LARS KJÆR NIELSEN

Rolling Stock Rescheduling in Passenger Railways

Applications in Short-term Planning and in Disruption Management

Rolling Stock Rescheduling in Passenger Railways

Applications in short-term planning and in disruption management

Rolling Stock Rescheduling in Passenger Railways

Applications in short-term planning and in disruption management

Herplanning van materieel voor reizigerstreinen
Met toepassingen in de korte termijn planning en de bijsturing

PhD Thesis

to obtain the doctoral degree from
Erasmus University Rotterdam
by command of the rector magnificus

Prof.dr. H.G. Schmidt

and in accordance with the decision of the Doctoral Board.

The public defense shall be held on
Friday 11 February 2011 at 13.30 o'clock

by

Lars Kjær Nielsen
born in Herning, Denmark

ERASMUS UNIVERSITEIT ROTTERDAM

Doctoral committee:

Promotor:	Prof.dr. L.G. Kroon
Other members:	Prof.dr. A.P.M. Wagelmans
	Prof.dr. A. Caprara
	Dr. D. Huisman
Co-promotor:	Dr. G. Maróti

Erasmus Research Institute of Management – ERIM
Rotterdam School of Management (RSM)
Erasmus School of Economics (ESE)
Erasmus University Rotterdam
Internet: `http://www.erim.eur.nl`

ERIM Electronic Series Portal: `http://hdl.handle.net/1765/1`

ERIM PhD Series in Management, 224
ERIM reference number: EPS-2011-224-LIS
ISBN 978-90-5892-267-0

Design: B&T Ontwerp en advies `www.b-en-t.nl`
Cover Art: Natalie Hanssen
Print: Haveka `www.haveka.nl`

Acknowledgments

The defense of this thesis marks the end of my time as a PhD candidate. I have had four great years in the department and would like to use these first pages to thank some of the people who have impacted my life as a PhD candidate and as an inhabitant in the Netherlands.

First of all, I owe much gratitude to Leo Kroon. I met Leo at a conference in Copenhagen after which I was invited for a job interview at the RSM. Not only has Leo been an outstanding promotor, but he also helped me settle in by providing a room in his house in Zoetermeer during my first two weeks in the Netherlands. Together with Cisca you made the transition period very easy and introduced me to Dutch culture. I think this is why I already after very short time referred to the Netherlands as home.

I would like to thank my promotor Jo van Nunen who tragically passed away in May 2010. I first met Jo during my job interview at the department in June 2006. Leo had informed me that Jo had a very intense and passionate character and encouraged me to "act the same way" during the interview. After meeting Jo I knew right away that I would fit in and have a great time at the department, and fortunately I was hired for the position.

I would like to thank Gábor Maróti. You were initially a colleague with whom I conducted most of my research, but you were later upgraded to co-promotor on my PhD project. During this time you have also become a good friend with whom I could share my fondness of silly and morbid humor as well as interesting discussions on anything work-related or not.

I would like to thank the organizers of the ARRIVAL project which supported my PhD education and I would like to thank the many people I have met through this project for inspiring presentations and conversations within the railway realm. I would also like to thank NS Reizigers for facilitating the opportunity to do research on railway planning and execution in the Netherlands and for putting the expertise of the practitioners at my disposal. Especially, I would like to thank the innovations group at NS for providing a great work environment during my visits there.

I have been to many exciting conferences during my time at the RSM. In particular, I remember enjoyable visits to Washington D.C. 2008, Bonn 2009 and Tromsø 2010. I would like to thank my travel companions during the conferences and events abroad for their company. Daniel and Gábor, I especially remember a certain summer school in Sevilla where we seemed to have less energy in the morning than in the evening.

I have enjoyed four inspiring years among colleagues and friends in the dynamic environment of Department Number One. We had many enjoyable social events and special thanks go to Niels, Bas, Muhammad, Sarita, Evelien, Evsen and Luuk for organizing and participating in the PhD dinners and other activities. Furthermore, I am very happy to delegate the task of paranymph to Sarita and Evsen as you are two of the next PhDs in the department's pipeline.

I would also like to thank my family for their support during my PhD studies. I know the eight-hour drive made it difficult to drop by for coffee. Luckily, discount airlines enabled me to visit my native country frequently so I did not seem so far away after all. Further thanks go to my old friends for staying in touch – or maybe I should rather thank Skype and various social networks for facilitating the contact.

Finally, I would like to thank the Danish community in Rotterdam for the many social activities and for providing a homelike sanctuary in a foreign land. Whenever I was fed up with the Dutch pork-and-potato based cuisine I could turn to the Danish society for the delicious Danish pork-and-potato based cuisine.

Lars Kjær Nielsen

Rotterdam, December 2010

Contents

Chapter 1

Introduction

Throughout Europe, millions of passengers travel daily in thousands of scheduled trains through heavily utilized networks connecting hundreds of stations. The intention of the railway companies is to increase the number of passengers by improving the quality of their services. A major contributor to the quality of the service is the handling of unexpected incidents since the operations of such large systems are inevitably subject to irregularities due to various internal and external factors. In the event of such incidents, quick and effective measures must be taken to restore operations. Determining adequate measures requires the solution of complex combinatorial problems in short time and the communication of the modifications to the affected parties. Currently, sophisticated computerized support for these problems is lacking; the problems are usually solved manually with rather simple automated support only.

The topic of this thesis is the development and evaluation of computerized support for the rescheduling of passenger railway rolling stock in the disruption management process. The developed models have applications in other settings in a passenger railway system as well, in particular, rolling stock rescheduling problems also occur in the short-term planning stage. Here the rolling stock schedules from the earlier planning stages are adapted to the circumstances of specific calendar days, while in the disruption management process the rolling stock is rescheduled to the disrupted situation. The common theme for these two decision making contexts is that they concern *rescheduling* rather than scheduling, i.e. existing schedules are adapted to a new situation rather than planned from scratch.

To test our methods in a real life disruption management setting would require a rather involved integration with the existing real-time information systems. However, the development of such an integration is out of the scope of this thesis. We therefore focus on the development of Operations Research models for rolling stock rescheduling in disruption management, but for the validation of the proposed methods in a real-life setting we turn to the short-term planning context.

1.1 Rescheduling

Since the modeling of rolling stock rescheduling and its applications are the main topics of this thesis we start by giving a formal definition of rescheduling.

> *Rescheduling is the task of adapting an existing schedule to a modified situation.*

Naturally, two questions rise from this definition; first, what is a modified situation? And second, when is rescheduling preferable over pure scheduling?

For the first question we broadly define a modified situation as a change in the system for which the existing schedule was planned; the change renders the existing schedule either infeasible or suboptimal. The modified situation may be characterized by one or more of the following properties: A change in availability of resources or infrastructure, a change in demand, or a change in the system environment.

For the second question we note that there is no strict consensus on when rescheduling is preferred over pure scheduling, but we give a number of informal reasons why a problem should be considered a rescheduling problem rather than a scheduling problem.

- The existing schedule is feasible (or even optimal) for most, but not all, of the modified situation. If the modified situation concerns a geographically or temporally restricted modification we may want to limit our response to the affected parts of the schedule only.

- There is only limited time available for creating a new schedule. If the problem exists in a time-critical environment, we may want to limit our efforts to rescheduling the subset of resources that cause infeasibilities.

- The problem is a subproblem of a larger and more complex problem. If there exist earlier, later or even parallel planning tasks that rely on the planning problem in question, we may want to minimize the implied costs for those other planning tasks.

- There may be costs associated with deviating from the existing schedule. Such costs may be direct operational costs such as fuel or over-time payment, or indirect costs through destabilization of the system or the implied need for more involved planning efforts in parallel planning tasks. Especially the need for communicating updates to many different actors in a time-critical environment may be detrimental to the system performance.

The rescheduling problem studied in this thesis is referred to as the *Rolling Stock Rescheduling Problem* and is the problem of rescheduling an existing rolling stock schedule to a modified situation. In this section we give an overview of the two decision making contexts where it occurs.

1.1.1 Disruption management

The execution of the railway operations involves monitoring the positions and movements of all resources, and reacting to any unexpected deviations from the plans.

Deviations from the operational plans may require rescheduling of some resources. Naturally, some deviations are more serious than others. In practice, there is a distinction between minor and major incidents. Minor incidents are called *disturbances* and major incidents are called *disruptions*. The distinction is purely practical and not well defined, but depends on the impact on the operations. As a rule of thumb, disruptions require significant changes of the pre-set resource schedules.

A disturbance is an event that causes part of the railway operations to deviate from the operational plans. For example, at a station the boarding of passengers may take longer than planned for a certain train leading to a delay in departure. Disturbances are usually absorbed by the slack in the system or can be handled by small changes to the resource allocation.

More serious incidents are known as disruptions and these incur major deviations from the planned operations. Disruptions may be caused by various internal or external factors such as a faulty switch on a busy track, broken down rolling stock, or damaged overhead wires. Disruptions generally cause serious timetable changes where trains are canceled or rerouted. The timetable changes invalidate the planned resource schedules so they will have to be updated to take the actual situation into account.

Disruption management refers to the process of dynamically revising an operational plan to obtain a new one that reflects the objectives and constraints of the actual situation (see Yu and Qi (2004)). In the railway context, disruption management refers to the task of adapting the three main entities, the *timetable*, the *rolling stock circulation* and the *crew schedules*, to the actual situation. When the railway infrastructure malfunctions, rolling stock breaks down or crews are unavailable, it may prevent the operational plan from being carried out as planned – in such a case the situation is disrupted and the disruption management process begins. This process is carried out by dispatchers at operations control centers and the process follows three steps. Each step is related to updating one of the three mentioned entities.

In the first step, the timetable is updated by canceling or changing trains. Depending on the nature of the disruption, dispatchers decide which services are affected by the disruption and update the timetable accordingly. The updates to the timetable usually follow a predetermined handling scenario.

In the second step, the rolling stock schedules are adapted to serve the updated timetable. The rolling stock circulation has to be updated to conform to the actual situation while fulfilling a number of operational constraints. These constraints include upper bounds on train length, compatibility of rolling stock with the route and complex shunting constraints while attempting to allocate the seat capacity in accordance with the demand, which is also affected by the disruption. It is generally a goal in this process to minimize the changes from the existing plans. Changes to the rolling stock circulation may require coordination with the local traffic controllers of the infrastructure managers to ensure that proposed shunting operations and other local issues are possible. In other words, the Rolling Stock Rescheduling Problem is applied in the second step in the disruption management process.

In the third step, the crew schedules are updated so that the correct number of qualified drivers and conductors are assigned to each train in the updated timetable. The Crew Rescheduling Problem is subject to a number of constraints as crews are entitled to breaks, there are limitations on the allowed length of duties and crews have to get back to their crew bases at the end of their duties. The proposed changes to the crew schedules are negotiated with the affected crews to ensure they can be realized in practice.

The three steps are interdependent in the sense that if it is not possible to assign rolling stock or crew in the second or third step to a train service, the timetable from the first step may have to be revised again. Furthermore, the process takes place in a dynamic environment that continuously affects the available options in the process. Consequently, integrated rescheduling may be useful, but is still considered too complex. Rolling stock rescheduling in the disruption management process is the topic of Chapters 4 and 5.

1.1.2 Short-term planning

Short-term planning refers to the period from around two months to two weeks before operations. In this planning stage the results of the tactical planning stage are transferred to the short-term planning department. The task is to adapt the rolling stock schedules to specific calendar days. Often, the adaptation requires only limited planning effort, but in case of large infrastructure maintenance projects or holidays the situation deviates significantly from the generic days. In such cases a more involved planning effort is needed.

The main difference between rolling stock rescheduling in short-term planning and in disruption management is the absence of uncertainty and the fact that short-term planning is much less time-critical.

Short-term planning is discussed in further detail in Section 2.2, and rolling stock rescheduling in short-term planning is the topic of Chapter 6.

1.2 Motivation for the research

The planning and operation of a busy passenger railway network is a complex task. Rolling stock and crew have to be scheduled to serve the timetable with ever growing demand for capacity. This has led to extensive research on optimizing the utilization of railway resources such as the infrastructure, the rolling stock and the crew (see Caprara et al. (2007b)). The developed methods have resulted in resource schedules that are highly efficient when operations run as planned (see Kroon et al. (2009)).

The success of OR methods in the early stages of the passenger railway planning process combined with the increasing computational power of modern computers motivates the development and application of OR methods in the real time control to assist the disruption management process. The challenges in real time decision making differ from the static planning tasks in a number of aspects: The decisions must be made within a tight time frame and a highly optimized plan already exists and any changes to that plan will have to be communicated to the involved parties (see Séguin et al. (1997), Grötschel et al. (2001)).

Another important characteristic of the problems investigated in this thesis is the underlying uncertainty of disruptions. In the event of a disruption the availability of resources is uncertain, for example it may not be known how long infrastructure is blocked and how long broken down rolling stock is out of service. We seek to contribute to the understanding of the underlying uncertainty in the disruption management process and develop methods for incorporating the uncertainty into real time decision support systems.

1.3 Research questions

As mentioned in the introduction, the central research question is:

> *How can rolling stock rescheduling be modeled and how can the models be applied in the decision making process in a passenger railway system?*

The main question naturally poses a number of sub-questions. First, to which problems can rolling stock rescheduling be applied:

1. *Which specific problems in the planning and operations of a passenger railway system can be characterized as rolling stock rescheduling problems?*

There is a number of goals in passenger railway rolling stock rescheduling. These goals are related to various aspects of the process and include providing satisfactory service while minimizing the changes to the existing plans. The second question is thus:

2. *Which aspects characterize the problems in rolling stock rescheduling in terms of managerial goals?*

Identifying the specific problems in rescheduling naturally leads to the question of modeling:

3. *How can the specific problems that arise in the rolling stock rescheduling context be modeled and solved?*

Furthermore, the managerial goals have to be taken into account in the models:

4. *How can the trade-off between the different aspects of the managerial goals be modeled?*

For the application of rescheduling models in disruption management, uncertainty plays a central role as a major complicating factor. The fifth question is thus:

5. *How can the inherent uncertainty of disruptions be incorporated in the rescheduling process?*

We also investigate how the specific rolling stock rescheduling problems relate to and depend on the passengers.

6. *How can we improve the service quality for the passengers in the rescheduling process?*

The six sub-questions above cannot be answered separately as they relate to different aspects of the main research question. Therefore, we seek to incorporate all aspects of the questions in each problem investigated. In the following section we discuss the methodology used to answer the research questions in this thesis.

1.4 Methodology

The research approach is model based and the main steps in the research process are:

1. Describe a framework for rolling stock rescheduling.

2. Identify relevant subproblems in the decision processes. Discussions with planners and dispatchers deliver insight into the challenges they are faced with when delays and disruptions occur.

3. Develop mathematical models describing the problems. Investigate possible solution methods to solve real life size instances to the models. Both exact and heuristic methods are investigated.

4. Evaluate the models and solution methods. The quality of the proposed models and solution methods should be evaluated in experimental settings. The evaluation methods involve both consultation with practitioners and extensive tests on real instances.

Several problems in the rescheduling context have corresponding static planning problems. However they differ in some crucial aspects; the objectives are very different in the early planning phase and the short-term/real time phase, and the time available to solve the problems is much longer in the early planning phase. Also, the presence of uncertainty offers further challenges for the real time problems. We investigate if it is possible to adapt methods from early planning to the real time context.

