\end{aligned} \tag{A.15}$$

$$\begin{aligned}
&t_{F_{\text{Zubringer}}}, \text{ restfahrdauer} \geq 52 \rightarrow \text{Maßnahme} = G_a(8.0, 2.0) \\
t_{H_B,F_{\text{Abbringer}}}, \text{ soll abfahrt} &\geq 2015-06-16 \ 10:21 \wedge \\
t_{H_B,F_{\text{Abbringer}}}, \text{ soll abfahrt} &\leq 2015-06-16 \ 16:34 \wedge
\end{aligned} \tag{A.16}$$

$$\begin{aligned}
&t_{F_{\text{Zubringer}}}, \text{ donefahrdauer} \geq 204 \wedge t_{F_{\text{Zubringer}}}, \text{ resthalte} \geq 4 \rightarrow \text{Maßnahme} = G_a(9.0, 1.0) \\
t_A, \text{ Regelwartezeit} \geq 3 \wedge t_{H_B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung} &\leq 11 \wedge \\
&t_A, \text{ soll umstieg} \geq 13 \wedge t_{F_{\text{Abbringer}}}, \text{ donefahrdauer} \leq 111 \rightarrow \text{Maßnahme} = W_r(45.0, 12.0)
\end{aligned} \tag{A.17}$$

$$\begin{aligned}
t_A, \text{ Regelwartezeit} \geq 3 \wedge t_{H_B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung} &\leq 11 \wedge \\
&t_A, \text{ soll umstieg} \geq 14 \rightarrow \text{Maßnahme} = W_r(79.0, 38.0)
\end{aligned} \tag{A.18}$$

$$\begin{aligned}
t_A, \text{ Regelwartezeit} \geq 3 \wedge t_{H_B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung} &\leq 5 \wedge \\
&t_A, \text{ soll umstieg} \geq 8 \wedge t_{F_{\text{Zubringer}}}, \text{ resthalte} \leq 1 \rightarrow \text{Maßnahme} = W_r(9.0, 1.0)
\end{aligned} \tag{A.19}$$

$$\begin{aligned}
&t_A, \text{ Regelwartezeit} \geq 3 \wedge \\
t_{H_B,F_{\text{Zubringer}}}, \text{ ist ankunft} \leq 2015-06-15 \ 19:30 \wedge t_A, \text{ soll umstieg} &\geq 26 \wedge
\end{aligned} \tag{A.20}$$

$$t_{H_B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung} \leq 23 \rightarrow \text{Maßnahme} = W_r(33.0, 16.0)$$

$$\begin{aligned}
t_A, \text{ Regelwartezeit} \geq 3 \wedge t_{H_B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung} &\leq 7 \wedge \\
t_A, \text{ soll umstieg} \geq 9 \wedge t_{F_{\text{Abbringer}}}, \text{ donefahrdauer} &\leq 121 \wedge
\end{aligned} \tag{A.21}$$

$$t_{F_{\text{Abbringer}}}, \text{ donehalte} \geq 3 \rightarrow \text{Maßnahme} = W_r(11.0, 1.0)$$

$$t_{F_{\text{Abbringer}}}, \text{ donehalte} \leq 0 \rightarrow \text{Maßnahme} = G_z(189.0, 43.0) \tag{A.22}$$

$$t_{H_B,F_{\text{Abbringer}}}, \text{ soll abfahrt} \geq 2015-06-16 \ 10:00 \rightarrow \text{Maßnahme} = G_z(81.0, 21.0) \tag{A.23}$$

$$\begin{aligned}
t_A, \text{ soll umstieg} \leq 10 \wedge t_{H_B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung} &\geq 8 \rightarrow \text{Maßnahme} = W_w(182.0, 68.0) \\
&
\end{aligned} \tag{A.24}$$

$$\begin{aligned}
t_A, \text{ soll umstieg} \leq 7 \wedge t_{F_{\text{Zubringer}}}, \text{ restfahrdauer} &\leq 36 \wedge \\
&t_A, \text{ soll umstieg} \leq 4 \rightarrow \text{Maßnahme} = W_w(43.0, 15.0)
\end{aligned} \tag{A.25}$$

$$t_A, \text{ soll umstieg} \leq 19 \wedge t_{H_B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung} \geq 14 \rightarrow \text{Maßnahme} = W_w(62.0, 22.0) \tag{A.26}$$

$$\begin{aligned}
&t_A, \text{ soll umstieg} \leq 9 \wedge t_{F_{\text{Abbringer}}}, \text{ resthalte} \geq 10 \wedge \\
&
\end{aligned} \tag{A.27}$$

$$\begin{aligned}
t_{H_B,F_{\text{Zubringer}}}, \text{ ankunft verspaetung} \geq 3 \wedge t_{F_{\text{Zubringer}}}, \text{ resthalte} &\leq 10 \rightarrow \text{Maßnahme} = W_w(44.0, 12.0) \\
&\rightarrow \text{Maßnahme} = W_{ii}(548.0, 210.0)
\end{aligned} \tag{A.28}$$

A.2 Klassenzusammenstellung \mathbb{K}_2

A.2.1 Featuremenge \mathbb{F}_1

$$\text{Linie}_{\text{Zubringer}} = 10.1 \wedge B_{\text{Abbringer Start}} = \text{MH} \rightarrow \text{Ma\ssnahme} = A, W_i(3.0, 1.0) \quad (\text{A.1})$$

$$t_{A, \text{soll umstieg}} \geq 139 \wedge t_{A, F_{\text{Abbringer}}, \text{eingangsversp\u00e4tung}} \leq 4 \rightarrow \text{Ma\ssnahme} = W_r, G(4.0, 1.0) \quad (\text{A.2})$$

$$t_{A, \text{ist umstieg}} \leq 0 \wedge t_{F_{\text{Abbringer}}, \text{donefahrdauer}} \geq 204 \wedge t_{H_{B, F_{\text{Abbringer}}}, \text{soll haltezeit}} \geq 7 \wedge \quad (\text{A.3})$$

$$t_{H_{B, F_{\text{Abbringer}}}, \text{abfahrt verspaetung}} \leq 1 \rightarrow \text{Ma\ssnahme} = A, W_w(7.0, 1.0)$$

$$\text{Richtung}_{\text{Zubringer}} = \text{Stuttgart Hbf} - \text{Hamburg-Altona} \wedge$$

$$t_{H_{B, F_{\text{Abbringer}}}, \text{soll haltezeit}} \geq 3 \wedge \quad (\text{A.4})$$

$$t_{H_{B, F_{\text{Abbringer}}}, \text{abfahrt verspaetung}} \leq 1 \rightarrow \text{Ma\ssnahme} = A, W_w(3.0, 0.0)$$

$$t_{A, \text{ist umstieg}} \leq 4 \wedge t_{A, F_{\text{Zubringer}}, \text{eingangsversp\u00e4tung}} \geq 24 \wedge \quad (\text{A.5})$$

$$t_{H_{B, F_{\text{Zubringer}}}, \text{ist haltezeit}} \geq 2 \wedge t_{F_{\text{Abbringer}}, \text{donefahrdauer}} \geq 53 \rightarrow \text{Ma\ssnahme} = W_w, G(13.0, 4.0)$$

$$t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{donefahrdauer}} \leq 70 \rightarrow \text{Ma\ssnahme} = A(37.0, 1.0) \quad (\text{A.6})$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 0 \wedge \quad (\text{A.7})$$

$$t_{F_{\text{Zubringer}}, \text{resthalte}} \geq 7 \rightarrow \text{Ma\ssnahme} = A(5.0, 0.0)$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{restfahrdauer}} \geq 105 \wedge \quad (\text{A.8})$$

$$t_{A, \text{soll umstieg}} \leq 17 \wedge t_{A, F_{\text{Zubringer}}, \text{eingangsversp\u00e4tung}} \geq 5 \rightarrow \text{Ma\ssnahme} = A(22.0, 7.0)$$

$$t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{donehalte}} \leq 10 \wedge t_{H_{B, F_{\text{Zubringer}}}, \text{soll ankunft uhrzeit}} \leq 15:15 \wedge \quad (\text{A.9})$$

$$t_{H_{B, F_{\text{Zubringer}}}, \text{soll ankunft}} \geq 2015-06-15 10:32 \rightarrow \text{Ma\ssnahme} = A(8.0, 0.0)$$

$$t_{H_{B, F_{\text{Zubringer}}}, \text{soll ankunft}} \geq 2015-06-16 10:41 \wedge \quad (\text{A.10})$$

$$t_{F_{\text{Zubringer}}, \text{restfahrdauer}} \geq 272 \wedge t_{F_{\text{Abbringer}}, \text{donehalte}} \geq 4 \rightarrow \text{Ma\ssnahme} = A(6.0, 0.0)$$

$$t_{H_{B, F_{\text{Zubringer}}}, \text{soll ankunft}} \geq 2015-06-16 19:18 \wedge \quad (\text{A.11})$$

$$t_{H_{B, F_{\text{Zubringer}}}, \text{soll haltezeit}} \geq 5 \wedge t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \rightarrow \text{Ma\ssnahme} = A(8.0, 1.0)$$

$$B_{\text{Zubringer Start}} = \text{FG} \rightarrow \text{Ma\ssnahme} = A(2.0, 0.0) \quad (\text{A.12})$$

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{H_{B, F_{\text{Abbringer}}}, \text{abfahrt verspaetung}} \leq 3 \wedge \quad (\text{A.13})$$

$$t_{A, \text{ist umstieg}} \geq 6 \rightarrow \text{Ma\ssnahme} = W_r(144.0, 16.0)$$

$$B_{\text{Abbringer Start}} = \text{NJS} \rightarrow \text{Ma\ssnahme} = W_r(3.0, 1.0) \quad (\text{A.14})$$

$$\text{Richtung}_{\text{Zubringer}} = \text{Karlsruhe Hbf} - \text{Kiel Hbf} \rightarrow \text{Ma\ssnahme} = W_r(2.0, 0.0) \quad (\text{A.15})$$

$$t_{A, \text{ist umstieg}} \leq 4 \wedge t_{F_{\text{Abbringer}}, \text{donehalte}} \geq 2 \wedge \quad (\text{A.16})$$

$$t_{\text{mindest\u00fcbergangszeit}} \geq 5 \wedge t_{H_{B, F_{\text{Zubringer}}}, \text{wartezeit}} \leq 6 \rightarrow \text{Ma\ssnahme} = W_w(173.0, 6.0)$$

$$t_{A, \text{ist umstieg}} \leq 4 \wedge t_{H_{B, F_{\text{Abbringer}}}, \text{ist haltezeit}} \leq 6 \wedge \quad (\text{A.17})$$

$$t_{A, \text{ist umstieg}} \leq 2 \wedge t_{F_{\text{Zubringer}}, \text{donehalte}} \geq 8 \rightarrow \text{Ma\ssnahme} = W_w(66.0, 3.0)$$

$$t_{A, \text{ ist umstieg}} \leq 5 \wedge t_{H_{B,F} \text{ Abbringer}, \text{ abfahrt verspaetung}} \leq 1 \wedge$$

$$t_{H_{B,F} \text{ Abbringer}, \text{ soll haltezeit}} \leq 12 \rightarrow \text{Ma\ssnahme} = W_w(68.0, 15.0) \quad (\text{A.18})$$

$$t_{A, \text{ ist umstieg}} \leq 6 \wedge t_{\text{mindest\u00fcbergangszeit}} \geq 7 \wedge$$

$$t_{H_{B,F} \text{ Abbringer}, \text{ ist haltezeit}} \leq 8 \wedge t_{F_{Zubringer}, \text{ donefahrtdauer}} \geq 117 \rightarrow \text{Ma\ssnahme} = W_w(14.0, 1.0) \quad (\text{A.19})$$

$$t_{A, \text{ ist umstieg}} \leq 4 \wedge t_{H_{B,F} \text{ Abbringer}, \text{ soll haltezeit}} \geq 8 \rightarrow \text{Ma\ssnahme} = W_w(22.0, 9.0) \quad (\text{A.20})$$

$$t_{F_{\text{Abbringer}}, \text{ donehalte}} \leq 0 \rightarrow \text{Ma\ssnahme} = G(233.0, 8.0) \quad (\text{A.21})$$

$$t_{H_{B,F} \text{ Abbringer}, \text{ soll abfahrt}} \geq 2015-06-16 \ 10:00 \rightarrow \text{Ma\ssnahme} = G(120.0, 6.0) \quad (\text{A.22})$$

$$t_{A, \text{ ist umstieg}} \leq 2 \wedge t_{H_{B,F} \text{ Zubringer}, \text{ soll haltezeit}} \leq 13 \rightarrow \text{Ma\ssnahme} = G(5.0, 0.0) \quad (\text{A.23})$$

$$\rightarrow \text{Ma\ssnahme} = W_{\ddot{u}}(544.0, 93.0) \quad (\text{A.24})$$

A.2.2 Featuremenge \mathbb{F}_2

$$\text{Linie}_{\text{Zubringer}} = 10.1 \wedge B_{\text{Abbringer Start}} = \text{MH} \rightarrow \text{Ma\ssnahme} = A, W_{\ddot{u}}(3.0, 1.0) \quad (\text{A.1})$$

$$\text{Richtung}_{\text{Abbringer}} = \text{Essen Hbf} - \text{W\u00fcrzburg Hbf} \rightarrow \text{Ma\ssnahme} = A, G(3.0, 1.0) \quad (\text{A.2})$$

$$t_{A, \text{ soll umstieg}} \geq 139 \wedge t_{F_{\text{Abbringer}}, \text{ restfahrtdauer}} \geq 146 \rightarrow \text{Ma\ssnahme} = W_r, G(2.0, 0.0) \quad (\text{A.3})$$

$$t_{H_{B,F} \text{ Zubringer}, \text{ ankunft verspaetung}} \geq 15 \wedge \text{Linie}_{\text{Abbringer}} = 28.2 \wedge$$

$$t_{A, \text{ soll umstieg}} \leq 18 \rightarrow \text{Ma\ssnahme} = W_w, G(4.0, 0.0) \quad (\text{A.4})$$

$$\text{RIL100} = \text{LF} \rightarrow \text{Ma\ssnahme} = W_{\ddot{u}}, G(2.0, 0.0) \quad (\text{A.5})$$

$$t_{F_{\text{Abbringer}}, \text{ resthalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{ donehalte}} \leq 8 \wedge$$

$$t_{F_{\text{Zubringer}}, \text{ donefahrtdauer}} \leq 98 \rightarrow \text{Ma\ssnahme} = A(40.0, 2.0) \quad (\text{A.6})$$

$$t_{F_{\text{Abbringer}}, \text{ donehalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{ restfahrtdauer}} \geq 105 \wedge$$

$$t_{A, F_{\text{Zubringer}}, \text{ eingangsversp\u00e4tung}} \geq 7 \wedge t_{A, \text{ soll umstieg}} \leq 12 \rightarrow \text{Ma\ssnahme} = A(14.0, 3.0) \quad (\text{A.7})$$

$$t_{F_{\text{Abbringer}}, \text{ donehalte}} \leq 0 \wedge t_{F_{\text{Abbringer}}, \text{ resthalte}} \leq 0 \wedge$$

$$t_{H_{B,F} \text{ Zubringer}, \text{ soll ankunft}} \geq 2015-06-16 \ 19:18 \rightarrow \text{Ma\ssnahme} = A(6.0, 1.0) \quad (\text{A.8})$$

$$t_{F_{\text{Abbringer}}, \text{ donehalte}} \leq 0 \wedge B_{\text{Zubringer Start}} = \text{LL} \wedge$$

$$t_{A, \text{ soll umstieg}} \leq 13 \rightarrow \text{Ma\ssnahme} = A(5.0, 0.0) \quad (\text{A.9})$$

$$t_{F_{\text{Abbringer}}, \text{ donefahrtdauer}} \leq 0 \wedge B_{\text{Zubringer Ziel}} = \text{BHF} \wedge$$

$$t_{F_{\text{Zubringer}}, \text{ donehalte}} \leq 10 \rightarrow \text{Ma\ssnahme} = A(8.0, 0.0) \quad (\text{A.10})$$

$$t_{H_{B,F} \text{ Zubringer}, \text{ soll ankunft}} \geq 2015-06-16 \ 09:23 \wedge$$

$$t_{F_{\text{Zubringer}}, \text{ restfahrtdauer}} \geq 272 \wedge t_{F_{\text{Abbringer}}, \text{ donehalte}} \geq 4 \rightarrow \text{Ma\ssnahme} = A(6.0, 0.0) \quad (\text{A.11})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_{B,F} \text{ Zubringer}, \text{ ankunft verspaetung}} \leq 11 \wedge$$

$$t_{A, \text{ soll umstieg}} \geq 13 \wedge t_{F_{\text{Zubringer}}, \text{ restfahrtdauer}} \leq 124 \wedge$$

$$t_{F_{\text{Abbringer}}, \text{ donehalte}} \leq 10 \wedge$$

$$t_{H_{B,F} \text{ Zubringer}, \text{ soll ankunft}} \leq 2015-06-15 \ 17:53 \rightarrow \text{Ma\ssnahme} = W_r(30.0, 5.0) \quad (\text{A.12})$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_{B,F} \text{ Zubringer, ankunft verspaetung}} \leq 10 \wedge \quad (\text{A.13})$$

$$t_{A, \text{ soll umstieg}} \geq 14 \rightarrow \text{Ma\ssnahme} = W_r(80.0, 37.0)$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_{B,F} \text{ Zubringer, ankunft verspaetung}} \leq 6 \wedge \quad (\text{A.14})$$

$$t_{A, \text{ soll umstieg}} \geq 9 \rightarrow \text{Ma\ssnahme} = W_r(38.0, 17.0)$$

$$t_{A, \text{ soll umstieg}} \leq 6 \wedge t_{H_{B,F} \text{ Zubringer, ist haltezeit}} \leq 7 \wedge \quad (\text{A.15})$$

$$t_{F \text{ Zubringer, restfahrdauer}} \leq 65 \wedge$$

$$t_{H_{B,F} \text{ Zubringer, soll ankunft uhrzeit}} \leq 17:20 \rightarrow \text{Ma\ssnahme} = W_w(60.0, 9.0)$$

$$t_{A, \text{ soll umstieg}} \leq 12 \wedge t_{H_{B,F} \text{ Zubringer, ankunft verspaetung}} \geq 9 \wedge$$

$$t_{F \text{ Abbringer, donehalte}} \geq 3 \wedge t_{F \text{ Abbringer, donefahrdauer}} \leq 136 \wedge \quad (\text{A.16})$$

$$t_{F \text{ Zubringer, resthalte}} \leq 6 \rightarrow \text{Ma\ssnahme} = W_w(46.0, 7.0)$$

$$t_{H_{B,F} \text{ Abbringer, soll abfahrt}} \leq 2015-06-15 \ 20:52 \wedge t_{A, \text{ soll umstieg}} \leq 8 \wedge$$

$$t_{H_{B,F} \text{ Abbringer, soll haltezeit}} \leq 12 \wedge \quad (\text{A.17})$$