1.5 Scientific contribution

A crucial difference between decision support in static planning and in a real time situation is the fact that the latter takes place in a dynamic environment. This means that resources have to be monitored and conflicts must be detected before available countermeasures can be established. Less time is available to determine and evaluate effective solutions than in static planning situations. The challenge for the algorithmic techniques is to deliver effective solutions in the shortest possible time while taking the effects of the dynamic situation into account. For this, major steps forward in algorithmic techniques are required for providing the necessary effectiveness.

The available literature on algorithmic support for static planning problems in railway systems is extensive. However, disruption management for railway systems is still a fairly uncharted field in the literature. More research has been conducted on disruption management in other time-critical logistic systems such as the airline industry, but the problems to be solved there are usually considerably smaller and are of a different nature than in railway systems, see Clausen et al. (2010), Ball et al. (2007) and Yu and Qi (2004) for overviews of disruption management in the airline industry.

This thesis contributes to the literature by investigating various approaches to solving large scale combinatorial rescheduling problems in a dynamic and time-critical environment. It also contributes to the understanding of the underlying uncertainty of disruptions. Although the focus is on railway systems, the results may very well be applicable to other time-critical logistic systems as well.

1.6 Managerial contribution

The quality of the operations of a passenger railway company is measured by a number of performance indicators. Some of the more important ones are seat capacity, traveling times, connection times, number of train changes on traveling routes, punctuality and reliability; Vromans (2005) provides an elaborate discussion of railway performance indicators. Punctuality is the percentage of trains arriving within a given margin of the scheduled arrival time and reliability refers to the number of canceled departures. Traveling time, connections and seat capacity are indicators of the quality of the planned operations whereas punctuality and reliability are indicators of the quality of the execution.

Railway companies have to fulfill their agreements with the government with respect to the quality of their operations. If the punctuality is too low, a company may have to pay certain penalties to the government. Thus increasing the performance of the railway system results in a monetary benefit for the company. More importantly, increasing the quality of the service has positive market effects in the form of higher customer satisfaction and more passengers.

The research conducted in this thesis contributes to developing decision support tools for disruption management, which will help planners and dispatchers deal with disruptions effectively. A better disruption management process improves the quality of the operations which in turn helps the railway company meet the quality goals. In particular, a decision support system for disruption management may increase punctuality and reliability, and ensures that seat capacity better meets demand during irregular operations.

As a more concrete managerial contribution we show how the application of Operations Research models for rolling stock rescheduling has helped streamline the short-term planning process at the largest railway operator in the Netherlands. The process of constructing a candidate rolling stock schedule for a specific setting is severely reduced compared to manual planning, and this reduction in through-put time allows the decision maker to consider multiple solutions before implementing one of them.

1.7 Societal contribution

Increasing the quality of railway services will attract additional passengers to the railways. Expanding the market share of public transportation, and railway transportation in particular, in the Dutch mobility market is a major goal for the Dutch government. In fact, easy access to the major cities is crucial for the Dutch society as millions of workers commute on a daily basis by various means of transportation. Also, attracting commuters from cars to the passenger railways has positive environmental benefits in the form of lowering CO_2 emissions (see Potter (2003) and Shapiro et al. (2002)).

The quality goals set by the government for the railway company intend to improve the customer satisfaction. Obviously, the first party that benefits from improvements in disruption management consists of the passengers: Better handling of disruptions leads to better service in the form of less delays and less canceled trains. Moreover, a more sophisticated disruption management process may lead to better information to passengers in case of disruptions. The latter is currently a major bottleneck in the railway system. The findings in this thesis contribute to attaining these objectives.

1.8 Structure of the thesis

In **Chapter 2** we describe the planning and operational processes in a passenger railway company. This includes the roles and responsibilities of the involved parties and the locality of information. Furthermore, we elaborate on the managerial goals in the different stages of the planning of the railway system. The chapter also contains a discussion of the terminology and a survey of the relevant scientific literature.

Chapter 3 is dedicated to the description of Operations Research models for the Rolling Stock Rescheduling Problem. The problem is divided into two subproblems; circulation generation, and duty generation. Two models are presented for the circulation generation step in this chapter. The first model is an extension of the model described in Fioole et al. (2006) whereas the second model is based on a constrained multi-commodity flow through the tasks of the rolling stock units and on sequences of tasks that may be assigned to the same rolling stock unit. For the duty generation step we use a MIP model to create the duty of each rolling stock unit from the rescheduled circulation in such a way that it minimizes the changes from the planned duties. Furthermore, we investigate the possibility of aggregating the circulation and duty generation steps into one model. Finally, we conclude the chapter with a comparison of the applicability of the involved models.

In **Chapter 4** we study the uncertainty in disruption management. The first part of of the chapter describes a framework for disruption management in passenger railway transportation. The core of the framework is the detection and tracking of deviations from the planned operations. Such deviations may lead to conflicts in the resource assignment that have to be solved. We discuss different ways of representing the underlying uncertainty in the process and how uncertainty affects the proposed solution methods. The second part of the chapter contains a method for disruption management of rolling stock where the problem is modeled as an online combinatorial decision problem. This part is based on the paper:

> Nielsen, L. K., Kroon, L., Maróti, G., *A Rolling Horizon Approach for Disruption Management of Railway Rolling Stock*, Technical report, ARRIVAL Project, 2009. Submitted to *European Journal of Operational Research*.

An early version of this paper won the 2008 INFORMS Railway Application Section Student Paper Award, and a revised version won the 2009 EURO Management Science Strategic Innovation Prize.

The models introduced in the previous chapters are based on the following assumption on passenger behavior: the premises for the demand for capacity either remain unchanged in a disrupted situation or the modified demand can be forecasted adequately and independently of the capacity allocation. This means that creating a new plan that is close to the original plan is considered as a major goal in the process. However, passengers respond to the changes in the system by finding alternative routes to their destination or by leaving the system if estimated traveling times become too long. In **Chapter 5** we introduce a mathematical model that takes the passenger response to the rescheduling into account. The model is based on an iterative method that uses three steps in each iteration. The first step creates a rolling stock schedule based on the current passenger demand. The second step simulates the passenger response to the timetable and rolling stock schedule, and the third step interprets the passenger response and feeds back penalties on capacity shortages to the first step in the next iteration. We finalize the chapter by discussing the practical implications of the model, and the availability of data for the model.

The models for disruption management of rolling stock developed throughout this thesis are thoroughly tested in a simulated environment. However, testing them in practice would require a sophisticated interface with the existing operational control system. The development and implementation of such interfaces is beyond the scope of this thesis. Instead we had the unique opportunity to apply the developed rolling stock rescheduling models to short-term planning cases at Netherlands Railways, the largest railway operator in the Netherlands. We describe in **Chapter 6** two case studies; one involving a period with infrastructure maintenance, and one involving the summer holiday schedule of 2009. The situation in the infrastructure maintenance case study is to some extent similar to a disrupted situation as an existing rolling stock schedule has to be adapted to a situation with further restrictions on the rolling stock assignment. We discuss the methods used in the case studies and the results achieved. Furthermore, we describe the challenges of working with non-OR practitioners and the interfaces between different data systems.

In **Chapter 7** we draw conclusions from the studies in the thesis, and discuss the results and future work. In particular, we analyze the practical implications of the findings and the challenges for adopting the methods in a real world setting.

The thesis contains three appendices. Appendix A contains a glossary where common railway terminology is explained, Appendix B contains a list of notations used throughout the thesis, and Appendix C contains some additional computational results for the computational framework presented in Chapter 4.

Chapter 2

Passenger Railway Operations

Running a passenger railway system is a complex process where several parties have different responsibilities. We start this chapter with a short discussion of the roles of the organizations involved in operating a passenger railway system. The discussion is based on the situation in the Netherlands. Next, we describe how the planning process is conducted at a passenger railway company. We then discuss how the real time control and disruption management process is organized in practice, and what this implies for the development of decision support tools for the involved problems. We finalize the chapter with a review of relevant literature on the planning and real time control phase.

2.1 Organizations

Operating a railway system involves several actors, each responsible for their own part of the process. Since the liberalization of the European railway market in the 1990s, the task of managing the railway traffic has been split between the *infrastructure manager* and the *operators*. The infrastructure manager is a governmental organization that is responsible for the maintenance and utilization of the infrastructure. The operators are the railway companies that operate the train services according to certain contracts with the government. We give here an overview of the organizations involved in the railway operations and their incentives.

The government

The government tenders the rights for operating train services on each part of the network. The contracts last several years and set the goals for the quality of the railway services. This way the government influences the development of the mobility market by requiring certain service levels on the tendered lines. Also, the limits for ticket prices are set by the government as part of the contract. Again, this is a tool for controlling the market share of public transportation.

As mentioned in Section 1.6, the operators are contractually bound to offer services with a certain frequency and seat probability, and if the realized punctuality is below a certain threshold, it incurs a fine. If the performance of an operator is consistently poor, it will be taken into account at the next tendering for the operating rights of the network.

The goal of the government is to ensure high mobility within the country through reliable access to the major cities, and passenger railway transportation plays a major role in achieving this goal.

The Netherlands is densely populated and commuters frequently experience congested highways. Of course, congestions can be alleviated by extending the highways and increasing the capacity of the access roads to the major cities. However, in recent years the focus has shifted to require not only reliable mobility, but also sustainable mobility. Public transportation, and railways in particular, is a crucial means for reducing CO_2 emissions (see Potter (2003) and Shapiro et al. (2002)).

The infrastructure manager

The infrastructure refers to the static elements of the railway network such as rails, switches, safety systems and power supply systems. The infrastructure manager is responsible for maintaining the existing infrastructure as well as administering the construction of new infrastructure. Furthermore, the infrastructure manager coordinates the division of capacity between the operators. This is performed on a yearly basis during the construction of the national railway timetable.

The infrastructure manager also conducts traffic control. This includes two tasks. First, in case of small disturbances the infrastructure manager decides, in cooperation with the operators, the order in which trains enter railway routes or stations. Second, in case of larger disruptions the infrastructure manager negotiates new time slots for the train services that are affected by the disruption. This includes trains that are delayed due to the disruption and trains that are rerouted due to infrastructure blockage, as such trains may conflict with other planned services.

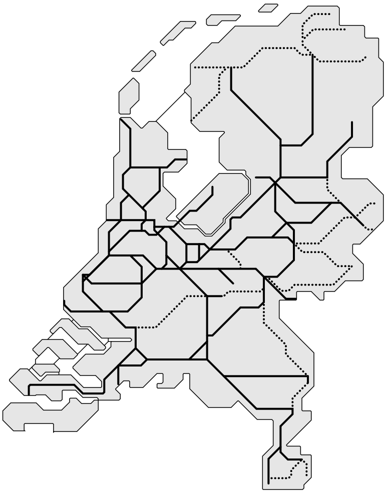


Figure 2.1: The Dutch railway network. The solid lines are operated by NS. The dotted lines are operated by other railway operators.

Before the liberalization of the railway market the railway infrastructure in the Netherlands was managed by the same company that operated the trains. Now however, the infrastructure manager is a separate company, named *ProRail*, that is a governmental organization.

Operators

An operator is a company that performs train services on part of the network. The operator is responsible for acquiring and maintaining rolling stock as well as hiring and training the necessary crew to run the trains.

Currently, a number of passenger railway companies operate on the Dutch railway network. NS is the largest operator currently carrying 1.1 million passengers daily; other operators are *Arriva*, *Syntus*, *Connexxion* and *Veolia*. Until 2015, NS has the exclusive right to run railway services on the core part of the Dutch railway network. Figure 2.1 illustrates the Dutch railway network where the lines operated by NS and other operators are shown.

In addition to the passenger operators, a number of cargo railway operators utilize the railway network. These include *DB Schenker*, *ERS Railways*, *Rail4chem* and *Veolia Cargo*. Although there is a part of the network dedicated to cargo rail transport, the cargo operators are allowed to operate on part of the network and thus compete with the passenger operators for infrastructure capacity.

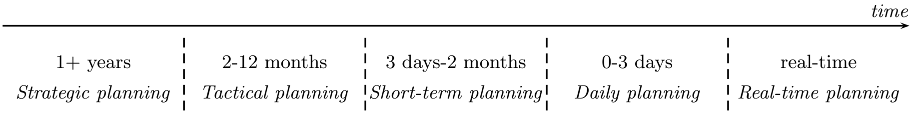


Figure 2.2: Railway planning phases and their time horizons.

The planning processes and the execution of the resulting plans in practice is the topic of the rest of this chapter. We therefore elaborate on the planning and operational processes of a passenger railway operator in Sections 2.2, 2.3 and 2.4.

2.2 Planning railway operations

The planning of the railway operations by an operator primarily concerns the timetable and the two main resources; the rolling stock and the crew. The planning of these resources undergoes several phases before the actual operations. In this section we give a short description of the problems in each phase and the implications for the final operations.

The planning phases can be classified by the time horizon of the involved decisions. Huisman et al. (2005) divide the planning process into four steps depending on the horizon of the decisions. We have refined the division into five phases, the first four of which are the traditional phases, *strategic*, *tactical*, *short-term* and *daily* planning. In addition to these four phases we define a fifth phase that concerns the replanning that takes place in real time during the actual operations. Figure 2.2 shows a time-line with the planning phases and their time horizons.

The reason for dividing the planning process into several planning phases in such a way is primarily organizational. The planning process is pipelined such that the output of one planning phase serves as input for the next planning phase. Also, there is the question of locality of information. In the early phases, the planning is highly centralized but as the time of operation draws closer the planning tasks are conducted by more localized organizations.

Strategic planning

The horizon of the strategic planning is several years and includes defining the overall objectives of the operator, purchasing and disassembling of rolling stock, hiring and training new crew, the basic structure of the timetable, and decisions on line planning.

For the rolling stock, strategic decisions are made on purchasing or leasing new rolling stock as well as refurbishing the existing rolling stock. These decisions have a very long horizon as rolling stock usually is in operation for decades. Also, it may take several years before newly ordered rolling stock is available. The strategic decisions on rolling stock naturally have implications for the daily operations as different rolling stock types have different characteristics. For example, double deck trains have a smaller operational cost per seat than single deck trains, but are less flexible when matching capacity to demand. Other characteristics such as acceleration, maximum speed and braking capabilities influence the railway performance.

A further strategic decision for the rolling stock is the maintenance strategy. This includes the locations and capacities of maintenance facilities and decisions on how often the rolling stock is maintained. A maintenance strategy dictates after how much usage or how much time certain spare parts are replaced and how often preventive maintenance is performed. The maintenance strategy thus balances the reliability of the rolling stock and the maintenance costs.

Hiring, training (or firing) crew are decisions with a horizon of several years. Even changing the conditions and labor rules of personnel has a long horizon. In particular training a train driver takes about two years, so the decisions take years to have an effect. At the same time, laying off personnel is an action that is subject to labor union agreements and therefore cannot be performed within a short horizon. Further, the positions and sizes of crew depots are decided in the strategic phase. The strategic crew decisions influence the availability of personnel in the daily operations, especially having only little extra crew decreases the flexibility in case of illness.

Line planning is the process of determining the train lines. This includes decisions on frequencies and stopping patterns of the involved train services. Line planning is based on estimated passenger demands between pairs of stations. At this early point in the planning process the passenger data is naturally rather uncertain. Therefore, the variation between peak and off-peak hours is not explicitly considered in the line planning phase. The fact that the line plans are based on highly aggregated passenger data and are fixed already in the strategic phase makes the railway system somewhat rigid with respect to the individual passengers' demand. The structure of the Dutch network allows for several traveling routes between most pairs of stations. This means that passengers often can choose a different route in case of disruptions.

Tactical planning

The planning horizon in tactical planning is two months up to a year. The planning steps conducted during this planning phase include constructing a generic one-week timetable that satisfies service demands, and allocating rolling stock and crew to the generic timetable.

The generic timetable sets the departure times of all train services for a generic week that has no special requirements. This means that eventual infrastructure maintenance projects are not taken into account in this phase – only the minor regular maintenance is accounted for. Neither are any events such as public holidays where passenger demand and traveling patterns significantly differ from normal weeks. The generic timetable is cyclical and is created by first generating a basic hourly pattern. The hourly plan is then copied to fill entire days before finally removing a number of services in off peak hours.

At the local level, plans are constructed for movements of trains inside the railway nodes according to the generic plans. These local plans primarily serve as a feasibility check for the generic timetable and rolling stock circulation.

An essential step in the tactical planning phase is the distribution of the rolling stock among the lines. Based on estimated passenger demand, the rolling stock is distributed between the lines so that on a given day each line is preferably served by its own set of rolling stock. The rolling stock is then assigned to each train service in the generic timetable. This detailed assignment also implies the need for shunting movements of rolling stock inside railway nodes. These movements are planned for the generic week in this phase as well. The existence of the generic resource schedules means that any later planning for specific days, that differ from the corresponding day in the generic week, has to take the generic schedules into account.

Short-term planning

The short-term planning phase refers to planning tasks with a time horizon of a few days up to two months. In this phase the generic week plan is adapted to the demands of the individual weeks or days.

For a particular week, infrastructure maintenance may require adaptations of the timetable. Departure and arrival times are adjusted and train services may be inserted or canceled depending on the available infrastructure. Such adaptations to the timetable also imply changes to the resource schedules. The resources are replanned while taking the structure of the generic plans into account. This is done to perturb the according shunting and cleaning schedules as little as possible.

National holidays and events that attract a lot of people, such as major sports events and Queen's day, generally require a different distribution of the seat capacity. The generic resource schedules are adapted to take this into account. The planning of the resources up to and including short-term planning is conducted at planning organizations, but after this phase the plans are sent to all parties that are going to implement them.

Daily planning

The last planning phase before the actual operations has a horizon of up to just a few days. We call this phase *daily planning* since it deals with issues that arise on a daily basis very close to operations. In this phase the plans from short-term planning are transferred to the relevant local controllers and dispatchers. In this transition some local issues may arise that require minor adaptations. This can be due to temporary unavailability of staff because of illness, or due to unexpected limitations in shunting capacity or in rolling stock availability. Most of these conflicts can be handled locally by exchanging duties between staff or rolling stock, and thus require little or no global coordination.

An important issue in this phase is the preventive *maintenance* of rolling stock. Rolling stock must undergo regular maintenance checks whenever a certain number of kilometers has been driven. The number of kilometers driven is tracked for each rolling stock unit so that the need for maintenance is detected well before the unit reaches the kilometer threshold. Once a unit needs its regular maintenance check, its duty is exchanged with the duty of a unit that ends at an appropriate maintenance facility. The need for regular preventive maintenance implies that there must be sufficient possibilities for duty exchanges in the rolling stock duties for the units to reach the maintenance facility within a few days. We note that at some railway operators the scheduling of preventive maintenance is taken into account already in tactical planning depending on the specific policy of the operator.

Real time planning

This phase concerns the replanning of resources during the actual operations. If an unexpected event renders the operational plans impossible to perform in practice, the affected resources have to be rescheduled. The process of monitoring the timetable and the resources, detecting the conflicts, and implementing proper reactions is described in the next section.

2.3 Passenger railway operations

In this section we first describe the process of operating a passenger railway system. We then elaborate on the actors involved in performing the railway operations and their roles during operations and in the disruption management process. The description in this section refers to the situation in the Netherlands. The situation at other operators may differ in the division between infrastructure management and train operation and in the exact division between global, regional and local responsibilities. For another discussion of the operational processes and the involved actors we refer to Jespersen-Groth et al. (2010). The authors describe the situation in the Netherlands as well as at DSB S-Tog A/S – a regional operator in the greater Copenhagen area in Denmark.

2.3.1 The operational process

For airline operations, Kohl et al. (2007) describe the disruption management process by Figure 2.3. The process is similar to that of a passenger railway system. We call this the *operational process* since the core of the process, namely the monitoring of the resources, continuously takes place also when no disruption has occurred. The process involves monitoring the resources and evaluating possible conflicts whenever the process deviates from the planned operations. If a conflict is indeed detected, an iterative process of identifying and evaluating possible options is performed. Finally, an appropriate decision is reached and the measures are implemented. In this section we elaborate on the different steps in the operational process, and in Section 2.3.2 we describe how the responsibilities are split between the actors in the process.

Monitoring the operations

When the operational plans are carried out in practice, they are monitored closely. This means that departure and arrival times of trains at stations and key infrastructure points are recorded as are the positions of rolling stock and the assignment of crew. Further, the performance of infrastructure components is monitored carefully to ensure the infrastructure provides the expected capacity. Any deviation between the planned and the actual positions of a resource may cause a conflict in the resource assignment. The first step when a deviation is detected is, therefore, to determine whether this results in a conflict. Predicting upcoming conflicts requires propagating delays and resource availabilities into the future. Also, the passenger flows are monitored to determine whether station and train capacities are sufficient.

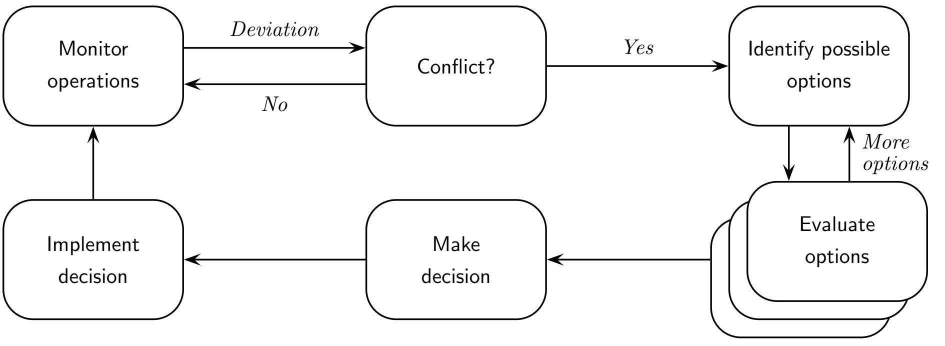


Figure 2.3: High-level view of the operational process based on Kohl et al. (2007)

A conflict can for example be a delayed train that blocks the arrival platform of another train in a station, or a delayed driver that is needed for driving the next train in his duty. More serious conflicts result from failing infrastructure and malfunctioning rolling stock. When a conflict is detected the process of identifying possible options begins.

Deciding countermeasures

In case of a conflict in the timetable or in the assignment of resources the dispatchers must decide how to react to the conflict. In case of small deviations it may be relatively simple to assess the possible options, but in case of a more involved disruption there are often numerous possible options with different impacts on the system performance. A further complicated part of the decision process is to estimate the duration of a conflict.

Possible options primarily involve canceling trains, changing the routes of trains between stations and inside stations, changing the timetable according to (expected) delays, cancelling passenger connections, changing the order of trains, and changing the assignment of resources to trains. Also, unusual passenger flows may require extra stops of some trains to relieve the pressure on others.

For conflicts affecting several timetabled trains and the involved resources the reaction from the dispatchers usually follows a predetermined handling scenario. The scenario describes a modified steady state for the system. In such a case it is still an important decision which scenario to use, and how to handle the transition from the originally planned situation to the disrupted steady state and later back to normal again. In Section 2.6 we elaborate on the possible options for handling conflicts in the railway resources with a particular focus on the rolling stock.

Implementing decisions

Once the overall framework of how to react to a conflict has been outlined, the decisions are implemented in practice. The practical implementation consists of two major tasks: Updating the appropriate information systems and communicating the changes to the proper parties.

The changes in expected arrival times of trains are automatically recorded in the information systems, while changes to the assignment of resources are entered into the information systems manually. Also, rerouted and canceled trains need to be recorded as well as the routes of trains inside stations. Updating the information systems allows the dispatchers to query the system for details on how the resources are utilized in the adapted situation. Thereby the new decisions are monitored rather than the originally planned operations.

In practice the timetable is updated before the resource schedules since the timetable forms the basis for the tasks to be performed by the resources. A further complicating factor is that different organizations perform the updates – the infrastructure manager updates the timetable, whereas the resource schedules are updated by the operator.

A crucial part of the implementation is to communicate the changes to the proper parties. Especially, the crew needs to know any changes to their duties, and shunting and cleaning personnel should be informed about changes in the rolling stock duties.

2.3.2 Actors in the operational process

The responsibility of monitoring the operations and reacting to any conflicts is divided between the operator and the infrastructure manager. Currently, NS is in a transition phase where the operational processes are becoming more centralized. The description of the actors in this section refers to the situation when the transition is complete.

Within the parties the responsibilities are divided between the network level and the regional level. Figure 2.4 shows the actors in the process and the communication between them. In the diagram the actors from the operator and the infrastructure manager on the network level are depicted in the *Operations Control Center Rail*

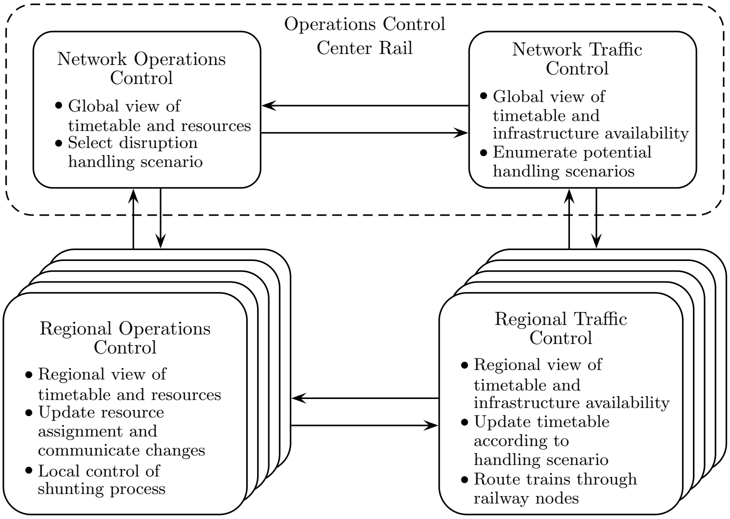


Figure 2.4: Schematic view of the actors involved in the operational process and the communication between them. As indicated by the dashed box the NOCC and the NTCC are physically located close to each other in the OCCR to ease the communication between them.

(OCCR). There they are located physically close to each other to ease the communication between them. The two main actors of the operator are:

- On the network level, the *Network Operations Control Center* (NOCC) monitors the operations. In particular, they trace the delays of trains and the maintenance requirements of the rolling stock. The NOCC decides the frame of how to react to disruptions by choosing which disruption handling scenario to apply. The latter is performed in close cooperation with the NTCC (see below).

- On the regional level, five *Regional Operations Control Centers* (ROCC) monitor the assignment of rolling stock and crew, and update the operational plans inside their region for these resources in cooperation with the NOCC. The ROCC also updates the shunting plans for the stations within its region.

The actors of the operator perform their tasks in close cooperation with the infrastructure manager who is represented in the process by the following actors:

- On the network level, the *Network Traffic Control Center* (NTCC) monitors the train traffic between stations and coordinates the track allocation between the operators. In case of a disruption they determine possible disruption handling scenarios and they select the disruption handling scenario to be used together with the NOCC.

- On the regional level, five *Regional Traffic Control Centers* (RTCC) update the timetable according to the disruption handling scenario, monitor the traffic inside the stations in their regions, and allocate tracks and time slots to trains at those stations.

At the center of the process lie the NOCC and the NTCC combined in the OCCR. Here, dispatchers from the operator evaluate any conflicts in the timetable and monitor the global allocation of rolling stock, while the dispatchers from the infrastructure manager monitor the performance of the infrastructure components. In case of a disruption, the NOCC together with the NTCC decide which disruption handling scenario to use and delegate the task of implementing the scenario to the appropriate ROCCs and RTCCs. If a conflict incurs changes in departure or arrival times of trains or a different route in the network, or even the need for canceling trains, the suggestion will have to be negotiated with the dispatchers of the NTCC and the RTCCs. The latter assign time slots to the trains in the network to facilitate the traffic.

The dispatchers of a ROCC monitor the assignment of crew and resolve any conflicts in the crew schedules. Furthermore, they work out the details of the rolling stock assignment to trains including how and when rolling stock is retrieved from or sent to storage facilities and assign the local personnel to perform these tasks. These tasks are performed in coordination with the RTCC since train movements require access to the available infrastructure. The dispatchers of a RTCC are responsible for routing trains through railway nodes and for assigning platforms to arriving trains.

With respect to the rolling stock and the timetable, the tasks in the process diagram in Figure 2.3 are split between the network and regional level as follows. The NOCC performs the monitoring, conflict detection, identification and evaluation of possible options, and the decisions on how to react. Their monitoring includes the tracing of the maintenance requirements of the rolling stock as well as the positions of units at the end of the day. The ROCCs are the primary actors for implementing the decisions and updating the information systems. The rolling stock and timetable are monitored globally rather than regionally since conflicts and updates in one region may have local effects in other regions. Furthermore, the combination of the timetable and rolling stock determine the service level offered to the passengers. The crew schedules and the shunting operations are monitored and updated regionally.

As indicated earlier, NS is currently in a transition phase where the operational processes are becoming more centralized. The responsibilities of the NOCC were split between several decentralized instances, and the local operations were managed by thirteen smaller centers located at the major hubs in the network. Recently, the decentralized network control was moved to one centralized unit – the current NOCC. Similarly, the local operations centers were integrated into the current five ROCCs where each is responsible for the major railway nodes in its region. The more centralized structure is expected to lead to less communication between different actors when evaluating potential decisions in the operational process. A similar centralizing process is currently being carried out at the traffic control centers of ProRail.

2.4 Perspectives of the operator goals

The goal of the railway operator is, simply put, to increase its competitiveness and make money. This long-term goal involves a number of strategies for the different parts of the company. One of them concerns the execution of the daily operations. We here discuss the goals for the operational process.

The overall managerial goals for the operational process consist of several aspects. We have divided the goals into four categories depending on the perspective: From the customers' perspective it is desirable to offer good service both in the planned operations and in the execution. The efficiency of the system emphasizes meeting the contractual target at minimum cost. From the perspective of the railway system itself, it is attractive to emphasize the robustness of the system so that deviations from the plans are either absorbed or at least handled in a way that minimizes the propagation of the effects of the deviations. Finally, for the employees involved in the operational processes it is desirable to deviate as little as possible from the operations planned earlier.

Service perspective

The customers are interested in high service levels in both the planned railway operations and during the actual execution of the plans.