$$t_{H_{B,F} \text{ Abbringer, soll haltezeit}} \geq 4 \wedge$$

$$t_{F \text{ Zubringer, restfahrdauer}} \geq 100 \rightarrow \text{Ma\ssnahme} = W_w(44.0, 15.0)$$

$$t_{H_{B,F} \text{ Abbringer, soll abfahrt}} \leq 2015-06-15 \ 21:41 \wedge$$

$$t_{A, \text{ soll umstieg}} \leq 19 \wedge t_{H_{B,F} \text{ Zubringer, ankunft verspaetung}} \geq 17 \wedge \quad (\text{A.18})$$

$$t_{H_{B,F} \text{ Zubringer, ankunft verspaetung}} \geq 23 \wedge t_{F \text{ Zubringer, resthalte}} \geq 3 \rightarrow \text{Ma\ssnahme} = W_w(17.0, 1.0)$$

$$t_{H_{B,F} \text{ Zubringer, ist ankunft uhrzeit}} \geq 10:14 \wedge t_{A, \text{ soll umstieg}} \leq 37 \wedge \quad (\text{A.19})$$

$$t_{H_{B,F} \text{ Zubringer, ankunft verspaetung}} \geq 33 \wedge t_{F \text{ Zubringer, resthalte}} \geq 2 \rightarrow \text{Ma\ssnahme} = W_w(17.0, 1.0)$$

$$t_{F \text{ Abbringer, donehalte}} \leq 0 \rightarrow \text{Ma\ssnahme} = G(240.0, 14.0) \quad (\text{A.20})$$

$$t_{H_{B,F} \text{ Abbringer, soll abfahrt}} \geq 2015-06-16 \ 10:00 \rightarrow \text{Ma\ssnahme} = G(117.0, 4.0) \quad (\text{A.21})$$

$$\rightarrow \text{Ma\ssnahme} = W_{\ddot{u}}(730.0, 336.0) \quad (\text{A.22})$$

A.2.3 Featuremenge \mathbb{F}_3

$$\text{Linie}_{\text{Zubringer}} = 10.1 \wedge B_{\text{Abbringer Start}} = \text{MH} \rightarrow \text{Ma\ssnahme} = A, W_{\ddot{u}}(3.0, 1.0) \quad (\text{A.1})$$

$$\text{Richtung}_{\text{Abbringer}} = \text{Essen Hbf} - \text{W\u00fcrzburg Hbf} \rightarrow \text{Ma\ssnahme} = A, G(3.0, 1.0) \quad (\text{A.2})$$

$$t_{F \text{ Abbringer, donefahrdauer}} \geq 204 \wedge$$

$$t_{H_{B,F} \text{ Abbringer, soll haltezeit}} \geq 7 \wedge \quad (\text{A.3})$$

$$t_{F \text{ Abbringer, restfahrdauer}} \leq 164 \wedge t_{F \text{ Zubringer, resthalte}} \geq 6 \wedge$$

$$t_{H_{B,F} \text{ Zubringer, ankunft verspaetung}} \geq 9 \rightarrow \text{Ma\ssnahme} = A, W_w(5.0, 0.0)$$

$$\text{RIL100} = \text{LF} \rightarrow \text{Ma\ssnahme} = W_{\ddot{u}}, G(2.0, 0.0) \quad (\text{A.4})$$

$$B_{\text{Zubringer Ziel}} = \text{MFSN} \rightarrow \text{Ma\ssnahme} = W_{\ddot{u}}, G(2.0, 0.0) \quad (\text{A.5})$$

$$t_{F \text{ Abbringer, resthalte}} \leq 0 \wedge t_{F \text{ Zubringer, donefahrdauer}} \leq 98 \rightarrow \text{Ma\ssnahme} = A(41.0, 2.0) \quad (\text{A.6})$$

$$\begin{aligned}
t_{F_{\text{Abbringer}}, \text{ donehalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{ restfahrdauer}} \geq 105 \wedge \\
t_{A, F_{\text{Zubringer}}, \text{ eingangverspätung}} \geq 7 \wedge \\
t_{H_{B, F_{\text{Zubringer}}}, \text{ soll ankunft uhrzeit}} \leq 12:58 \wedge
\end{aligned} \tag{A.7}$$

$$\begin{aligned}
t_{\text{mindestübergangszeit}} \leq 7 \wedge t_{F_{\text{Zubringer}}, \text{ donehalte}} \leq 12 \rightarrow \text{Maßnahme} = A(13.0, 1.0) \\
t_{F_{\text{Abbringer}}, \text{ donehalte}} \leq 0 \wedge t_{H_{B, F_{\text{Zubringer}}}, \text{ soll haltezeit}} \geq 5 \wedge \\
t_{H_{B, F_{\text{Zubringer}}}, \text{ soll ankunft uhrzeit}} \geq 19:08 \wedge
\end{aligned} \tag{A.8}$$

$$\begin{aligned}
t_{F_{\text{Zubringer}}, \text{ donefahrdauer}} \geq 175 \rightarrow \text{Maßnahme} = A(13.0, 0.0) \\
d_{H_{B, F_{\text{Zubringer}}}, \text{ soll ankunft wochentag}} \geq 2 \wedge B_{\text{Zubringer Ziel}} = \text{MH} \wedge \\
t_{H_{B, F_{\text{Zubringer}}}, \text{ soll ankunft uhrzeit}} \geq 10:41 \wedge \\
t_{H_{B, F_{\text{Zubringer}}}, \text{ soll ankunft uhrzeit}} \leq 15:23 \rightarrow \text{Maßnahme} = A(7.0, 0.0)
\end{aligned} \tag{A.9}$$

$$\begin{aligned}
t_A, \text{ Regelwartezeit} \geq 3 \wedge t_{H_{B, F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \leq 10 \wedge \\
t_A, \text{ soll umstieg} \geq 9 \rightarrow \text{Maßnahme} = W_r(169.0, 84.0)
\end{aligned} \tag{A.10}$$

$$\begin{aligned}
t_A, \text{ soll umstieg} \leq 8 \wedge t_{H_{B, F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \geq 8 \wedge \\
t_{H_{B, F_{\text{Abbringer}}}, \text{ soll haltezeit}} \leq 13 \rightarrow \text{Maßnahme} = W_w(127.0, 40.0)
\end{aligned} \tag{A.11}$$

$$\begin{aligned}
t_A, \text{ soll umstieg} \leq 7 \wedge t_{H_{B, F_{\text{Zubringer}}}, \text{ ist ankunft uhrzeit}} \leq 19:32 \wedge \\
t_{H_{B, F_{\text{Abbringer}}}, \text{ soll haltezeit}} \leq 11 \wedge
\end{aligned} \tag{A.12}$$

$$t_{H_{B, F_{\text{Zubringer}}}, \text{ ist haltezeit}} \leq 7 \wedge t_A, \text{ soll umstieg} \leq 5 \rightarrow \text{Maßnahme} = W_w(49.0, 11.0)$$

$$\begin{aligned}
t_{H_{B, F_{\text{Zubringer}}}, \text{ ist ankunft uhrzeit}} \leq 20:24 \wedge t_A, \text{ soll umstieg} \leq 19 \wedge \\
t_{H_{B, F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \geq 14 \wedge \\
t_{H_{B, F_{\text{Abbringer}}}, \text{ soll haltezeit}} \leq 14 \rightarrow \text{Maßnahme} = W_w(61.0, 18.0)
\end{aligned} \tag{A.13}$$

$$\begin{aligned}
t_A, \text{ soll umstieg} \leq 11 \wedge t_{H_{B, F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \geq 3 \wedge \\
t_{H_{B, F_{\text{Abbringer}}}, \text{ soll haltezeit}} \leq 11 \wedge t_{F_{\text{Abbringer}}, \text{ resthalte}} \geq 12 \wedge
\end{aligned} \tag{A.14}$$

$$t_{F_{\text{Zubringer}}, \text{ resthalte}} \leq 9 \wedge t_{\text{mindestübergangszeit}} \geq 5 \rightarrow \text{Maßnahme} = W_w(31.0, 7.0)$$

$$\begin{aligned}
t_{H_{B, F_{\text{Zubringer}}}, \text{ ist ankunft uhrzeit}} \geq 08:32 \wedge t_A, \text{ soll umstieg} \leq 10 \wedge \\
t_{F_{\text{Abbringer}}, \text{ restfahrdauer}} \leq 43 \wedge t_A, \text{ soll umstieg} \leq 7 \rightarrow \text{Maßnahme} = W_w(41.0, 16.0)
\end{aligned} \tag{A.15}$$

$$t_{H_{B, F_{\text{Abbringer}}}, \text{ soll haltezeit}} \geq 0 \wedge \tag{A.16}$$

$$\begin{aligned}
t_{H_{B, F_{\text{Zubringer}}}, \text{ ankunft verspaetung}} \geq 27 \wedge t_A, \text{ soll umstieg} \leq 38 \rightarrow \text{Maßnahme} = W_w(36.0, 14.0) \\
t_{F_{\text{Abbringer}}, \text{ donehalte}} \leq 0 \rightarrow \text{Maßnahme} = G(245.0, 14.0)
\end{aligned} \tag{A.17}$$

$$d_{H_{B, F_{\text{Abbringer}}}, \text{ soll abfahrt wochentag}} \geq 2 \wedge \tag{A.18}$$

$$\begin{aligned}
t_{H_{B, F_{\text{Abbringer}}}, \text{ soll abfahrt uhrzeit}} \geq 10:19 \rightarrow \text{Maßnahme} = G(117.0, 5.0) \\
\rightarrow \text{Maßnahme} = W_{ii}(547.0, 207.0)
\end{aligned} \tag{A.19}$$