The service in the planned operations involves the characteristics of the train services, i.e. the frequency of departures and the traveling times, but also the environment in which the train services exist affect the customers' perception of the service. It is perceived as good service if the train services are well connected with other modes of transportation such as public transport options like trams and buses, but also private transportation in the form of parking facilities. Further aspects concern whether the stations have certain facilities available, such as a kiosk and proper waiting facilities. These issues are taken into account in the strategic planning.

The perception of the service of the actual operations is affected by any unexpected inconvenience, such as whether the passenger reaches his destination on time, whether the trains are crowded, and whether the passengers are given appropriate and adequate information on the development of the situation during irregular operations.

The service perspective is considered already in the strategic planning phase when the line plan is constructed and when the passenger demands are estimated. The service perspective is also important in the short-term planning phase where the generic resource schedules are adapted to account for extraordinary capacity demands.

Efficiency perspectives

The railway operators usually set their operational goals to meet the contractual targets for each of the performance indicators at minimum operational cost. An important means in achieving this is to optimize the resource schedules. First, the timetable is constructed in such a way that it meets the contractual service requirements for traveling times, connections and frequencies. Second, the rolling stock is scheduled to meet the required seat demand while minimizing the direct cost associated with the rolling stock. Third, the crew is scheduled to operate the rolling stock. Each train needs a driver and a number of conductors depending on the length of the train. The crew schedules are constructed to minimize the number of drivers and conductors needed within the negotiated set of rules.

The efficiency perspective is the primary focus of the tactical planning stage where the generic resource schedules are set up.

System perspective

From the perspective of the system, it is desirable to be able to keep the system running in the event of irregularities. This is referred to as the *robustness* of the system. The robustness consists of several elements, which can be classified as either *pro-active* or *re-active*.

Pro-active robustness refers to the slack in the system such as running time supplements. This is the extra time added to the traveling time of a train so that the service can be performed without necessarily running at the technical maximum speed. Adding slack to the system increases its ability to absorb small disturbances without interference from controllers. On the other hand, slack increases the operational costs of the resource schedules which contradicts the efficiency goals.

Re-active robustness refers to the system's ability to recover from disruptions. This capability is increased for example by placing reserve crew and rolling stock at major stations, and simplifying the structure of resource duties so that recovery is less involved. Again, adding standby resources is detrimental to the efficiency of the system.

The robustness of the system is considered in the strategic and tactical planning phases where the timetable is set up and where the total amount of resources necessary to run the system is estimated and evaluated. In Section 2.8.3 we review the literature on robust planning in passenger railway systems.

Process perspective

A great deal of personnel is involved in the operations; in addition to the drivers and conductors that operate the trains, the work of service personnel, cleaning crews and shunting personnel is essential for the operational process. This means that any changes to the planned operations have to be negotiated with and communicated to a lot of different parties. Naturally, the more parties involved the higher the risk of miscommunication when the plans are rescheduled during operation.

From the perspective of the process it is desirable to implement the planned services with as few deviations as possible. The process perspective is often the primary concern in the late planning phases, i.e. short-term and daily planning, and during the actual operations where the practical implementability of the plans is far more important than any other perspective.

Specifically for the rolling stock, changes to the schedules during operations have consequences for the local operations and thereby affect shunting and cleaning crews. Furthermore, changes in the assignment of rolling stock to trains affect the need for conductors on the trains. Therefore, planners prefer rescheduling the rolling stock in such a way that the rescheduling has the least consequences for these parallel processes.

2.4.1 Conflicting goals

There are several conflicting objectives in the described perspectives of the operational process. The robustness considerations in the system perspective are, as mentioned, conflicting with the efficiency perspective as slack and stand-by resources make the resource schedules less efficient. Also, the process and service perspectives are conflicting since providing better service often requires more flexibility which again requires more changes to the plans on short notice.

The implications of the conflicting goals are that any solution implemented in practice involves a trade-off between the different objectives. The current methods for handling deviations from the planned operations are based on rules derived from experiences of similar situations. This makes the process somewhat inflexible since automated decision support tools for dealing with the specific situation are lacking. The inflexibility means that the focus is primarily on the process perspective, so clearly the perspective of the customer service has a lower priority.

Measuring the performance of the railway operator is in itself problematic. The consequence of the governmental contracts is that the railway operators usually set their operational goals to meet the target numbers for the performance indicators discussed in Section 1.6 at minimum cost. However, performing well with respect to the given indicators in a disrupted situation is not necessarily correlated with the passengers' perception of the service quality as is demonstrated by Vromans (2005).

Consider for example the case where the trains from Rotterdam to Utrecht are canceled due to some infrastructure blockage. Then passengers are advised to travel to Schiphol and take the train to Utrecht from there. However, the train from Rotterdam to Schiphol arrives one minute after the departure of the train from Schiphol to Utrecht which incurs a delay of 29 minutes for the involved passengers. The given travel advise thus causes either a delay in the departure of the train from Schiphol to Utrecht or a long connection time for the rerouted passengers. Deliberately delaying the departure of a train is detrimental to the punctuality, but providing a fast connection for rerouted passengers increases the service quality during the disruption.

The example above implies that the usual evaluation of the service quality may be an inaccurate indicator of the actual performance of the operator in the eyes of the customers.

2.5 Rolling stock planning features and terminology

In this section we introduce the terminology we use for rolling stock planning throughout this thesis. The features described are based on the situation at NS. Other operators have similar features, but some features are very characteristic for NS.

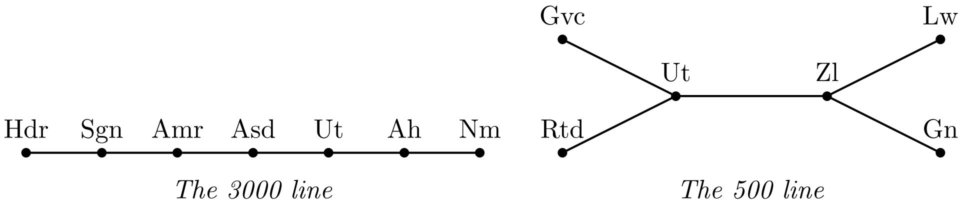


Figure 2.5: The 3000 line between Den Helder (Hdr) and Nijmegen (Nm), and the 500 line between The Hague (Gvc)/Rotterdam (Rtd) and Leeuwarden (Lw)/Groningen (Gn).

Timetable

The basis of the passenger railway operations is the *timetable* which consists of a set of *train services*. A train service is a train that goes from one terminal station to another terminal station with a number of intermediate stops. At NS, the timetable is organized into *lines*. A line is a sequence of stations visited by the same train service. The timetable is cyclic, which means that the same train services are operated with a certain frequency. Each train service has planned departure and arrival times at all stations at which it calls. Figure 2.5 shows some example lines, the 3000 line has a simple structure and runs between Den Helder (Hdr) and Nijmegen (Nm) with a 30 minute frequency while calling at a number of intermediate stations. The 500 line has a more complex structure. Trains from The Hague (Gvc) and Rotterdam (Rtd) are combined in Utrecht (Ut) and later split in Zwolle (Zl) to go to terminal stations Leeuwarden (Lw) and Groningen (Gn) with a one hour frequency. The combining and splitting are reversed on the equivalent service in the other direction. The depicted lines both perform short stops at a number of additional minor stations.

Rolling stock

The passenger train services at NS are primarily operated by electrically powered *rolling stock units*. A unit is self propelled and consists of a fixed number of carriages. Units are available in different types with different characteristics. Units of the same type are considered as interchangeable up to and including short-term planning. In daily planning and during operations a few units may be singled out due to maintenance requirements but the remaining ones are treated as fully interchangeable.

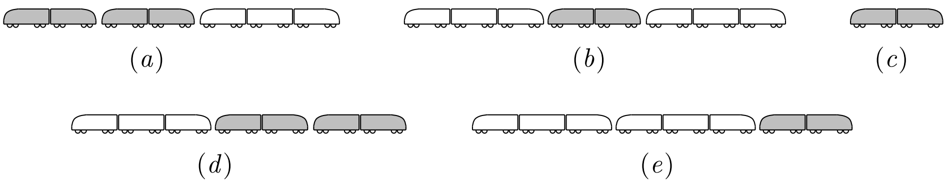


Figure 2.6: Examples of compositions with (a) three units consisting of 7 carriages (b) three units consisting of 8 carriages (c) one unit consisting of 2 carriages (d) three units consisting of 7 carriages (e) three units consisting of 8 carriages.

Rolling stock units can be combined with each other to form *compositions*. This way capacity can better meet demand during peak hours. Figure 2.6 shows different examples of compositions composed of two different types of units. In this thesis we use the convention for drawing compositions that a composition always drives from left to right. The white rolling stock unit in composition (a) is thus in the front position. If composition (a) changes driving direction, it becomes composition (d). Note that although compositions (b) and (e) consist of the same types of units, they are not considered the same compositions because of the order of the units in the composition. Composition (c) consists of a single unit with two carriages.

An essential feature of the rolling stock planning at NS is the fact that, due to efficiency reasons, train compositions are frequently adapted during daily operations by *uncoupling* units from or *coupling* units to trains. The (un)coupling of rolling stock units is performed between trips either at a terminal or underway.

Trips and connections

For the planning process the train services of the timetable are divided into smaller components called *trips*. These trips denote train services from a departure station to an arrival station at specific departure and arrival times. The division into trips is performed in such a way that the assignment of rolling stock to trains cannot be changed during a trip.

The trips that make up a particular train service are connected by fixed *rolling stock connections* (or just *connections*) at the intermediate stations. This means that the same rolling stock units are used on connected trips except from any units coupled or uncoupled at the connection. For simple line structures, such as the 3000 line, the connections involve one incoming trip and one outgoing trip. However, for

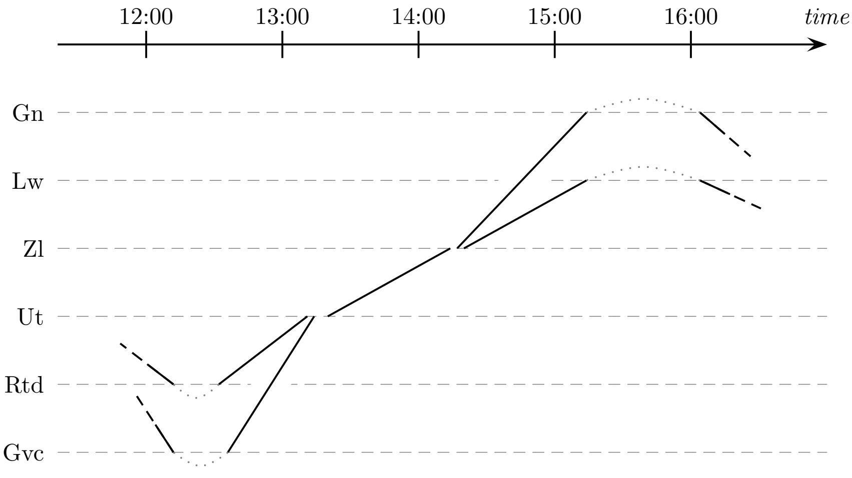


Figure 2.7: Time-space diagram of a train service on the 500 line. This particular service consists of five trips including an underway combination in Utrecht (Ut) and a splitting in Zwolle (Zl). The dotted curves denote turnings at the terminal stations.

more complex lines, such as the 500 line, a connection may involve combining the trains from two incoming trips to a train on an outgoing trip, or splitting the train on an incoming trip into two trains on two outgoing trips.

In the planning process fixed rolling stock connections are also used at the end of a line. Usually, the driving direction of the train is reversed at connections at the end of the line, this is known as *turning* the train. Generally, a certain pattern is used so that each incoming trip at an end station is matched with a specific outgoing trip there.

Figure 2.7 shows a train service on the 500 line in a *time-space diagram* where time is shown in the horizontal dimension and stations in the vertical dimension. The trips that make up the train service are shown as solid lines from a station at a certain time to another station at a certain time. The turnings are shown as dotted curves at the terminal stations, and the underway connections involve a combining of trains in Utrecht and a splitting of the train in Zwolle.

Shunting

The non-timetabled movements of rolling stock units inside railway nodes are called *shunting*. This includes the uncoupling of units from trains and the coupling of units to trains, which is also known as *composition changes*; these are the operations that can be performed between consecutive trips. Usually only certain types of composition changes are allowed depending on the layout of the station. For example the standard allowed composition changes are to couple up to two units to the front end of a train or to uncouple up to two units from the rear end of a train.

Tasks and duties

The position of a rolling stock unit in a composition is relevant with respect to whether the unit can be uncoupled from the composition at a station. The combination of a trip and a position in a composition is referred to as a *task* of a rolling stock unit. For example, if the composition assigned to a trip consists of three units, like in Figure 2.6(a), it results in three tasks, namely one for the front position, one for the middle position, and one for the rear position. The series of tasks performed by a rolling stock unit over a certain time horizon, for example a day, is called the *duty* of the unit.

Circulations and inventory

The global assignment of rolling stock compositions to trips is called the rolling stock circulation or just *circulation*. At NS, each line usually has a fixed number of rolling stock units for each day, so that each line has a closed circulation of rolling stock on a daily basis.

According to the circulation of the units each unit will end up at a specific station at the end of the planning period. The number of units of each type that end their duties at each station is called the *rolling stock balance*. The balance connects the rolling stock circulation to that of the next planning period. Deviations in the rolling stock balance caused by units ending their duty at a different station than originally planned are known as *off-balances*. A means to resolve off-balances is to *reposition* rolling stock units. Repositioning means moving a train consisting of one or more rolling stock units from one station to another without passengers.

The number of units of each type that is available at a station at any given time is called the *inventory* of the station at that time. The rolling stock units are stored at the stations on dedicated storage tracks. The monitoring of this process is a task for the ROCC. Therefore, the inventory is a high-level assumption on the storage of rolling stock units. The inventory does not take into account exactly how the units are stored at the station, but only the number of units. When a unit is coupled to a departing train, it is taken from the inventory. Similarly, when a unit is uncoupled from an arriving train, it is added to the inventory.

A circulation is practically implementable if the composition changes performed at the rolling stock connections are allowed and if the circulation uses no more rolling stock units than are available, i.e. if the inventory of each type of rolling stock unit is non-negative at all times.

The operational rolling stock schedule that planners and dispatchers work with consists of the assignment of rolling stock units to duties. Figure 2.8(a) shows the time-space diagram of an example timetable with two trips in each direction between two stations. Figure 2.8(b) shows a rolling stock circulation for the example timetable. In the circulation, a black, a white and a gray unit cover the trips. The duties can be regarded as paths in the time-space diagram as they are shown in Figure 2.8(b). Note that the gray unit is uncoupled after the first trip.

Another view of the duties is shown in Figure 2.8(c) where each unit has a list of tasks i.e. trips along with the position in the corresponding compositions. In the diagram, the tasks are denoted by rectangles whose left and right sides denote the departure and arrival times on the time line, and the departure and arrival stations are found above the top corners. Inside the rectangles the position of the task and the total number of units in the train are depicted.

In the example in Figure 2.8 the gray and the black units are of the same type but the white unit is of a different type as it consists of 3 carriages. The inventory at the beginning of the planning period is thus two units of length 2 at station `A` and one unit of length 3 at station `B`. The inventory at the end of the planning period is one unit of length 2 at station `A` and one unit of length 2 and one of length 3 at station `B`.

Note that the duties of a set of rolling stock units can be translated uniquely to a circulation and also imply the inventories of the stations at all time instants. However, a circulation and an inventory do not necessarily translate to a unique set of rolling stock duties. This is because the inventory only states the number of units available at stations at all times, and does not specify which exact unit is taken from the inventory when a unit is coupled to a departing train. There may thus be several ways to connect uncoupled units with later couplings.

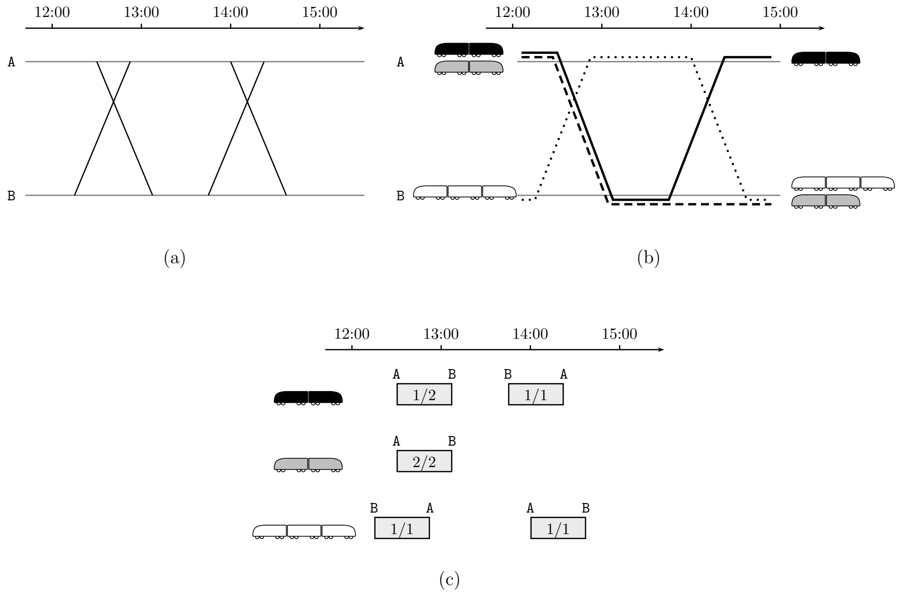


Figure 2.8: (a) Time-space diagram of a small example timetable. (b) Rolling stock circulation. (c) Rolling stock duties.

2.6 Rolling stock rescheduling decisions

In the event of a disruption, the timetable, rolling stock and crew schedules have to be adjusted in a series of steps during the disruption management process. The overall process involves three major steps.

In the first step, the timetable is adjusted according to the disruption. If the disruption involves a blockage of the infrastructure, it immediately incurs a conflict for all trains that were meant to utilize that part of the infrastructure. These conflicts are resolved by canceling or rerouting the trains. To ease this part of the process, the traffic control organization of the infrastructure manager has constructed a number of *disruption handling scenarios* in cooperation with the railway operators. Such a handling scenario assumes a blockage of a particular piece of infrastructure and states which trips are canceled. The trains are then turned at the last appropriate station on either side of the blockage which means that the rolling stock connections between the trips are modified. Outside the disrupted area the timetable usually remains the same according to the handling scenario.

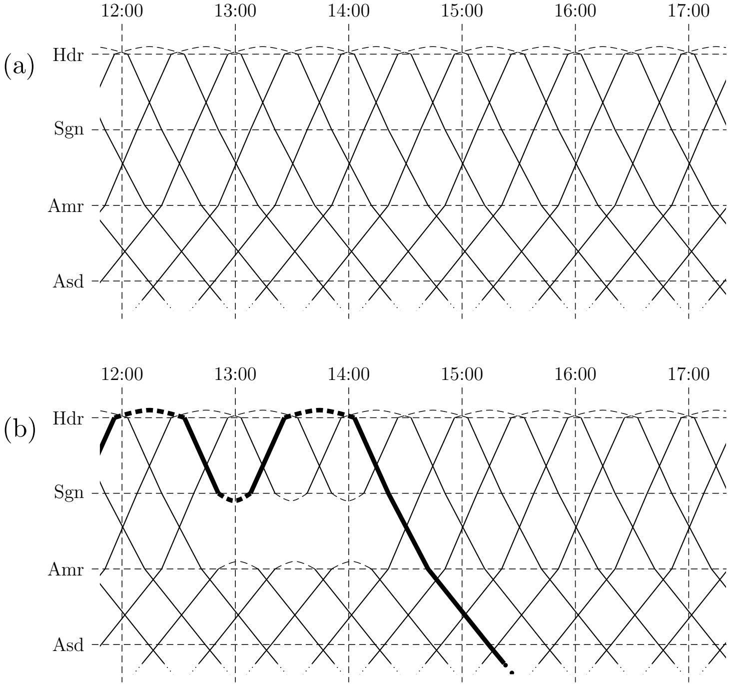


Figure 2.9: (*a*) Time-space diagram for part of the 3000 line. (*b*) Adapted timetable after the occurrence of a disruption between Schagen (Sgn) and Alkmaar (Amr).

An example of the timetable adjustment is given in Figure 2.9. The time-space diagram in Figure 2.9(a) shows the timetable for a part of the 3000 line in the afternoon of a weekday. If a disruption blocks the infrastructure between Schagen (Sgn) and Alkmaar (Amr), the handling scenario states that the trains are turned at Schagen on one side of the disruption and at Alkmaar on the other side. The turning is performed in such a way that the first incoming train is matched with the first outgoing train on either side according to the regular timetable. Based on the estimated duration of the blockage the resulting adapted timetable is shown in Figure 2.9(b).

In the second step, the rolling stock is rescheduled to serve the adjusted timetable. If the rolling stock connections between some trips are changed during the adjustment of the timetable, the planned assignment of compositions to trips cannot be implemented in practice. In the rolling stock rescheduling step, the goal is to assign the rolling stock in such a way that the capacity meets passenger demand as much as possible. The means for adjusting the assignment is by modifying the planned composition changes.

Consider the train arriving at Schagen (Sgn) shortly before 13:00 in Figure 2.9. In the adapted timetable the train is turned rather than continuing through. The train now follows the path of trips marked with bold in the diagram in Figure 2.9(b). This means that different compositions may have to be assigned to the train services which may render any planned composition changes impossible to perform.

In the third step, the crew is rescheduled. Each train needs a driver and a number of conductors to operate. The crew duties are rescheduled according to the result of the rolling stock rescheduling. This is performed through exchanging tasks between crew members and reserve crew.

Even though the three major rescheduling steps are presented as sequential here, they are interdependent to some extent. If there is no practical solution to one step the previous step may be revisited to take into account the lack of resources. This is for example the case if there is no available rolling stock for a trip, then the timetable may be updated by canceling that trip. This is also the case if no crew is available to serve a trip, then both rolling stock and timetable have to be revised accordingly.

2.6.1 Options for rolling stock rescheduling

The purpose of the assignment of rolling stock to trips is to provide capacity to satisfy the passenger demand. This also holds true in the rescheduling phase. Therefore the goal of the rolling stock rescheduling process is to provide adequate capacity to the planned trains that are not canceled as well as to any rerouted trains during the disruption. In current practice, the goal is to return to the planned situation after the disruption. The reason for this is pragmatic – by returning to the plan a highly optimized feasible schedule for the operations is already available. Returning to schedule is complicated by the fact that a number of rolling stock units are at wrong positions at this point. The available options for the dispatchers performing the rolling stock rescheduling can be grouped into three categories: (*i*) Changing shunting operations, (*ii*) adapting turning patterns, and (*iii*) repositioning train units. These options are explained next.

Changing shunting operations

According to the short-term plan there may be a planned shunting operation at the connection between the arrival of a trip and the departure of the next trip. However, due to a disruption, a different composition may be assigned to the arriving train. This may render the planned shunting operation invalid.

Consider an example where the planned shunting operation involves uncoupling a unit of a specific type, and the arriving composition does not contain a unit of that type. Then it is impossible to perform the planned composition change. Also, it may be the case that performing such an uncoupling reduces the seat capacity to an inappropriate level. Similarly, if the planned shunting operation involves coupling one or more units to the train, then it may result in a train that violates the restrictions on train length. We illustrate the different ways to change the planned shunting operations through the examples in Figure 2.10. Each row in the figure shows the planned shunting operation on the left side and the potential new one on the right. The examples use two types of units; small units with two carriages and large units with three carriages.

A potentially problematic modification of the shunting plans is the introduction of a composition change at a connection where no composition change was planned. This modification may cause problems because it implies changes to the local plans. These changes will have to be negotiated with the appropriate local dispatchers. A new shunting operation may require free storage capacity and available tracks for the movement of the train units. In Figure 2.10(a) no shunting operation is planned, thus uncoupling a unit is considered a new shunting operation.

A change to the shunting plans, that is usually less challenging, is to cancel a planned shunting operation. This change still has to be communicated to the local dispatchers, but it means less work for the involved shunting personnel and thus they are likely to agree to it. Problems with storage capacity in the shunting yard may occur if a planned coupling of several units is canceled. In Figure 2.10(b) a short unit was planned to be uncoupled. However if the units arrive in a different order, the planned uncoupling is not possible and it is an option to cancel the planned shunting operation.

Changing a planned coupling to an uncoupling may also potentially cause difficulties. If a coupling was planned then shunting personnel is available but rather than retrieving a unit from the shunting yard they now have to uncouple a unit and bring it to the shunting yard. This requires significant changes to the movements inside the railway node. Figure 2.10(c) illustrates how a planned coupling may be changed to an uncoupling to achieve the originally planned departing composition. Similar concerns relate to changing a planned uncoupling to a coupling as illustrated in the example in Figure 2.10(d).

Coupling a different type of unit or a different number of units is usually not too problematic. The rolling stock movements inside the station are relatively unchanged, but the unit has to be retrieved from a different track in the storage yard. Figure 2.10(e) shows an example where a planned coupling of a long unit is changed to a coupling of a short unit. Similarly, the example in Figure 2.10(f) illustrates how a coupling is changed to involve a different number of units since two short units are coupled instead of one. Uncoupling a different type or a different number of units is also not too tricky although capacity problems in the shunting yard may occur. The example in Figure 2.10(g) shows how a planned uncoupling of a long unit is changed to an uncoupling of a short unit. Similarly, the example in Figure 2.10(h) illustrates how an uncoupling is changed to involve a different number of units.

Adapting turning patterns at terminal station

In tactical and short-term planning it is often preferred to use a certain turning pattern at terminal stations. This means that each incoming train is matched with a particular departing train. The application of such a turning pattern is useful in the planning phase as it eases the task of local planning by continuously repeating turning schemes. However, in the real time rescheduling phase it may be too restrictive to uphold the same turning patterns.

The concept is shown by an example: Figure 2.11 shows the time-space diagram for a part of a timetable. In particular it shows a number of trips arriving at or departing from a terminal station, along with the turning pattern. In this case each arriving trip is matched with the second departing trip after its own arrival. Rolling stock units are assigned to the trips and during the shown period one white unit is coupled to the second departing train. This turning pattern implies that at any point in time either one or two trains are idle inside the station.

Consider the example where, due to some disruption elsewhere in the network, different compositions than expected are assigned to the first two incoming trains in the example. Then, as shown in Figure 2.12, major changes to the planned shunting operations are needed if the departing trains are to have the originally planned compositions assigned. In particular, the first train will need to both uncouple a gray unit and couple a white one. Uncoupling from and coupling to the same composition within this time frame is generally not possible so that is not an option.

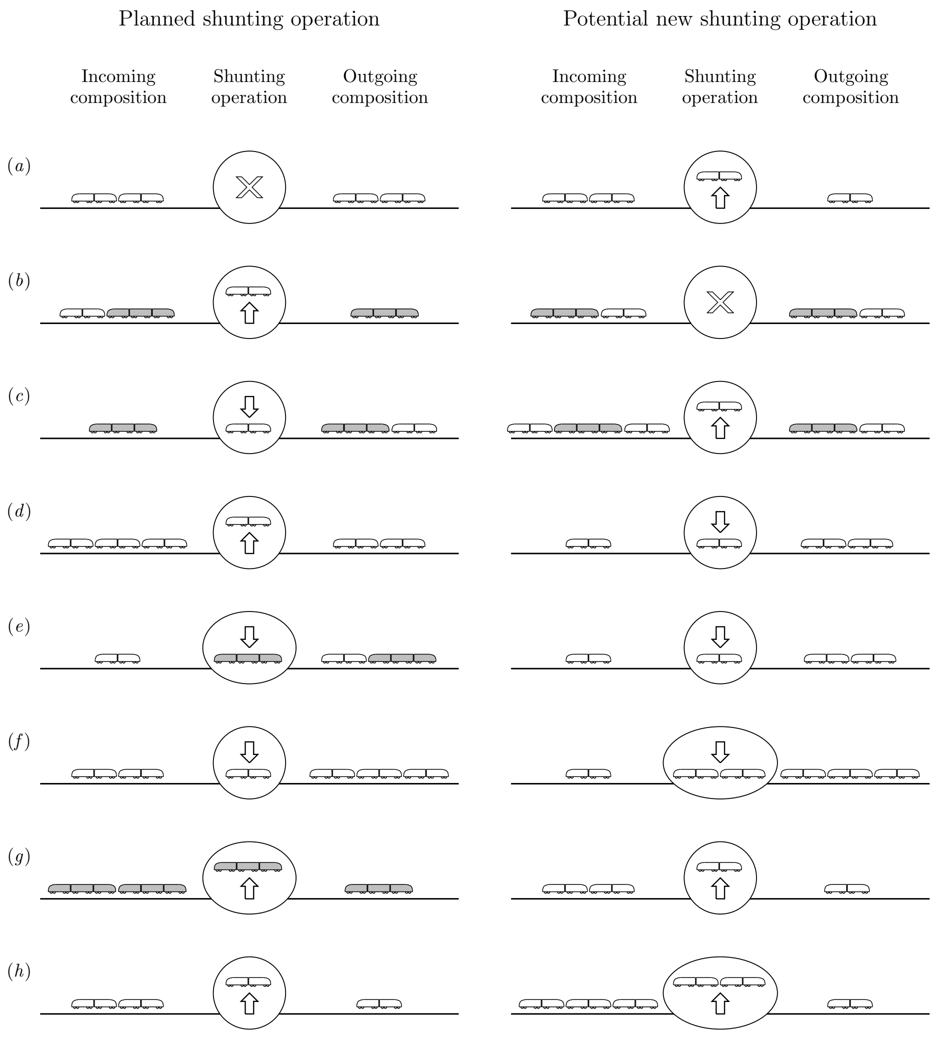


Figure 2.10: Examples of changes to the shunting operations. Each row shows a planned shunting operation in the left column and a potential new shunting operation in the right column. In the examples, it is possible to couple units to the front (right side) of the train and uncouple units from the rear (left side) of the train. Each shunting operation is illustrated with an incoming composition on the left, an outgoing composition on the right, and the composition change encircled in the middle. Coupling is denoted by a downward arrow, uncoupling by an upward arrow, and shunting operations that involve no composition change are denoted by an X.