A.2.4 Featuremenge \mathbb{F}_4

$$\begin{aligned} \text{Richtung}_{\text{Abbringer}} &= \text{München Hbf} - \text{Hamburg-Altona} \wedge \\ &\text{Linie}_{\text{Zubringer}} = 10.1 \rightarrow \text{Maßnahme} = A, W_{\ddot{u}}(3.0, 1.0) \end{aligned} \quad (\text{A.1})$$

$$\begin{aligned} \text{Richtung}_{\text{Abbringer}} &= \text{Essen Hbf} - \text{Würzburg Hbf} \rightarrow \text{Maßnahme} = A, G(3.0, 1.0) \end{aligned} \quad (\text{A.2})$$

$$\begin{aligned} t_{F_{\text{Abbringer}}, \text{donefahrtdauer}} &\geq 204 \wedge \\ t_{H_{B,F_{\text{Abbringer}}}, \text{soll haltezeit}} &\geq 7 \wedge \\ t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} &\geq 13 \wedge t_{A, \text{soll umstieg}} \leq 12 \rightarrow \text{Maßnahme} = A, W_w(4.0, 0.0) \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned} t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} &\geq 15 \wedge \text{Linie}_{\text{Abbringer}} = 28.2 \wedge \\ t_{A, \text{soll umstieg}} &\leq 18 \rightarrow \text{Maßnahme} = W_w, G(4.0, 0.0) \end{aligned} \quad (\text{A.4})$$

$$\begin{aligned} t_{H_{B,F_{\text{Zubringer}}}, \text{ankunft verspaetung}} &\geq 11 \wedge B_{\text{Zubringer Start}} = \text{XDKH} \rightarrow \text{Maßnahme} = W_{\ddot{u}}, G(3.0, 0.0) \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned} \text{RIL100} &= \text{LF} \rightarrow \text{Maßnahme} = W_{\ddot{u}}, G(2.0, 0.0) \end{aligned} \quad (\text{A.6})$$

$$\begin{aligned} \text{Richtung}_{\text{Zubringer}} &= \text{Frankfurt (Main) Flugh Fernbf} - \text{Ostseebad Binz} \rightarrow \text{Maßnahme} = W_{\ddot{u}}, G(4.0, 1.0) \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned} \text{RIL100} &= \text{WLO} \rightarrow \text{Maßnahme} = W_{\ddot{u}}, G(2.0, 0.0) \end{aligned} \quad (\text{A.8})$$

$$\begin{aligned} B_{\text{Zubringer Ziel}} &= \text{MFSN} \rightarrow \text{Maßnahme} = W_{\ddot{u}}, G(2.0, 0.0) \end{aligned} \quad (\text{A.9})$$

$$\begin{aligned} \text{Linie}_{\text{Abbringer}} &= 78.1 \rightarrow \text{Maßnahme} = W_{\ddot{u}}, G(3.0, 1.0) \end{aligned} \quad (\text{A.10})$$

$$\begin{aligned} \text{RIL100} &= \text{BWP} \rightarrow \text{Maßnahme} = W_{\ddot{u}}, G(3.0, 1.0) \end{aligned} \quad (\text{A.11})$$

$$\begin{aligned} t_{F_{\text{Abbringer}}, \text{resthalte}} &\leq 0 \wedge t_{F_{\text{Zubringer}}, \text{donehalte}} \leq 8 \rightarrow \text{Maßnahme} = A(52.0, 6.0) \end{aligned} \quad (\text{A.12})$$

$$\begin{aligned} t_{F_{\text{Abbringer}}, \text{donehalte}} &\leq 0 \wedge t_{F_{\text{Zubringer}}, \text{restfahrtdauer}} \geq 105 \wedge \\ t_{A, \text{soll umstieg}} &\leq 14 \wedge t_{A,F_{\text{Zubringer}}, \text{eingangsverspaetung}} \geq 7 \wedge \\ t_{F_{\text{Zubringer}}, \text{resthalte}} &\leq 6 \rightarrow \text{Maßnahme} = A(10.0, 0.0) \end{aligned} \quad (\text{A.13})$$

$$\begin{aligned} t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft}} &\geq \text{2015-06-16 09:23} \wedge \\ t_{H_{B,F_{\text{Zubringer}}}, \text{soll haltezeit}} &\geq 3 \wedge \\ t_{F_{\text{Zubringer}}, \text{restfahrtdauer}} &\geq 272 \rightarrow \text{Maßnahme} = A(8.0, 2.0) \end{aligned} \quad (\text{A.14})$$

$$\begin{aligned}
& t_{F_{\text{Abbringer}}, \text{ donehalte}} \leq 0 \wedge t_{H_{B,F_{\text{Zubringer}}, \text{ soll haltezeit}} \geq 5 \wedge \\
& t_{A, \text{ soll umstieg}} \leq 14 \wedge t_{\text{mindestübergangszeit}} \geq 7 \rightarrow \text{Maßnahme} = A(11.0, 2.0)
\end{aligned}
\tag{A.15}$$

$$\begin{aligned}
& t_{H_{B,F_{\text{Zubringer}}, \text{ soll ankunft}} \geq 2015-06-16 \text{ 09:23} \wedge \\
& t_{A,F_{\text{Zubringer}}, \text{ eingangsverspätung}} \geq 9 \wedge
\end{aligned}
\tag{A.16}$$

$$\begin{aligned}
& t_{A,F_{\text{Zubringer}}, \text{ eingangsverspätung}} \leq 11 \wedge \\
& t_{F_{\text{Abbringer}}, \text{ donefahrdauer}} \leq 113 \rightarrow \text{Maßnahme} = A(5.0, 0.0) \\
& t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_{B,F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \leq 11 \wedge \\
& t_{A, \text{ soll umstieg}} \geq 9 \wedge t_{F_{\text{Abbringer}}, \text{ donefahrdauer}} \leq 111 \wedge \\
& t_{H_{B,F_{\text{Zubringer}}, \text{ soll ankunft}} \leq 2015-06-15 \text{ 11:55} \wedge \\
& t_{A, \text{ soll umstieg}} \geq 13 \rightarrow \text{Maßnahme} = W_r(18.0, 1.0)
\end{aligned}
\tag{A.17}$$

$$\begin{aligned}
& t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_{B,F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \leq 10 \wedge \\
& t_{A, \text{ soll umstieg}} \geq 14 \wedge t_{F_{\text{Zubringer}}, \text{ donehalte}} \geq 5 \rightarrow \text{Maßnahme} = W_r(61.0, 22.0)
\end{aligned}
\tag{A.18}$$

$$\begin{aligned}
& t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_{B,F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \leq 5 \wedge \\
& t_{A, \text{ soll umstieg}} \geq 8 \wedge t_{F_{\text{Zubringer}}, \text{ donefahrdauer}} \geq 133 \wedge \\
& t_{F_{\text{Zubringer}}, \text{ donefahrdauer}} \leq 181 \rightarrow \text{Maßnahme} = W_r(20.0, 5.0)
\end{aligned}
\tag{A.19}$$

$$\begin{aligned}
& t_{A, \text{ soll umstieg}} \leq 12 \wedge t_{H_{B,F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 8 \wedge \\
& t_{H_{B,F_{\text{Zubringer}}, \text{ soll ankunft}} \leq 2015-06-15 \text{ 12:58} \wedge \\
& t_{H_{B,F_{\text{Abbringer}}, \text{ soll haltezeit}} \leq 14 \rightarrow \text{Maßnahme} = W_w(69.0, 14.0)
\end{aligned}
\tag{A.20}$$

$$\begin{aligned}
& t_{A, \text{ soll umstieg}} \leq 8 \wedge t_{H_{B,F_{\text{Zubringer}}, \text{ wartezeit}} \leq 6 \wedge \\
& t_{A, \text{ soll umstieg}} \leq 5 \rightarrow \text{Maßnahme} = W_w(96.0, 34.0)
\end{aligned}
\tag{A.21}$$

$$\begin{aligned}
& t_{A, \text{ soll umstieg}} \leq 10 \wedge t_{H_{B,F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 4 \wedge \\
& t_{H_{B,F_{\text{Abbringer}}, \text{ soll haltezeit}} \geq 0 \wedge B_{\text{Zubringer Start}} = AA \rightarrow \text{Maßnahme} = W_w(30.0, 9.0)
\end{aligned}
\tag{A.22}$$

$$\begin{aligned}
& t_{H_{B,F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 3 \wedge t_{A, \text{ soll umstieg}} \leq 9 \wedge \\
& t_{F_{\text{Abbringer}}, \text{ restfahrdauer}} \geq 306 \rightarrow \text{Maßnahme} = W_w(20.0, 5.0)
\end{aligned}
\tag{A.23}$$

$$\begin{aligned}
& t_{H_{B,F_{\text{Zubringer}}, \text{ ist ankunft}} \geq 2015-06-15 \text{ 13:30} \wedge \\
& t_{H_{B,F_{\text{Abbringer}}, \text{ soll haltezeit}} \geq 1 \wedge \\
& t_{H_{B,F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 23 \wedge t_{A, \text{ soll umstieg}} \leq 33 \rightarrow \text{Maßnahme} = W_w(41.0, 8.0)
\end{aligned}
\tag{A.24}$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \rightarrow \text{Maßnahme} = G(237.0, 12.0) \quad (\text{A.25})$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{soll abfahrt}} \geq 2015-06-16 \ 10:10 \rightarrow \text{Maßnahme} = G(115.0, 4.0) \quad (\text{A.26})$$