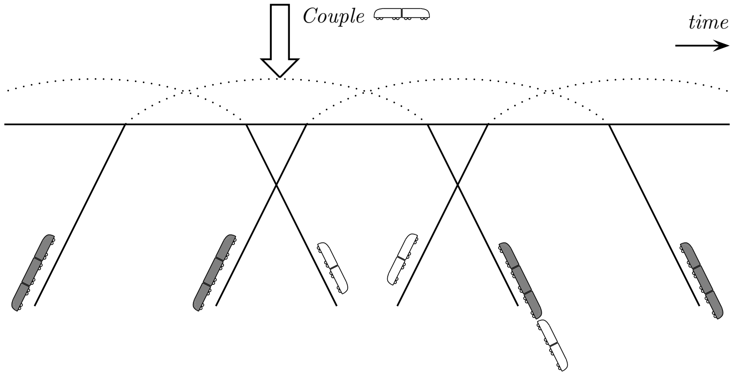


Figure 2.11: Part of rolling stock circulation including turning pattern.

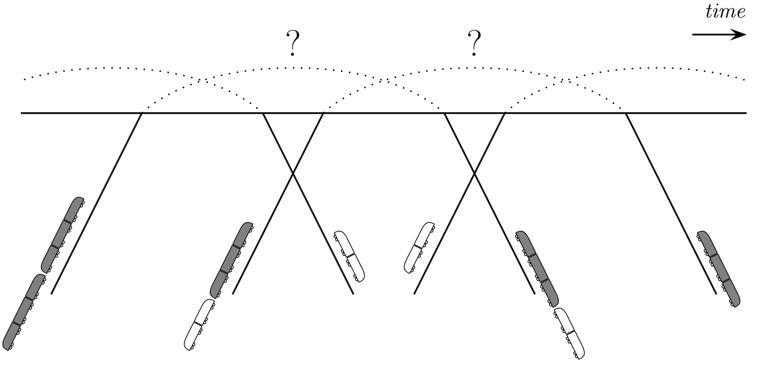


Figure 2.12: Part of rolling stock circulation with different incoming compositions.

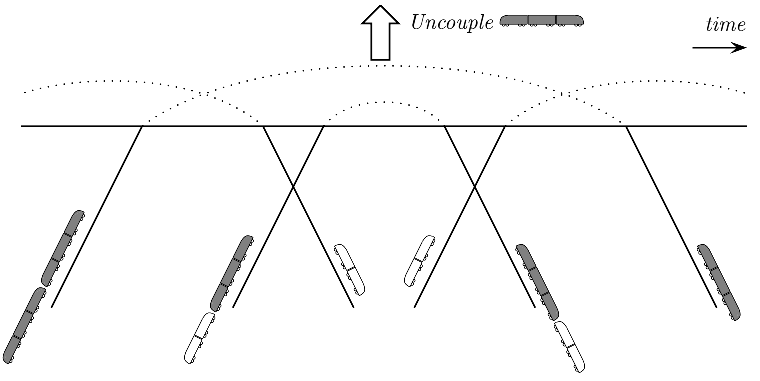


Figure 2.13: Solving the problem by adapting the turning pattern.

A more elegant solution is to locally change the turning pattern by matching the first incoming trip with the last outgoing and the second incoming trip with the trip departing immediately after. Then the compositions fit if a gray unit is uncoupled from the first incoming trip. The changed turning pattern is shown in Figure 2.13. We note that the planned turning pattern leaves up to two compositions idle in the station at the same time. The changed turning pattern also leaves at most two compositions idle in the station at any point in time. This means that the same number of platforms are needed at the station during operation, only the routes inside the station are different.

Changing the turning patterns has a number of consequences for the ROCC. They have to either inform the passengers that a train departs from a different platform than expected, or they may switch platforms of the trains upon arrival, if possible. In either case, deviating from the planned turning patterns is a decision that has to be made jointly by NOCC and ROCC since it has both global and local consequences.

Repositioning of rolling stock

An option in the rolling stock rescheduling phase is to reposition rolling stock units. This option may be necessary if there is an acute shortage of rolling stock at a particular station. During daytime the railway network is heavily utilized so adding more trains may be problematic, but in some situation it may be the only option for providing adequate rolling stock for certain trains.

Repositioning of trains is on the other hand a frequently used option during nighttime. Here it is a useful tool for solving off-balances, i.e. deviations in the rolling stock balance caused by units ending their duty at a different station than originally planned. Resolving off-balances is necessary to ensure that the appropriate rolling stock units are available the next day for the morning peak.

2.7 Literature overview

There exists a vast amount of literature on algorithmic support for planning problems in passenger railway systems. Assad (1980) provides a survey that describes the problems that arise in all planning phases of a railway system, and Cordeau et al. (1998) give an overview of models for the involved problems. A recent overview of Operations Research models applied in railway optimization is given by Caprara et al. (2007b).

Tools based on Operations Research have been applied to a range of planning problems at NS. We here go through each step in the planning phase and discuss the related literature. For an elaborate overview we refer to Huisman et al. (2005) and Kroon et al. (2009).

2.7.1 Strategic planning

In the strategic planning phase, Operations Research methods have mainly been applied to line planning. The methods in the literature typically focus on the trade-off between the efficiency of the line plan and the service it provides. For line planning at NS, Goossens et al. (2004) formulate the problem as a Mixed Integer Linear Programming (MIP) problem where the objective is to minimize the estimated operational costs. They apply a branch-and-cut approach for solving the model and test it for several practical instances.

For further research on line planning in the Netherlands, we refer to Goossens (2004), while line planning in general is discussed by Bussieck et al. (2004), Scholl (2005), Schöbel and Scholl (2006) and Borndörfer et al. (2008).

2.7.2 Tactical planning

The tactical planning phase is a prominent application area of Operations Research in railway planning. At NS, Operations Research models have been applied to both timetabling and resource planning.

Timetabling

In the tactical planning phase, the basic timetable is constructed once per year. For this task, NS uses the software package DONS (Designer Of Network Schedules), that contains two complementary algorithmic modules: CADANS and STATIONS. CADANS generates the tentative timetable based on the line plan and the network layout (see Schrijver and Steenbeek (1994) and Schrijver (1998)). The core of CADANS is based on the Periodic Event Scheduling Problem (PESP) (introduced by Serafini and Ukovich (1989)) and is solved by constraint programming with a limited post processing optimization module. STATIONS generates routes and schedules for trains through stations to check whether the timetable is locally implementable at the stations (see Zwaneveld et al. (2001)).

Since the integration of DONS in the timetabling process at NS, a number of models for timetable generation have been proposed. Like the core model in DONS, they are based on PESP, such as Peeters and Kroon (2001). Kroon and Peeters (2003) describe an extended model that includes variable trip times. Recently, Kroon et al. (2008b) presented a stochastic approach to cyclic timetable optimization that aims at minimizing the average delay of the trains. This model has been applied in practice to a set of lines at NS resulting in a significant increase in on-time performance of the involved trains. For an overview of cyclic timetabling based on PESP we refer to the work by Liebchen (2006) which includes optimization approaches based on MIP.

The PESP concept is extended by Caimi et al. (2007) to Flexible PESP in which a solution consists of intervals for event times rather than exact event times. This allows for more flexibility in the succeeding routing of trains through railway nodes.

In other railway settings, especially in cargo settings, non-cyclic timetabling problems occur (see Cacchiani et al. (2010a), Caprara et al. (2002) and Brännlund et al. (1998)). Non-cyclic timetabling aims at determining the timetable for a set of trains subject to track capacity constraints. In the cargo context the objective is often to maximize the utilization of the network and thereby the throughput of cargo.

Rolling stock scheduling

A rich list of publications address the problem of constructing rolling stock schedules under various assumptions. Since this thesis focuses on the operations of rolling stock, we here list a number of the papers on tactical rolling stock scheduling and describe the assumptions and solution approaches. We start with the approaches that have been applied at NS.

For rolling stock scheduling at NS, Abbink et al. (2004) propose a MIP model for distributing the available rolling stock between lines. In this model, rolling stock types are assigned to lines while minimizing the anticipated shortage of seat capacity during peak hours. The model is solved by applying a commercial solver.

For setting up the generic rolling stock circulations at NS, Alfieri et al. (2006), Peeters and Kroon (2008) and Fioole et al. (2006) all consider the problem of assigning compositions to trips. The models optimize the weighted service and efficiency of the schedules while also accounting for some robustness aspects. Alfieri et al. (2006) present a MIP model for the problem for simple line structures and solve the model for a number of real life instances from NS. Peeters and Kroon (2008) and Fioole et al. (2006) tighten the formulation but propose different solution strategies. Peeters and Kroon (2008) use a branch-and-price approach for solving the model. This dedicated solution procedure achieves a speed up compared to a generic commercial solver, but the model can only handle simple line structures. Fioole et al. (2006) extend the formulation to handle more complicated line structures such as splitting and combining and solve it with a commercial solver. Currently, this latter model is used for constructing the generic rolling stock circulations at NS. The model forms the basis for the models used in Chapters 3 to 6.

Cordeau et al. (2000) study the problem of simultaneously assigning locomotives and locomotive-hauled carriages to trains. Shunting is considered possible from or to any position in the train so the model used does not consider the order of units in the compositions. The resulting MIP model is solved using Benders decomposition within short computation times for a number of real life instances. Cordeau et al. (2001b) extend the model to facilitate maintenance constraints and car switching penalties, and solve the model with a heuristic branch-and-bound scheme. Cordeau et al. (2001a) further extend the model and apply it to instances from VIA Rail Canada. Rouillon et al. (2006) elaborate on the branching strategies for the branch-and-bound schemes for the solution methods.

Ahuja et al. (2005) address the problem of assigning locomotives to cargo trains at the US cargo operator CSX Transportation. They model the possibilities for assigning several locomotives to the same trains as well as repositioning locomotives. The assignment has to obey requirements on pulling power and availability of locomotives, and is planned cyclically for an entire week at the time. The resulting MIP formulation is solved with a solution technique based on integer programming and very large-scale neighborhood search. The authors report considerable potential savings on real life instances. The model is extended by Vaidyanathan et al. (2008) to incorporate a number of realistic constraints such as requirements for locomotives fitted with special equipment, power requirements on foreign tracks, and repeatability of the schedule.

Mellouli and Suhl (2007) consider the problem of assigning locomotives and carriages to sets of trips. The models consider several types of locomotives and carriages. The objective is to minimize operational costs while taking special restrictions into account, in particular, the given capacity requirements must be obeyed. The solution approach is based on a multi commodity network flow in a multi-layered network with additional cover constraints. The authors are able to solve practical problems from the German Railways.

Cacchiani et al. (2010b) present a case study for assigning locomotives and carriages to the passenger trains of a regional Italian operator. The problem is modeled with an integer program formulation and sophisticated techniques are used to ensure a strong LP relaxation. The instances are characterized by a large number of distinct rolling stock types which results in a particularly large number of allowed compositions on the involved trips. The authors solve the model with a heuristic based on the LP relaxation of the problem and report potential improvements compared to manually created solutions on all instances in their case study.

Folkmann et al. (2007) introduce models for evaluating the amount of rolling stock and crew needed to run a candidate timetable. The models are applied to the timetable of DSB S-tog A/S, an operator in the suburban area of Copenhagen, Denmark.

Crew scheduling

Planning tools based on Operations Research have been successfully implemented for tactical crew planning at NS. Since 1998 the crew planning tool CREWS (Morgado and Martins (1998) and Siscog (2010)) has been used by NS for assisting the manual crew planning process. In 2000 the tool TURNI was added as a supplementary optimization tool for the tactical crew scheduling. The optimization module in TURNI is based on a set covering model that is solved with a column generation scheme based on Caprara et al. (1999a) and Caprara et al. (1999b). The software was tailored to the special Dutch requirements on crew duties as described by Kroon and Fischetti (2001).

In 2001, TURNI played a major role in designing the new crew duty rules which finally led to a new agreement between the labor unions and NS (see Abbink et al. (2005)). In 2008 an optimization module based on the models by Caprara et al.

(1999b), was integrated in CREWS which has been used for crew schedule optimization at NS since then.

2.7.3 Short-term planning

In the short-term planning phase, the scheduling and rescheduling of the resources have been the subject of a number of studies.

Rolling stock scheduling

Rolling stock scheduling in the short-term planning phase involves adapting the generic timetable and rolling stock schedules to days or weeks with special requirements. Budai et al. (2009) study the rolling stock balancing problem. This is the problem of ensuring that the rolling stock ends up at the correct stations at the end of the day. Balancing is an issue when the rolling stock schedules are adapted to a modified timetable. Their models take the exact order of the rolling stock units in the trains into account and minimize the number of rolling stock units that end up at different stations than planned. The rescheduling is subject to a number of constraints on capacity demand and shunting possibilities. The problem is solved for fairly large instances by iterative heuristics with short running times.

Ben-Khedher et al. (1998) describe an approach for short-term rolling stock planning at the French passenger railway operator SNCF. An important assumption is that passengers book seats in advance of their journey. The approach contains two modules; one is focused on revenue management for the seat reservation system, and the other later adjusts the rolling stock assignment based on actual seat reservations. The second module is applied with a very short time horizon and can be considered an application in the daily planning phase.

Lingaya et al. (2002) extend the model by Cordeau et al. (2001b) for rescheduling the carriage assignment to locomotive hauled trains to take a number of short-term aspects into account. The carriages are rescheduled to meet short-term demand revisions formulated by the marketing department which is responsible for tracing the ticket sales. The model considers only a short planning horizon and ensures that the duties of the rescheduled carriages connect with their original duties at both ends of the horizon. Operational constraints on maintenance requirements and minimal switching times are also incorporated.

Brucker et al. (2003) suggest a model for the problem of assigning carriages to locomotive hauled trains. The model is based on an existing assignment of compositions to trains and the model considers both trains in the timetable and optional empty trains for repositioning rolling stock between stations. The model decides on the detailed routing of railway carriages for a given planning horizon. The objective is to minimize the weighted number of carriages used and the number of repositioning trips. The approach is based on an integer multi-commodity network flow formulation with a non-linear objective for satisfying passenger demand while minimizing cost. The model is solved with a local search approach and the method is able to solve real life instances.

In Chapter 6 we describe a case study on rolling stock rescheduling in the short-term planning phase at NS. We introduce a model for the problem and discuss how the results of the model are implemented in practice. A prototype implementation of the model has recently been adopted to become a stand-alone decision support tool in the short-term planning phase at NS.

Crew scheduling

As with short-term rolling stock planning, crew planning in the short-term phase involves adapting the generic crew schedules to days or weeks with special requirements. For short-term crew planning at NS, Huisman (2007) formulates the problem as a large-scale set covering problem and solves it with a column generation algorithm. The algorithm has been used for several real life instances for which some of the solutions were implemented in practice. The application of this algorithm leads to a speedup of the short-term planning process. In 2006 the optimization module of the Crews planning software was extended to handle short-term crew planning as well.

The model described by Potthoff et al. (2009) is designed for crew rescheduling in the disruption management process, but it has also been applied to short-term rescheduling problems at NS. The model is based on a set covering problem and solved using column generation techniques combined with Lagrangian heuristics. The algorithm defines a core problem of crew duties that are affected by the short-term specifications and dynamically increases the problem size if tasks are uncovered. See also Potthoff (2010).

2.7.4 Daily planning

An important aspect of daily rolling stock planning is the monitoring and routing of rolling stock units that need maintenance. Maróti and Kroon (2007) describe the maintenance routing problem. It is the problem of exchanging duties between rolling stock units to ensure that units with maintenance requirements reach the proper maintenance facility before certain deadlines. The authors study the complexity of the problem and formulate an integer programming model for solving the problem. The problem is also studied by Maróti and Kroon (2005), who suggest an approach based on a network flow model that requires less data. Both papers provide computational results for instances from NS.

2.7.5 Shunting, platforming and movements inside railway nodes

Along with the centralized planning problems concerning the timetable, the rolling stock and the crew, there is a number of more localized planning problems. These problems involve the shunting process, the assignment of platforms to trains and the routing of trains inside railway nodes. The time horizons of these problems span several planning phases since they are relevant already when constructing the timetable, and in later planning phases when the local consequences of rolling stock circulations are considered.

Shunting problems occur across several planning horizons and are thus considered in both tactical, short-term, daily and real-time planning. Shunting refers to the movement of rolling stock inside stations and railway nodes, but the storage of rolling stock units during idle time is also part of the shunting problem. The planning of shunting movements and storage of rolling stock is related to rolling stock scheduling in the sense that a rolling stock circulation immediately implies the need for shunting at the stations. However, in the rolling stock scheduling literature shunting is generally only accounted for in an implicit manner, i.e. by ensuring certain minimum connection times between tasks that require shunting in between.

For the shunting problems at NS, Freling et al. (2005) develop a framework for shunt planning in both short-term and daily planning. They propose mathematical models for the involved problems and solve them with a solution method based on column generation. For an overview of Operations Research models in shunt planning, and at NS in particular, we refer to Lentink (2006) and Kroon et al. (2008a). We further refer to Di Stefano and Koči (2004) and Cornelsen and Di Stefano (2007) for complexity studies of certain classes of shunting problems.

The routing of trains through railway nodes is considered in several publications. Zwaneveld et al. (1996) and Zwaneveld et al. (2001) present a conflict graph model for the routing of trains through stations. The models are based on the node packing problem where the possible decision of a specific route for a specific train at a specific time is modeled by a node, while an edge represents a conflict between two possible routings. The models are solved by using valid inequalities and a branch-and-cut approach. Kroon et al. (1997) discuss complexity considerations for these types of node packing models.

Fuchsberger (2007) presents an alternative model for routing trains through a station based on a resource tree conflict graph. Here a node represents a specific train at a specific time passing a network topology element with a given velocity. Then the possible routes through the station topology correspond to trees, and the objective is to find a path from the root to a leaf for all trains subject to the conflict constraints. The resulting IP models are solved using a commercial-strength IP solver in short computation times for realistic instances.

The problem of deciding which trains call at which platforms in the stations is known as the platforming problem. The problem is related to shunting and train routing as the platform assignment restricts the available routes for arriving and departing trains as well as shunting movements inside the station. The platforming problem is studied by Caprara et al. (2007a) who formulate the problem as an IP model with a quadratic objective. The model is solved by linearizing the objective and then applying a diving heuristic. Kroon and Maróti (2008) consider a similar formulation of the problem and add extra terms to the objective function to account for some robustness considerations. These considerations concern the probability of routing conflicts in case of small delays. For other studies of the platforming problem we refer to Carey and Carville (2003) and Billionnet (2003).

2.7.6 Disruption management and real time control

In Section 2.3 we described the disruption management process of a passenger railway company and discussed the roles of the involved actors. Jespersen-Groth et al. (2010) study the practical and organizational aspects of the process and identify the problems related to each step in the process. The authors specifically study the situation at DSB S-tog A/S, an operator in the suburban area of Copenhagen, Denmark, as well as at NS. They conclude that there is a potential for improving the quality of the operations through the development and application of Operations Research models for decision support in the disruption management process.

A problem in the real-time operations of railways that has received much attention, is the problem of dispatching train traffic through corridors and stations. Even though train dispatching is the responsibility of the infrastructure manager, it has implications for the performance of the involved operators as well. Dispatching strategies are a major factor in the management of small delays and their cumulative effects. Törnquist (2005) provides a comprehensive overview of the literature on the dispatching problem and classifies the surveyed models with respect to problem type, solution mechanism and type of evaluation. D'Ariano (2008) and Corman et al. (2010) model the dispatching problem as a job shop scheduling problem and suggests several solution methods including local search and branch-and-bound techniques. The authors apply the methods to several cases from bottlenecks in the Dutch network. The objective is measured by the total knock-on delay experienced in the system due to conflicts incurred by delayed trains. The models achieve significant improvements over the simple rule based dispatching methods applied in practice.

We have already mentioned the models by Budai et al. (2009) and their application in short-term planning. However, as the authors claim, the models are also relevant in the disruption management setting. The problem of balancing the rolling stock after a disruption is relevant: off-balances require either expensive repositioning of rolling stock units or may cause poor utilization of seat capacity in the peak hours of the following morning. The short computation times of the models presented by the authors are attractive in a real-time setting.

Jespersen-Groth and Clausen (2006) present an approach for integrated rolling stock rescheduling and timetable adaption in case of a disruption. The authors consider the option of canceling an entire line in case of a disruption and reinsert the trains as soon as the disruption is over. They formulate a MIP model which minimizes train cancellations on the disrupted lines. The authors report average computation times of less than one second on instances from the Danish railway operator DSB S-Tog A/S. Due to the size and structure of the involved railway network, it is possible to enumerate all possible disruption scenarios, and the model has been used for generating offline reinsertion plans for all scenarios.

Real-time crew rescheduling in passenger railways is the problem of adapting the crew schedules to serve an updated timetable and rolling stock circulation. Potthoff et al. (2009) describe an algorithm based on column generation techniques for rescheduling the drivers when a disruption occurs in a passenger railway system. They test their methods on several realistic cases from NS and achieve high quality solutions in few minutes of computation time. The model is extended by Veelenturf et al. (2010) to allow the possibility for delaying the departure of some trains. Thereby, timetable adjustments and crew rescheduling decisions are partly integrated. Computational experiments show that the retiming option improves the solutions along the crew rescheduling cost metrics. Rezanova and Ryan (2010) present a fast set partitioning heuristic for the train driver recovery problem and test it on instances generated from historical real-life operations of the Danish railway operator DSB S-Tog A/S.

Abbink et al. (2009) propose an approach to train crew recovery based on the Actor-Agent paradigm: Crew members are represented by software agents who negotiate rescheduling options among them. Coordination between agents is based on team formation for the evaluation of alternatives subject to constraints and preferences given by dispatchers and the involved crew. The authors test the approach on several real-life instances of NS and find feasible solutions in all cases.

Walker et al. (2005) present an integrated model for train and crew recovery in the event of disruptions. The approach incorporates the complex crew constraints into the timetable modification decisions and minimizes the train idle times. The model does not explicitly consider rolling stock as it does not contain possibilities for changing the assignment of rolling stock to the trains.

2.8 Related fields

The application of Operations Research in the disruption management process of railway systems is still a relatively uncharted field. More research has been conducted, however, on the topic in other logistic settings. Especially the airline industry has attracted extensive research on the disruption management process. In this section we list a number of successful applications of Operations Research in disruption management in related fields.

For an extensive framework on disruption management, we refer to the work by Yu and Qi (2004). The authors introduce a number of modeling paradigms for disruption management and apply them to different settings. We will discuss the applicability of their methods to passenger railway operations in Chapter 4 where we develop a framework for disruption management in passenger railways.

2.8.1 Disruption management in the airline industry

There is a number of similarities between disruption management in passenger railways and in the airline industry. In both cases the timetable is published in advance and thus has to be taken into account, and aircraft and rolling stock as well as crews have to be rescheduled in case of a disruption. However, disruptions are handled quite differently in the airline industry compared to the railway industry. This is due to a number of key differences in the features of the two industries. First of all, trains are bound to the infrastructure network. Second, the number of operators that share both airports and airspace is usually much larger in the airline industry leading to different challenges than experienced in passenger railways. Third, aircraft usually only perform a few flights per day whereas a rolling stock unit may be assigned to dozens of train services during a day. Fourth, trains may consist of several coupled rolling stock units. Also the much stricter maintenance requirements for aircraft pose additional challenges. At the same time, more knowledge on passenger flows is available in the airline industry due to reservation systems. This means that passengers can be routed individually in case of delays or cancellations.

Disruption management in the context of airlines is a well studied field. For an overview of the research on rescheduling in the airline industry we refer to Clausen et al. (2010) and Kohl et al. (2007).

Aircraft make up the resource that provides seat capacity in the airline setting where rolling stock is the corresponding resource in the railway setting. Yan and Lin (1997) and Thengvall et al. (2001) study the problem of recovering the circulation of aircraft after a disruption that involves the closure of an airport. The process is oriented to minimizing the cost related to delays and cancellations. Rosenberger et al. (2003) present a model for aircraft recovery that not only minimizes operational cost, but also implicitly takes crew and passenger recovery into account by penalizing crew disruptions and disruptions in the passenger flows. Aircraft recovery is also studied by among others Thengvall et al. (2000), Yan et al. (2005) and Ball et al. (2007).

2.8.2 Disruption management in vehicle scheduling applications

The vehicle scheduling problem is addressed in a vast number of publications and variants of the problem have applications in many logistic settings such as pick-up and delivery, dispatching of services and scheduling of shuttle services (see Bodin and Golden (1981) and Bunte and Kliewer (2009)). Many applications concern scheduling of buses or trucks from one or more depots with or without time-windows. A number of those applications address online variants of the problem where vehicles are dynamically dispatched at the arrival of new requests (see Ghiani et al. (2003)).

Both online and disruption management in vehicle scheduling consider problems where existing schedules have to be adapted to a change in the environment. However, although the online version considers an important and realistic operational aspect, it differs from disruption management in the sense that disruption management concerns unwanted restrictions rather than new requests by customers.

Recently, disruption management in vehicle scheduling has received more attention such as by Li et al. (2007), who consider the vehicle rescheduling problem for buses. There are several similarities between the rescheduling of buses and railway rolling stock. In both cases, vehicles are assigned to trips in a timetable, and vehicles are dispatched from one or more depots, or shunting yards in the railway case.

There is a number of differences between the two settings as well. Most notable is the difference in the underlying infrastructure, buses utilize the road network whereas trains of course run on rails. Rescheduling vehicles on the roads is often less involved as there may be many alternative routes to the destination, and buses are not subject to the complicated headway security requirements experienced in the one-dimensional railway setting. Another difference is that several railway rolling stock units are usually combined to make longer trains.

2.8.3 Robust planning and recoverability

Where disruption management is concerned with the actual reassignment of resources, robust planning is about incorporating disruption management measures already in the planning phase. As mentioned in Section 2.4, robust planning concerns pro-active features such as slack and reserve resources, and re-active features such as recoverability.

Several papers consider robust planning of railway resources. One robustness aspect that is considered in several publications is the built-in option of exchanging crew duties in case of a disruption. This is introduced as the concept of *move-up crews* by Shebalov and Klabjan (2006) in the airline industry. The concept is adapted to rolling stock schedules by Cacchiani et al. (2008b), who use move-up rolling stock units to increase the recoverability of a rolling stock schedule. They explore the trade-off between the efficiency and the recoverability of the resulting schedules.

Cacchiani et al. (2008a) further explore the recoverability of rolling stock schedules and quantify the trade-off as the price of recoverability. The authors define it as the loss in efficiency for the schedules to be inexpensively recovered in a number of disruption scenarios. A mathematical model is proposed for the robust version of the problem and a solution method based on Benders decomposition is applied. The approach is tested on a case from NS, and the authors observe that a robust solution with a slightly higher cost than an optimal non-robust solution performs dramatically better in a simulated disrupted scenario.

De Almeida et al. (2008) propose a model for robust rolling stock scheduling at the French railway company SNCF. Their model incorporates two aspects of robustness, limiting delay propagation and the possibility of swapping resources. They also observe that accounting for these aspects of robustness leads to schedules that propagate less delays.

The concept of robust crew planning in passenger railways is addressed by Flier et al. (2008) who study the complexity of the problem and develop algorithms for both creating robust crew schedules and for optimally performing the duty swaps in case of a disruption.

Potthoff (2010) introduces the concept of quasi robustness in the railway crew rescheduling context. It is an approach to rescheduling under uncertainty that considers the crew rescheduling problem under uncertainty as a multi-stage program. Rather than utilizing a stochastic programming approach, the quasi robust approach attempts to ensure that all crews have feasible alternative duties in case the disruption lasts longer than expected.

2.9 Contributions to the literature

This thesis contributes to the literature on rolling stock rescheduling and its applications in disruption management and short-term planning in four ways.

First, we develop several mathematical models for rolling stock rescheduling in Chapter 3. We compare the models and discuss their implementability in practice. We also propose several extensions to incorporate more realistic aspects of the process.

Second, we develop a framework for disruption management of railway resources in Chapter 4. The framework consists of the monitoring of the resources, the detection of conflicts, and ways to represent the uncertainty in the process. The framework further contains models for the online problem of rolling stock rescheduling under uncertainty based on a rolling horizon approach.

Third, we incorporate passenger flows in the rolling stock rescheduling process in Chapter 5 where we suggest an iterative approach to matching seat capacity to passenger demand in an environment where passengers greedily choose their own routes.

Fourth, we study rolling stock rescheduling in the short-term planning phase in Chapter 6. We perform a case study on rescheduling the rolling stock in the event of planned track maintenance.

Chapter 3

Models for Rolling Stock Rescheduling

The Rolling Stock Rescheduling Problem (RSRP) amounts to adapting a given rolling stock schedule to a change in the premises. RSRP has applications in several stages of the planning and operations of a railway system; it is the core problem in the disruption management of railway rolling stock, and can be utilized in the short-term planning as well. In this chapter we formalize the problem and introduce mathematical models for each step in the process. The different approaches are compared and we close the chapter by discussing the applicability of our models in practice.

3.1 The Rolling Stock Rescheduling Problem

Informally, RSRP can be stated as a combinatorial optimization problem where the current rolling stock schedule is to be adjusted to a new situation with a modified timetable, possibly with a different passenger demand, and with a number of new restrictions on rolling stock availability and shunting possibilities. RSRP is the central problem in disruption management of rolling stock, but has also applications in the later stages of the planning process. In Chapters 4, 5 and 6 we utilize RSRP in a number of concrete applications in the disruption management process and in the short-term planning stage. In this section we formalize the input data, the desired output, and the objectives for the problem.

3.1.1 Input data

The input to the problem can be divided into two categories consisting of (*i*) the *current* situation and (*ii*) the *modified* situation to which the schedules of the current situations must be adapted. The current situation contains the following elements:

The current timetable consists of a list of trips. These trips make up the train services that are currently communicated to all crew, local planners and passengers. The current timetable also describes the connections between preceding and succeeding trips.

The current rolling stock schedule is the assignment of rolling stock to the trips in the current timetable. The schedule is given as a list of rolling stock units with corresponding duties.

The modified situation is characterized by a description of the changes in the environment, in the available resources, and in the local shunting capacity. The input for the modified situation is summarized by the following list:

The modified timetable consists of a list of trips. These are the trips that are intended to be served in the modified situation. The modified timetable takes any unavailability of infrastructure into account and is therefore implementable in the modified situation. In practice the modified timetable is constructed by canceling, changing or inserting trips in the current timetable, and therefore has a significant overlap with the current situation. Depending on the rescheduling context, the modified timetable may be constructed from a predetermined disruption handling scenario.