$$\rightarrow \text{Maßnahme} = W_{\ddot{u}}(686.0, 289.0) \quad (\text{A.27})$$

A.3 Klassenzusammenstellung \mathbb{K}_3

A.3.1 Featuremenge \mathbb{F}_1

$$t_{F_{\text{Abbringer}}, \text{resthalte}} \leq 0 \wedge t_{F_{\text{Zubringer}}, \text{donefahrtdauer}} \leq 98 \rightarrow \text{Maßnahme} = A(41.0, 2.0) \quad (\text{A.1})$$

$$t_{F_{\text{Abbringer}}, \text{donefahrtdauer}} \leq 3 \wedge t_{F_{\text{Zubringer}}, \text{restfahrtdauer}} \geq 105 \wedge$$

$$t_{A,F_{\text{Zubringer}}, \text{eingangsverspätung}} \geq 7 \wedge t_A, \text{soll umstieg} \leq 17 \wedge \quad (\text{A.2})$$

$$t_{F_{\text{Zubringer}}, \text{resthalte}} \leq 12 \rightarrow \text{Maßnahme} = A(14.0, 1.0)$$

$$t_{F_{\text{Abbringer}}, \text{donehalte}} \leq 0 \wedge B_{\text{Zubringer Start}} = \text{LL} \wedge$$

$$t_{F_{\text{Zubringer}}, \text{donehalte}} \leq 6 \rightarrow \text{Maßnahme} = A(9.0, 1.0) \quad (\text{A.3})$$

$$t_{F_{\text{Abbringer}}, \text{donefahrtdauer}} \leq 0 \wedge B_{\text{Zubringer Ziel}} = \text{BHF} \wedge$$

$$t_{F_{\text{Zubringer}}, \text{donehalte}} \leq 10 \rightarrow \text{Maßnahme} = A(8.0, 0.0) \quad (\text{A.4})$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft}} \geq 2015-06-16 \ 10:41 \wedge$$

$$B_{\text{Zubringer Ziel}} = \text{MH} \wedge \quad (\text{A.5})$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft}} \leq 2015-06-16 \ 15:23 \rightarrow \text{Maßnahme} = A(7.0, 0.0)$$

$$B_{\text{Zubringer Start}} = \text{FG} \rightarrow \text{Maßnahme} = A(2.0, 0.0) \quad (\text{A.6})$$

$$t_{H_{B,F_{\text{Zubringer}}}, \text{soll ankunft}} \geq 2015-06-16 \ 09:23 \wedge$$

$$B_{\text{Zubringer Start}} = \text{MH} \wedge t_{F_{\text{Zubringer}}, \text{donefahrtdauer}} \geq 364 \rightarrow \text{Maßnahme} = A(5.0, 0.0) \quad (\text{A.7})$$

$$t_A, \text{Regelwartezeit} \geq 3 \wedge t_{H_{B,F_{\text{Abbringer}}}, \text{abfahrt verspätung}} \leq 3 \wedge$$

$$t_A, \text{ist umstieg} \geq 6 \wedge t_{H_{B,F_{\text{Abbringer}}}, \text{wartezeit}} \geq 1 \wedge \quad (\text{A.8})$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{abfahrt verspätung}} \geq 1 \rightarrow \text{Maßnahme} = W_r(128.0, 0.0)$$

$$\text{Richtung}_{\text{Zubringer}} = \text{Karlsruhe Hbf} - \text{Kiel Hbf} \rightarrow \text{Maßnahme} = W_r(2.0, 0.0) \quad (\text{A.9})$$

$$t_A, \text{ist umstieg} \leq 4 \wedge t_{H_{B,F_{\text{Abbringer}}}, \text{ist haltezeit}} \leq 20 \wedge$$

$$t_{H_{B,F_{\text{Abbringer}}}, \text{abfahrt verspätung}} \leq 1 \rightarrow \text{Maßnahme} = W_w(196.0, 4.0) \quad (\text{A.10})$$

$$t_A, \text{ist umstieg} \leq 5 \wedge t_{F_{\text{Abbringer}}, \text{donehalte}} \geq 2 \wedge t_A, \text{ist umstieg} \leq 3 \wedge$$

$$t_{F_{\text{Abbringer}}, \text{donefahrtdauer}} \leq 75 \rightarrow \text{Maßnahme} = W_w(52.0, 3.0) \quad (\text{A.11})$$

$$t_A, \text{ist umstieg} \leq 4 \wedge t_{H_{B,F_{\text{Abbringer}}}, \text{soll haltezeit}} \geq 8 \wedge$$

$$t_{\text{mindestübergangszeit}} \geq 5 \rightarrow \text{Maßnahme} = W_w(20.0, 0.0) \quad (\text{A.12})$$

$$t_{A, \text{ ist umstieg}} \leq 5 \wedge t_{H_{B,F} \text{ Abbringer}, \text{ wartezeit}} \leq 3 \wedge \quad (\text{A.13})$$

$$t_{\text{mindestübergangszeit}} \geq 6 \rightarrow \text{Maßnahme} = W_w(32.0, 1.0)$$

$$t_{A, \text{ ist umstieg}} \leq 4 \wedge t_{F \text{ Abbringer}, \text{ donehalte}} \geq 2 \wedge \quad (\text{A.14})$$

$$t_{H_{B,F} \text{ Abbringer}, \text{ wartezeit}} \leq 2 \wedge t_{\text{mindestübergangszeit}} \geq 5 \rightarrow \text{Maßnahme} = W_w(15.0, 0.0)$$

$$t_{A, \text{ ist umstieg}} \leq 6 \wedge t_{\text{mindestübergangszeit}} \geq 7 \wedge \quad (\text{A.15})$$

$$t_{H_{B,F} \text{ Abbringer}, \text{ ist haltezeit}} \leq 8 \rightarrow \text{Maßnahme} = W_w(10.0, 1.0)$$

$$t_{A, \text{ ist umstieg}} \leq -4 \wedge t_{H_{B,F} \text{ Abbringer}, \text{ soll haltezeit}} \geq 2 \rightarrow \text{Maßnahme} = W_w(8.0, 2.0) \quad (\text{A.16})$$

$$t_{F \text{ Abbringer}, \text{ donehalte}} \leq 0 \rightarrow \text{Maßnahme} = G(240.0, 5.0) \quad (\text{A.17})$$

$$t_{H_{B,F} \text{ Abbringer}, \text{ soll abfahrt}} \geq 2015-06-16 \ 10:00 \rightarrow \text{Maßnahme} = G(117.0, 3.0) \quad (\text{A.18})$$

$$t_{H_{B,F} \text{ Abbringer}, \text{ abfahrt verspaetung}} \leq 0 \rightarrow \text{Maßnahme} = G(5.0, 0.0) \quad (\text{A.19})$$

$$t_{A, \text{ ist umstieg}} \leq 2 \wedge t_{H_{B,F} \text{ Zubringer}, \text{ soll ankunft}} \leq 2015-06-16 \ 08:53 \rightarrow \text{Maßnahme} = G(5.0, 0.0) \quad (\text{A.20})$$

$$t_{H_{B,F} \text{ Abbringer}, \text{ wartezeit}} \leq 0 \rightarrow \text{Maßnahme} = G(2.0, 0.0) \quad (\text{A.21})$$

$$\rightarrow \text{Maßnahme} = W_{\ddot{u}}(492.0, 33.0) \quad (\text{A.22})$$

A.3.2 Featuremenge \mathbb{F}_2

$$t_{F \text{ Abbringer}, \text{ donehalte}} \leq 0 \wedge t_{F \text{ Abbringer}, \text{ resthalte}} \leq 0 \wedge$$

$$t_{F \text{ Zubringer}, \text{ donefahrtdauer}} \leq 98 \rightarrow \text{Maßnahme} = A(41.0, 2.0)$$

(A.1)

$$t_{F \text{ Abbringer}, \text{ donefahrtdauer}} \leq 3 \wedge t_{F \text{ Zubringer}, \text{ restfahrtdauer}} \geq 105 \wedge$$

$$B_{\text{Zubringer Ziel}} = \text{BHF} \rightarrow \text{Maßnahme} = A(9.0, 1.0)$$

(A.2)

$$t_{F \text{ Abbringer}, \text{ donehalte}} \leq 0 \wedge t_{F \text{ Zubringer}, \text{ donefahrtdauer}} \geq 202 \wedge$$

$$t_{F \text{ Zubringer}, \text{ donefahrtdauer}} \leq 208 \rightarrow \text{Maßnahme} = A(14.0, 2.0)$$

(A.3)

$$t_{H_{B,F} \text{ Zubringer}, \text{ soll ankunft}} \geq 2015-06-16 \ 09:28 \wedge$$

$$B_{\text{Zubringer Ziel}} = \text{MH} \wedge$$

$$t_{H_{B,F} \text{ Zubringer}, \text{ soll ankunft}} \leq 2015-06-16 \ 15:23 \rightarrow \text{Maßnahme} = A(9.0, 0.0)$$

(A.4)

$$\text{Richtung}_{\text{Zubringer}} = \text{München Hbf} - \text{Berlin-Gesundbrunnen} \wedge \text{RIL100} = \text{BL} \rightarrow \text{Maßnahme} = A(3.0, 0.0)$$

(A.5)

$$B_{\text{Zubringer Start}} = \text{FG} \rightarrow \text{Maßnahme} = A(2.0, 0.0)$$

(A.6)

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{A, \text{ soll umstieg}} \geq 13 \wedge$$

$$t_{H_{B,F} \text{ Zubringer}, \text{ ankunft verspaetung}} \leq 11 \wedge$$

$$t_{F \text{ Abbringer}, \text{ donefahrtdauer}} \leq 111 \rightarrow \text{Maßnahme} = W_r(45.0, 12.0)$$

(A.7)

$$\begin{aligned}
t_{A, \text{ Regelwartezeit}} &\geq 3 \wedge t_{A, \text{ soll umstieg}} \geq 9 \wedge \\
t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} &\leq 11 \wedge \\
t_{H_{B,F} \text{Zubringer}, \text{ soll haltezeit}} &\leq 2 \rightarrow \text{Maßnahme} = W_r(28.0, 8.0)
\end{aligned}
\tag{A.8}$$

$$\begin{aligned}
t_{A, \text{ Regelwartezeit}} &\geq 3 \wedge t_{A, \text{ soll umstieg}} \geq 10 \wedge \\
t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} &\leq 23 \wedge \\
t_{F \text{Zubringer}, \text{ donefahrtdauer}} &\leq 179 \wedge t_{F \text{Zubringer}, \text{ donehalte}} \geq 5 \rightarrow \text{Maßnahme} = W_r(31.0, 11.0)
\end{aligned}
\tag{A.9}$$

$$\begin{aligned}
t_{A, \text{ soll umstieg}} &\leq 9 \wedge t_{H_{B,F} \text{Zubringer}, \text{ wartezeit}} \leq 6 \wedge \\
t_{A, \text{ soll umstieg}} &\leq 6 \wedge t_{F \text{Zubringer}, \text{ donefahrtdauer}} \leq 103 \rightarrow \text{Maßnahme} = W_w(74.0, 19.0)
\end{aligned}
\tag{A.10}$$

$$\begin{aligned}
t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} &\geq 8 \wedge t_{A, \text{ soll umstieg}} \leq 15 \wedge \\
t_{F \text{Zubringer}, \text{ donefahrtdauer}} &\geq 258 \wedge t_{F \text{Abbringer}, \text{ donehalte}} \geq 6 \rightarrow \text{Maßnahme} = W_w(29.0, 3.0)
\end{aligned}
\tag{A.11}$$

$$\begin{aligned}
t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} &\geq 6 \wedge t_{A, \text{ soll umstieg}} \leq 13 \wedge \\
t_{H_{B,F} \text{Abbringer}, \text{ soll haltezeit}} &\leq 14 \wedge \\
t_{F \text{Abbringer}, \text{ donefahrtdauer}} &\leq 130 \wedge t_{F \text{Zubringer}, \text{ donefahrtdauer}} \geq 80 \rightarrow \text{Maßnahme} = W_w(102.0, 28.0)
\end{aligned}
\tag{A.12}$$

$$\begin{aligned}
t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} &\geq 10 \wedge t_{A, \text{ soll umstieg}} \leq 31 \wedge \\
t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} &\geq 21 \wedge \\
t_{H_{B,F} \text{Abbringer}, \text{ soll haltezeit}} &\leq 4 \rightarrow \text{Maßnahme} = W_w(36.0, 7.0)
\end{aligned}
\tag{A.13}$$

$$t_{F \text{Abbringer}, \text{ donehalte}} \leq 0 \rightarrow \text{Maßnahme} = G(237.0, 12.0)
\tag{A.14}$$

$$t_{H_{B,F} \text{Abbringer}, \text{ soll abfahrt}} \geq 2015-06-16 \ 10:00 \rightarrow \text{Maßnahme} = G(117.0, 3.0)
\tag{A.15}$$

$$\rightarrow \text{Maßnahme} = W_{\dot{i}}(633.0, 228.0)
\tag{A.16}$$

A.3.3 Featuremenge \mathbb{F}_3

$$t_{F \text{Abbringer}, \text{ resthalte}} \leq 0 \wedge t_{F \text{Zubringer}, \text{ donefahrtdauer}} \leq 98 \rightarrow \text{Maßnahme} = A(41.0, 2.0)
\tag{A.1}$$

$$\begin{aligned}
t_{F \text{Abbringer}, \text{ donehalte}} &\leq 0 \wedge t_{F \text{Zubringer}, \text{ restfahrtdauer}} \geq 105 \wedge \\
t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} &\geq 11 \wedge \\
t_{F \text{Zubringer}, \text{ resthalte}} &\leq 10 \wedge
\end{aligned}
\tag{A.2}$$

$$t_{H_{B,F} \text{Zubringer}, \text{ ankunft verspaetung}} \leq 20 \rightarrow \text{Maßnahme} = A(13.0, 1.0)$$

$$t_{F\text{Abbringer}, \text{donehalte}} \leq 0 \wedge t_{F\text{Zubringer}, \text{donefahrtdauer}} \geq 202 \wedge \quad (\text{A.3})$$

$$t_{F\text{Zubringer}, \text{donefahrtdauer}} \leq 208 \rightarrow \text{Ma\ssnahme} = A(14.0, 2.0)$$

$$d_{H_B, F\text{Zubringer}, \text{soll ankunft wochentag}} \geq 2 \wedge B_{\text{Zubringer Ziel}} = \text{MH} \wedge$$

$$t_{H_B, F\text{Zubringer}, \text{soll ankunft uhrzeit}} \geq 09:28 \wedge \quad (\text{A.4})$$

$$t_{H_B, F\text{Zubringer}, \text{soll ankunft uhrzeit}} \leq 15:23 \rightarrow \text{Ma\ssnahme} = A(9.0, 0.0)$$

$$B_{\text{Zubringer Start}} = \text{FG} \rightarrow \text{Ma\ssnahme} = A(2.0, 0.0) \quad (\text{A.5})$$

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{H_B, F\text{Zubringer}, \text{ankunft verspaetung}} \leq 11 \wedge$$

$$t_{A, \text{soll umstieg}} \geq 14 \wedge t_{F\text{Zubringer}, \text{restfahrtdauer}} \leq 124 \rightarrow \text{Ma\ssnahme} = W_r(49.0, 14.0) \quad (\text{A.6})$$

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{A, \text{soll umstieg}} \geq 13 \wedge$$

$$t_{H_B, F\text{Zubringer}, \text{ankunft verspaetung}} \leq 9 \wedge \quad (\text{A.7})$$

$$t_{H_B, F\text{Zubringer}, \text{wartezeit}} \leq 4 \wedge t_{F\text{Zubringer}, \text{resthalte}} \geq 7 \rightarrow \text{Ma\ssnahme} = W_r(23.0, 4.0)$$

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{A, \text{soll umstieg}} \geq 9 \wedge$$

$$t_{H_B, F\text{Zubringer}, \text{ist ankunft uhrzeit}} \leq 18:59 \wedge$$

$$t_{F\text{Abbringer}, \text{restfahrtdauer}} \geq 273 \wedge \quad (\text{A.8})$$

$$t_{H_B, F\text{Abbringer}, \text{soll haltezeit}} \leq 5 \wedge$$

$$t_{H_B, F\text{Zubringer}, \text{ankunft verspaetung}} \leq 10 \rightarrow \text{Ma\ssnahme} = W_r(18.0, 3.0)$$

$$t_{A, \text{soll umstieg}} \leq 7 \wedge t_{H_B, F\text{Zubringer}, \text{ankunft verspaetung}} \geq 9 \wedge$$

$$t_{F\text{Abbringer}, \text{resthalte}} \geq 4 \wedge t_{F\text{Zubringer}, \text{donehalte}} \leq 9 \rightarrow \text{Ma\ssnahme} = W_w(50.0, 5.0) \quad (\text{A.9})$$

$$t_{A, \text{soll umstieg}} \leq 12 \wedge t_{H_B, F\text{Zubringer}, \text{ankunft verspaetung}} \geq 5 \wedge$$

$$t_{H_B, F\text{Abbringer}, \text{soll haltezeit}} \leq 12 \wedge \quad (\text{A.10})$$

$$t_{F\text{Abbringer}, \text{donefahrtdauer}} \leq 130 \wedge t_{F\text{Zubringer}, \text{donefahrtdauer}} \geq 86 \rightarrow \text{Ma\ssnahme} = W_w(111.0, 35.0)$$

$$t_{H_B, F\text{Zubringer}, \text{wartezeit}} \leq 6 \wedge t_{A, \text{soll umstieg}} \leq 5 \wedge$$

$$t_{A, F\text{Zubringer}, \text{eingangsverspaetung}} \leq 5 \rightarrow \text{Ma\ssnahme} = W_w(48.0, 13.0) \quad (\text{A.11})$$

$$t_{H_B, F\text{Zubringer}, \text{ankunft verspaetung}} \geq 14 \wedge t_{A, \text{soll umstieg}} \leq 33 \wedge$$

$$t_{H_B, F\text{Abbringer}, \text{soll haltezeit}} \geq 1 \wedge \quad (\text{A.12})$$

$$t_{H_B, F\text{Zubringer}, \text{ankunft verspaetung}} \geq 25 \rightarrow \text{Ma\ssnahme} = W_w(34.0, 4.0)$$

$$t_{H_B, F\text{Abbringer}, \text{soll haltezeit}} \geq 4 \wedge t_{A, \text{soll umstieg}} \leq 6 \wedge$$

$$t_{F\text{Abbringer}, \text{restfahrtdauer}} \leq 43 \rightarrow \text{Ma\ssnahme} = W_w(21.0, 5.0) \quad (\text{A.13})$$

$$t_{H_B, F\text{Zubringer}, \text{ankunft verspaetung}} \geq 14 \wedge t_{A, \text{soll umstieg}} \leq 19 \wedge$$

$$t_{H_B, F\text{Abbringer}, \text{soll haltezeit}} \geq 3 \rightarrow \text{Ma\ssnahme} = W_w(26.0, 8.0) \quad (\text{A.14})$$

$$t_{F\text{Abbringer}, \text{donehalte}} \leq 0 \rightarrow \text{Ma\ssnahme} = G(242.0, 11.0) \quad (\text{A.15})$$

$$d_{H_B, F\text{Abbringer}, \text{soll abfahrt wochentag}} \geq 2 \wedge$$

$$t_{H_B, F\text{Abbringer}, \text{soll abfahrt uhrzeit}} \geq 10:10 \rightarrow \text{Ma\ssnahme} = G(117.0, 4.0) \quad (\text{A.16})$$

$$\rightarrow \text{Ma\ssnahme} = W_{ii}(592.0, 199.0) \quad (\text{A.17})$$

A.3.4 Featuremenge \mathbb{F}_4

$$t_{F\text{Abbringer, resthalte}} \leq 0 \wedge t_{F\text{Zubringer, donefahrtdauer}} \leq 43 \rightarrow \text{Ma\ssnahme} = A(35.0, 0.0) \quad (\text{A.1})$$

$$t_{F\text{Abbringer, donehalte}} \leq 0 \wedge t_{H_B, F\text{Zubringer, soll haltezeit}} \geq 5 \wedge t_{F\text{Abbringer, resthalte}} \leq 3 \rightarrow \text{Ma\ssnahme} = A(10.0, 2.0) \quad (\text{A.2})$$

$$t_{F\text{Abbringer, donefahrtdauer}} \leq 3 \wedge t_{F\text{Zubringer, restfahrtdauer}} \geq 113 \wedge t_{A, F\text{Zubringer, eingangversp\u00e4tung}} \geq 7 \wedge \text{Ma\ssnahme} = A(16.0, 3.0) \quad (\text{A.3})$$

$$t_{F\text{Zubringer, donefahrtdauer}} \leq 259 \wedge t_{F\text{Abbringer, resthalte}} \leq 20 \rightarrow \text{Ma\ssnahme} = A(16.0, 3.0) \wedge t_{H_B, F\text{Zubringer, soll ankunft}} \geq 2015-06-16 \ 11:09 \wedge t_{F\text{Zubringer, restfahrtdauer}} \geq 273 \wedge t_{A, \text{Regelwartezeit}} \geq 3 \rightarrow \text{Ma\ssnahme} = A(6.0, 1.0) \quad (\text{A.4})$$

$$t_{F\text{Abbringer, donehalte}} \leq 0 \wedge \text{RIL100} = \text{FF} \wedge t_{A, \text{soll umstieg}} \leq 14 \wedge t_{H_B, F\text{Zubringer, soll ankunft}} \geq 2015-06-15 \ 19:08 \rightarrow \text{Ma\ssnahme} = A(7.0, 0.0) \quad (\text{A.5})$$

$$\text{Richtung}_{Z\text{ubringer}} = \text{N\u00fcrnberg Hbf} - \text{M\u00fcnchen Hbf} \rightarrow \text{Ma\ssnahme} = A(2.0, 0.0) \quad (\text{A.6})$$

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{A, \text{soll umstieg}} \geq 13 \wedge t_{H_B, F\text{Zubringer, ankunft verspaetzung}} \leq 17 \rightarrow \text{Ma\ssnahme} = W_r(171.0, 82.0) \quad (\text{A.7})$$

$$t_{A, \text{soll umstieg}} \leq 8 \wedge t_{H_B, F\text{Zubringer, ankunft verspaetzung}} \geq 8 \rightarrow \text{Ma\ssnahme} = W_w(136.0, 45.0) \quad (\text{A.8})$$

$$t_{H_B, F\text{Zubringer, ist ankunft}} \geq 2015-06-15 \ 07:27 \wedge t_{A, \text{soll umstieg}} \leq 7 \wedge t_{H_B, F\text{Abbringer, soll haltezeit}} \leq 11 \wedge \text{Ma\ssnahme} = W_w(43.0, 15.0) \quad (\text{A.9})$$

$$t_{H_B, F\text{Abbringer, soll haltezeit}} \geq 6 \rightarrow \text{Ma\ssnahme} = W_w(43.0, 15.0)$$

$$t_{H_B, F\text{Zubringer, ankunft verspaetzung}} \geq 3 \wedge t_{H_B, F\text{Abbringer, soll haltezeit}} \leq 16 \wedge t_{A, \text{soll umstieg}} \leq 19 \wedge \text{Ma\ssnahme} = W_w(61.0, 15.0) \quad (\text{A.10})$$

$$t_{H_B, F\text{Abbringer, soll haltezeit}} \leq 16 \wedge t_{A, \text{soll umstieg}} \leq 19 \wedge t_{H_B, F\text{Zubringer, ankunft verspaetzung}} \geq 14 \rightarrow \text{Ma\ssnahme} = W_w(61.0, 15.0) \quad (\text{A.10})$$

$$t_{H_B, F\text{Zubringer, ankunft verspaetzung}} \geq 14 \rightarrow \text{Ma\ssnahme} = W_w(61.0, 15.0)$$

$$t_{H_B, F\text{Zubringer, ankunft verspaetzung}} \geq 4 \wedge t_{H_B, F\text{Abbringer, soll haltezeit}} \leq 14 \wedge t_{A, \text{soll umstieg}} \leq 9 \wedge t_{F\text{Abbringer, resthalte}} \geq 12 \rightarrow \text{Ma\ssnahme} = W_w(36.0, 10.0) \quad (\text{A.11})$$

$$t_{H_B, F\text{Abbringer, soll haltezeit}} \leq 14 \wedge t_{A, \text{soll umstieg}} \leq 9 \wedge t_{F\text{Abbringer, resthalte}} \geq 12 \rightarrow \text{Ma\ssnahme} = W_w(36.0, 10.0) \quad (\text{A.11})$$

$$t_{F\text{Abbringer, resthalte}} \geq 12 \rightarrow \text{Ma\ssnahme} = W_w(36.0, 10.0)$$

$$t_{H_B, F\text{Zubringer, ankunft verspaetzung}} \geq 29 \wedge t_{A, \text{soll umstieg}} \leq 37 \rightarrow \text{Ma\ssnahme} = W_w(30.0, 11.0) \quad (\text{A.12})$$

$$t_{F\text{Abbringer, donehalte}} \leq 0 \rightarrow \text{Ma\ssnahme} = G(227.0, 13.0) \quad (\text{A.13})$$

$$t_{H_B, F\text{Abbringer, soll abfahrt}} \geq 2015-06-16 \ 10:19 \rightarrow \text{Ma\ssnahme} = G(115.0, 4.0) \quad (\text{A.14})$$

$$\text{Richtung}_{Z\text{ubringer}} = \text{Dresden Hbf} - \text{Frankfurt (Main) Hbf} \rightarrow \text{Ma\ssnahme} = G(2.0, 0.0) \quad (\text{A.15})$$

$$B_{\text{Abbringer Ziel}} = \text{SGR} \rightarrow \text{Ma\ssnahme} = G(2.0, 0.0) \quad (\text{A.16})$$

$$\rightarrow \text{Ma\ssnahme} = W_{\ddot{u}}(511.0, 175.0) \quad (\text{A.17})$$

A.4 Klassenzusammenstellung \mathbb{K}_4

A.4.1 Featuremenge \mathbb{F}_1

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{H_B, F_{\text{Abbringer}}, \text{abfahrt verspaetung}} \leq 3 \wedge \quad (A.1)$$

$$t_{A, \text{ist umstieg}} \geq 6 \rightarrow \text{Maßnahme} = W_r(128.0, 0.0)$$

$$\text{Richtung}_{Zubringer} = \text{Karlsruhe Hbf} - \text{Kiel Hbf} \rightarrow \text{Maßnahme} = W_r(2.0, 0.0) \quad (A.2)$$

$$t_{A, \text{Regelwartezeit}} \geq 5 \wedge t_{A, \text{ist umstieg}} \geq 5 \wedge t_{A, \text{ist umstieg}} \leq 17 \rightarrow \text{Maßnahme} = W_r(4.0, 0.0) \quad (A.3)$$

$$t_{A, \text{ist umstieg}} \leq 4 \wedge t_{\text{mindestübergangszeit}} \geq 5 \rightarrow \text{Maßnahme} = W_w(264.0, 0.0) \quad (A.4)$$

$$t_{A, \text{ist umstieg}} \leq 3 \wedge t_{\text{mindestübergangszeit}} \geq 4 \rightarrow \text{Maßnahme} = W_w(40.0, 0.0) \quad (A.5)$$

$$t_{A, \text{ist umstieg}} \leq 5 \wedge t_{\text{mindestübergangszeit}} \geq 6 \rightarrow \text{Maßnahme} = W_w(16.0, 0.0) \quad (A.6)$$

$$\rightarrow \text{Maßnahme} = W_{\ddot{u}}(471.0, 9.0) \quad (A.7)$$

A.4.2 Featuremenge \mathbb{F}_2

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{A, \text{soll umstieg}} \geq 13 \wedge \quad (A.1)$$

$$t_{H_B, F_{Zubringer}, \text{ankunft verspaetung}} \leq 11 \wedge$$

$$t_{F_{\text{Abbringer}}, \text{donefahrtdauer}} \leq 111 \rightarrow \text{Maßnahme} = W_r(40.0, 7.0)$$

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{A, \text{soll umstieg}} \geq 14 \wedge$$

$$t_{H_B, F_{Zubringer}, \text{ankunft verspaetung}} \leq 12 \wedge \quad (A.2)$$

$$t_{H_B, F_{Zubringer}, \text{soll haltezeit}} \leq 2 \rightarrow \text{Maßnahme} = W_r(20.0, 4.0)$$

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{A, \text{soll umstieg}} \geq 17 \wedge$$

$$t_{F_{\text{Abbringer}}, \text{resthalte}} \geq 11 \wedge t_{F_{\text{Abbringer}}, \text{restfahrtdauer}} \leq 188 \rightarrow \text{Maßnahme} = W_r(11.0, 0.0) \quad (A.3)$$

$$t_{A, \text{Regelwartezeit}} \geq 3 \wedge t_{A, \text{soll umstieg}} \geq 22 \wedge$$

$$t_{H_B, F_{Zubringer}, \text{ankunft verspaetung}} \leq 25 \wedge t_{A, \text{soll umstieg}} \leq 67 \wedge$$

$$t_{F_{\text{Abbringer}}, \text{donefahrtdauer}} \leq 262 \wedge \quad (A.4)$$

$$t_{H_B, F_{\text{Abbringer}}, \text{soll haltezeit}} \geq 3 \wedge$$

$$t_{H_B, F_{Zubringer}, \text{soll ankunft uhrzeit}} \geq 09:18 \rightarrow \text{Maßnahme} = W_r(20.0, 2.0)$$

$$t_{A, \text{soll umstieg}} \leq 8 \wedge t_{H_B, F_{Zubringer}, \text{ankunft verspaetung}} \geq 8 \wedge$$

$$t_{F_{\text{Abbringer}}, \text{donefahrtdauer}} \leq 136 \rightarrow \text{Maßnahme} = W_w(84.0, 10.0) \quad (A.5)$$

$$t_{A, \text{soll umstieg}} \leq 7 \wedge t_{A, \text{soll umstieg}} \leq 4 \wedge$$

$$t_{F_{Zubringer}, \text{restfahrtdauer}} \leq 46 \rightarrow \text{Maßnahme} = W_w(40.0, 8.0) \quad (A.6)$$

$$t_{H_B, F_{Zubringer}, \text{ankunft verspaetung}} \geq 14 \wedge t_{A, \text{soll umstieg}} \leq 19 \rightarrow \text{Maßnahme} = W_w(67.0, 13.0) \quad (A.7)$$

$$t_{A, \text{soll umstieg}} \leq 9 \wedge t_{H_B, F_{\text{Abbringer}}, \text{soll haltezeit}} \geq 7 \wedge$$

$$t_{H_B, F_{Zubringer}, \text{wartezeit}} \geq 3 \rightarrow \text{Maßnahme} = W_w(26.0, 7.0) \quad (A.8)$$

$$t_{H_B, F_{Zubringer}, \text{ankunft verspaetung}} \geq 10 \wedge t_{A, \text{soll umstieg}} \leq 38 \wedge$$

$$t_{H_B, F_{Zubringer}, \text{ankunft verspaetung}} \geq 24 \rightarrow \text{Maßnahme} = W_w(28.0, 7.0) \quad (A.9)$$

$$t_{A, \text{ soll umstieg}} \leq 12 \wedge t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 5 \wedge \quad (\text{A.10})$$

$$t_{F_{\text{Abbringer}}, \text{ donefahrtdauer}} \leq 130 \wedge t_{F_{\text{Abbringer}}, \text{ resthalte}} \geq 13 \rightarrow \text{Ma\ssnahme} = W_w(30.0, 7.0)$$

$$t_{A, \text{ soll umstieg}} \leq 12 \wedge t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 10 \wedge \quad (\text{A.11})$$

$$t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \leq 11 \rightarrow \text{Ma\ssnahme} = W_w(23.0, 6.0)$$

$$t_{A, \text{ soll umstieg}} \leq 7 \wedge t_{F_{\text{Abbringer}}, \text{ donefahrtdauer}} \leq 31 \wedge \quad (\text{A.12})$$

$$t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 4 \rightarrow \text{Ma\ssnahme} = W_w(9.0, 0.0)$$

$$\text{RIL100} = \text{EDO} \rightarrow \text{Ma\ssnahme} = W_w(8.0, 1.0) \quad (\text{A.13})$$

$$\rightarrow \text{Ma\ssnahme} = W_{\ddot{u}}(519.0, 126.0) \quad (\text{A.14})$$

A.4.3 Featuremenge \mathbb{F}_3

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{A, \text{ soll umstieg}} \geq 13 \wedge \quad (\text{A.1})$$

$$t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \leq 11 \wedge t_{A, \text{ soll umstieg}} \geq 18 \wedge$$

$$t_{F_{\text{Zubringer}}, \text{ donefahrtdauer}} \geq 154 \rightarrow \text{Ma\ssnahme} = W_r(52.0, 14.0)$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{A, \text{ soll umstieg}} \geq 11 \wedge$$

$$t_{F_{\text{Abbringer}}, \text{ restfahrtdauer}} \geq 177 \wedge \quad (\text{A.2})$$

$$t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \leq 20 \wedge$$

$$t_{A, F_{\text{Zubringer}}, \text{ eingangverspaetung}} \geq 10 \rightarrow \text{Ma\ssnahme} = W_r(26.0, 8.0)$$

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \leq 6 \wedge \quad (\text{A.3})$$

$$t_{A, \text{ soll umstieg}} \geq 9 \rightarrow \text{Ma\ssnahme} = W_r(66.0, 28.0)$$

$$t_{A, \text{ soll umstieg}} \leq 7 \wedge t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 7 \wedge \quad (\text{A.4})$$

$$t_{F_{\text{Abbringer}}, \text{ donefahrtdauer}} \leq 113 \rightarrow \text{Ma\ssnahme} = W_w(69.0, 5.0)$$

$$t_{A, \text{ soll umstieg}} \leq 12 \wedge t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 10 \wedge \quad (\text{A.5})$$

$$t_{H_B, F_{\text{Zubringer}}, \text{ soll ankunft uhrzeit}} \leq 17:14 \rightarrow \text{Ma\ssnahme} = W_w(52.0, 9.0)$$

$$t_{A, \text{ soll umstieg}} \leq 5 \wedge t_{F_{\text{Abbringer}}, \text{ donefahrtdauer}} \leq 97 \wedge \quad (\text{A.6})$$

$$t_{H_B, F_{\text{Zubringer}}, \text{ ist haltezeit}} \leq 2 \rightarrow \text{Ma\ssnahme} = W_w(21.0, 3.0)$$

$$t_{A, \text{ soll umstieg}} \leq 22 \wedge t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 10 \wedge \quad (\text{A.7})$$

$$t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 22 \rightarrow \text{Ma\ssnahme} = W_w(24.0, 1.0)$$

$$t_{A, \text{ soll umstieg}} \leq 15 \wedge t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 10 \wedge \quad (\text{A.8})$$

$$\text{RIL100} = \text{HH} \rightarrow \text{Ma\ssnahme} = W_w(14.0, 1.0)$$

$$t_{A, \text{ soll umstieg}} \leq 7 \wedge t_{\text{mindestuebergangszeit}} \geq 6 \wedge \quad (\text{A.9})$$

$$t_{F_{\text{Zubringer}}, \text{ restfahrtdauer}} \leq 117 \rightarrow \text{Ma\ssnahme} = W_w(37.0, 11.0)$$

$$t_{F_{\text{Abbringer}}, \text{ donefahrtdauer}} \leq 68 \wedge$$

$$t_{H_B, F_{\text{Zubringer}}, \text{ ankunft verspaetung}} \geq 27 \wedge t_{A, \text{ soll umstieg}} \leq 35 \rightarrow \text{Ma\ssnahme} = W_w(13.0, 0.0) \quad (\text{A.10})$$

$$\text{RIL100} = \text{EDO} \rightarrow \text{Ma\ssnahme} = W_w(12.0, 4.0) \quad (\text{A.11})$$

$$t_{F_{\text{Abbringer}}, \text{ donefahrtdauer}} \leq 4 \rightarrow \text{Maßnahme} = W_w(9.0, 2.0) \quad (\text{A.12})$$

$$\rightarrow \text{Maßnahme} = W_{ii}(530.0, 149.0) \quad (\text{A.13})$$

A.4.4 Featuremenge \mathbb{F}_4

$$t_{A, \text{ Regelwartezeit}} \geq 3 \wedge t_{A, \text{ soll umstieg}} \geq 13 \wedge t_{H_{B,FZubringer}, \text{ ankunft verspaetung}} \leq 11 \rightarrow \text{Maßnahme} = W_r(133.0, 54.0) \quad (\text{A.1})$$

$$t_{A, \text{ soll umstieg}} \leq 7 \wedge t_{H_{B,FZubringer}, \text{ ankunft verspaetung}} \geq 7 \wedge t_{F_{\text{Abbringer}}, \text{ donefahrtdauer}} \leq 113 \rightarrow \text{Maßnahme} = W_w(69.0, 5.0) \quad (\text{A.2})$$

$$t_{A, \text{ soll umstieg}} \leq 19 \wedge t_{H_{B,FZubringer}, \text{ ankunft verspaetung}} \geq 10 \rightarrow \text{Maßnahme} = W_w(152.0, 51.0) \quad (\text{A.3})$$

$$t_{A, \text{ soll umstieg}} \leq 5 \wedge t_{F_{\text{Abbringer}}, \text{ resthalte}} \geq 7 \wedge t_{H_{B,FZubringer}, \text{ ist haltezeit}} \leq 7 \wedge t_{\text{mindestübergangszeit}} \geq 4 \rightarrow \text{Maßnahme} = W_w(35.0, 7.0) \quad (\text{A.4})$$

$$t_{H_{B,FZubringer}, \text{ ankunft verspaetung}} \geq 29 \wedge t_{A, \text{ soll umstieg}} \leq 38 \rightarrow \text{Maßnahme} = W_w(26.0, 6.0) \quad (\text{A.5})$$

$$\rightarrow \text{Maßnahme} = W_{ii}(510.0, 167.0) \quad (\text{A.6})$$