The available rolling stock is a list of rolling stock units that can be used to serve the trips in the modified timetable. Usually these units are the same units as in the current rolling stock schedules. Exceptions are when stand-by units are used or when units are broken down and therefore removed from the circulation.

The shunting possibilities are described in a list of rules describing which shunting operations are allowed at each station.

Forecasted passenger demand is the estimated number of passengers per trip in the modified timetable. Often there is little or no information available on passenger demand in the modified situation in practice. In that case the estimates from the current timetable are used. In Chapter 5 an attempt is made to take the modified demand into account in a more integrated way.

3.1.2 Output

The desired output of RSRP is an assignment of the available rolling stock to the modified timetable. The output should be in the form of a rolling stock duty for each available rolling stock unit. In addition to the hard restrictions for the output, there are several soft requirements as well. These requirements concern the implementability of the rolling stock schedule for the modified timetable, but are considered in the objective rather than as constraints.

Output specifications

The set of duties implies a composition for each trip in the modified timetable. The duties and compositions in the output are allowed if the following conditions hold:

- There are no 'holes' in the compositions: If a unit is assigned to a task in position one in a train and another unit is assigned to a task in position four, then other units must be assigned to tasks in positions two and three. The task in position one is the front of the composition and the remaining positions are numbered increasingly.
- At most one unit is assigned to a task. The order of the units in a composition must be specified and therefore at most one unit may be assigned to each task.
- The composition assigned to a trip must respect certain length restrictions because of the lengths of the platforms of the stations served by the trip.
- The shunting operations implied by the set of duties respect the available shunting possibilities.

3.1.3 Objectives

The objective of RSRP consists of a trade-off between several elements of the overall managerial goals from Section 2.4. The following list summarizes the objectives of RSRP grouped by the corresponding perspectives.

The service perspective

- Minimize the number of canceled trips. There may not be enough rolling stock to run all trips in which case some trips may be canceled in addition to any trips already canceled in the modified timetable.

- Provide sufficient seat capacity according to the operator's policy. Since it is possible, and indeed may be necessary, to insert rolling stock with less capacity than the policy prescribes we minimize the amount by which a train is undercapacitated. We do this by minimizing the number of *seat shortage kilometers.* This is a kilometer of train travel where one passenger is expected not to be assigned a seat. In other words it is a measure of undercapacity of trains with respect to the given passenger demand.

The efficiency perspective

- Minimize the number of *carriage kilometers.* A carriage kilometer is a kilometer of train travel with one carriage. Carriage kilometers are directly connected to the cost of the rolling stock plan and should therefore be minimized.

The system perspective

- Minimize the number of off-balances. The importance of this objective depends on the rescheduling horizon, but in general the rescheduled rolling stock has to fit into the existing rolling stock schedules that are beyond the horizon of the current rescheduling process.

The process perspective

- Minimize the number of changes to the shunting plans. This is a process related soft requirement for keeping the shunting plans intact as far as possible. Shunting is locally planned and there may not be full knowledge available on the possibilities for locally changing the details of the shunting operations. Therefore, it is desirable to minimize these changes.

- Minimize the changes to the rolling stock duties. It eases the practical rescheduling process if there is only a limited number of changes in the new duties.

Clearly, all objectives cannot be minimized at the same time. We therefore use weights for each element in the objective to allow the decision maker to express the trade-off between the different elements.

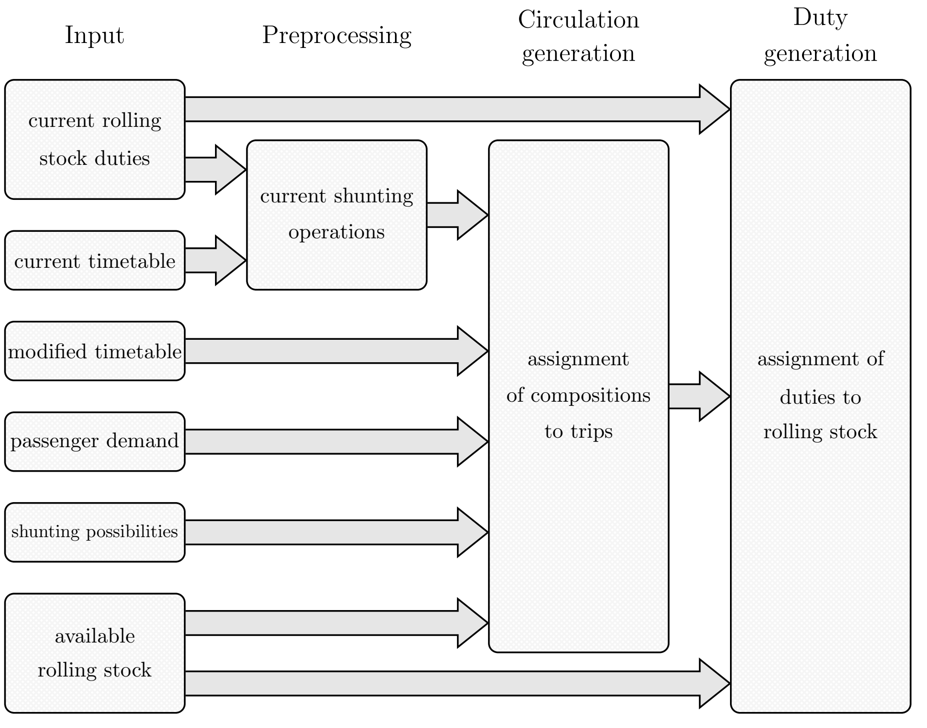


Figure 3.1: Solution framework for RSRP.

3.2 Solution framework

We here suggest a solution framework based on mathematical modeling for solving RSRP. The approach is sketched in Figure 3.1 and involves three steps; preprocessing, circulation generation and duty generation. We also present an integrated approach where circulation and duties are created in one step.

3.2.1 Preprocessing

In the preprocessing step we analyze the current timetable and rolling stock duties to identify all shunting operations. The analysis is performed by examining each pair of consecutive tasks in a duty and determining whether the tasks belong to connected trips in the timetable. If the tasks do not belong to connected trips we know that the rolling stock unit was uncoupled after the first task and coupled before the second task. Similarly, a unit is always coupled before its first task and uncoupled after its last task.

After the analysis in the preprocessing step, we know at which stations and between which trips shunting is performed. This information is needed since it is part of the overall objective to minimize the changes to the shunting plans. Although it is not a trivial step, the preprocessing itself is uninteresting from a modeling point of view since it involves only data analysis. We therefore treat the current shunting operations as if they were given as input.

3.2.2 Circulation generation

In the circulation generation step we construct a circulation for the modified timetable. The circulation consists of an assignment of compositions to trips and the choice of composition changes at the connections between consecutive trips.

We propose two different models for this step; the *Composition Model* and the *Task Model*. The Composition Model is an extension of the model by Fioole et al. (2006) which is based on the compositions assigned to the trips in the modified timetable. We describe the Composition Model in Section 3.4. The Task Model is based on the flow of rolling stock units trough the potential tasks for consecutive trips. We describe the Task Model in Section 3.5.

Most of the RSRP objectives can be taken into account in this step. The requirement of minimizing the changes to the rolling stock duties cannot be taken directly into account in the circulation generation. However, it is possible to encourage features of the circulation that are likely to provide less changes to the structure of the duties. We discuss how to extend the Composition Model to take this aspect into account in Section 3.4.7.2.

3.2.3 Duty generation

In the duty generation step we translate the circulation from the previous step to a set of duties and assign each available rolling stock unit to a duty. We present two models for duty generation; the *Duty Flow Model* and the *Duty Path Model*. The Duty Flow Model is based on the flow of rolling stock units in a graph induced by the circulation. The Duty Path Model is an extension of the previous model where we specifically enumerate all potential duties in the graph.

This step relies on the circulation from the previous step, the current rolling stock duties and the available rolling stock. In this step we take the requirement of minimizing the changes to the structure of the duties into account. We discuss the models for the duty generation process and how to quantify the result in Section 3.6.

Further we propose a model that integrates the duty generation step in the circulation generation in Section 3.7. The model is an extension of the Task Model and is called the Duty-Task Model.

3.2.4 Notations

Throughout this chapter we use the following notation. We note that the objects described in this section were introduced and discussed in Section 2.5. Let $\mathcal{T}$ be the set of trips in the modified timetable. The trips are linked through connections, that indicate the relation between consecutive trips. We denote the set of connections in the modified timetable by $\mathcal{C}$. Each trip $t \in \mathcal{T}$ is preceded by its predecessor connection, $\pi(t)$, and succeeded by its successor connection, $\sigma(t)$. A connection thus has one of the following forms:

- It has exactly one incoming trip and one outgoing trip.
- It consists of an incoming trip that is *split* into two outgoing trips.
- It consists of two incoming trips that are *combined* into an outgoing trip.
- It consists of an outgoing trip and no incoming trips. It thus denotes a train service whose rolling stock is taken from the inventory where it has been stored for some time.
- It consists of an incoming trip and no outgoing trips. It thus denotes a train service whose rolling stock is sent to the inventory and is not to be used in the immediate future.

A connection $c \in \mathcal{C}$ thus has zero, one or two incoming trips; we denote those trips by the set IN_c. Similarly the connection has zero, one or two outgoing trips; we denote those by OUT_c. Even though splitting into more than two trips never occurs in the instances we encountered, it is possible to model splitting into more trips by replacing an outgoing trip by a dummy trip that departs and arrives at the same station and is immediately split upon arrival. The same applies for combining more than two trips.

Let $\mathcal{M}$ denote the set of rolling stock types and let $\mathcal{S}$ denote the set of stations. Then $s(c) \in \mathcal{S}$ denotes the station at which connection c takes place.

A composition is a string of rolling stock types $m_k m_{k-1} \cdots m_1$ that denotes the rolling stock type in each position in the train. Recall that by convention the driving direction of a composition is from left to right which means that m_1 is the rolling

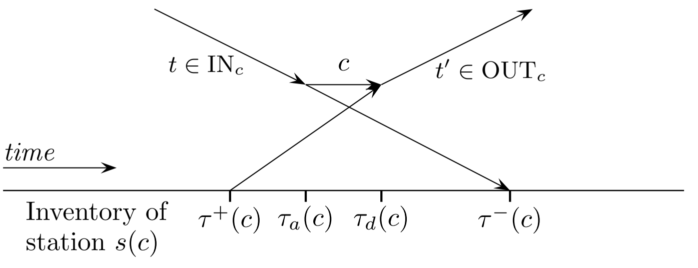


Figure 3.2: Schematic view of a connection c.

stock type in position 1 in the composition. The set of all possible compositions is denoted by $\mathcal{P}$. For a composition $p \in \mathcal{P}$ we denote by $\nu_m(p)$ the number of units of type $m \in \mathcal{M}$ in the composition. For each trip $t \in \mathcal{T}$, $\eta(t)$ denotes the set of allowed compositions on the trip.

A composition change encodes the composition on each incoming and outgoing trip and which units are uncoupled from each incoming trip along with which units are coupled to each outgoing trip. For connection $c \in \mathcal{C}$, $\varrho(c)$ denotes the set of composition changes allowed at the connection, i.e.

$$\varrho(c) \subseteq \prod_{t \in \mathrm{IN}_c} \eta(t) \times \prod_{t \in \mathrm{OUT}_c} \eta(t)$$

For a given connection, $c \in \mathcal{C}$, and an allowed composition change $q \in \varrho(c)$, $p_{q,t}$ denotes the composition of the incoming trip $t \in \mathrm{IN}_c$ and $p'_{q,t}$ denotes the composition of the outgoing trip $t \in \mathrm{OUT}_c$. This way the composition changes encode the shunting possibilities between incoming and outgoing trips.

For a given connection $c \in \mathcal{C}$ and a composition change $q \in \varrho(c)$, let $\alpha_{q,m}$ and $\beta_{q,m}$ denote how many units are uncoupled and coupled respectively of type m in composition change q. For $c \in \mathcal{C}$, the time at which the first incoming trip arrives is denoted by $\tau_a(c)$ and the time at which the last outgoing trip departs is denoted by $\tau_d(c)$. If the connection either has no incoming or no outgoing trips, then we define $\tau_a(c) = \tau_d(c)$. Furthermore, the time at which coupling takes place is denoted by $\tau^+(c)$, and the time at which an uncoupled unit is available after uncoupling is denoted by $\tau^-(c)$. This way it is possible to account for the time it takes to perform the actual (un)coupling and move the units to/from the inventory.

A schematic view of a connection c is shown in Figure 3.2. Here $t \in \mathrm{IN}_c$ and $t' \in \mathrm{OUT}_c$ are an incoming trip and an outgoing trip respectively. The time line shows the arrival time $\tau_a(c)$ and the departure time $\tau_d(c)$, as well as the coupling time $\tau^+(c)$ and the uncoupling time $\tau^-(c)$.

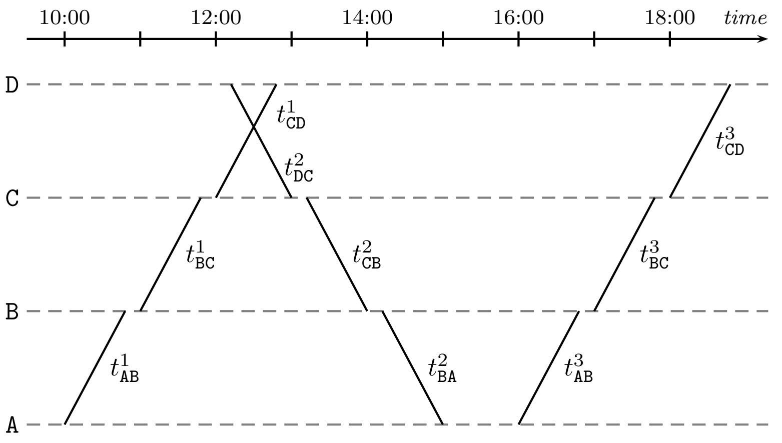


Figure 3.3: Example timetable consisting of three train services.

The rolling stock circulation is connected to the circulation in the earlier and later planning periods through the inventories. The available number of units of type $m \in \mathcal{M}$ at station $s \in \mathcal{S}$ at the beginning of the planning period is given by the parameter $i^0_{s,m}$ and the desired number at the end of the planning period is given by the parameter $i^\infty_{s,m}$.

3.3 Circulation generation

In the circulation generation step we assign a composition to every trip in the modified timetable and a composition change to each connection. We here show an example of the circulation generation process. Consider the timetable in Figure 3.3. It connects four stations, `A`, `B`, `C` and `D` through three train services; two from `A` to `D` and one from `D` to `A` calling at every intermediate station. The trip in the train service 1 from `A` to `B` is named $t^1_{\texttt{AB}}$, the other trips are named similarly. The departure and arrival times of the trips are derived from the time axis in the diagram. There is a connection between trip $t^1_{\texttt{AB}}$ and $t^1_{\texttt{BC}}$ at station `B`, and similarly there is a connection between trip $t^1_{\texttt{BC}}$ and $t^1_{\texttt{CD}}$ at station `C`. Similar patterns exist for services 2 and 3. At stations `B` and `C` it is possible to couple a unit to the front or to uncouple a unit from the rear of a through going train. At terminal stations `A` and `D`, departing trains retrieve their units from the inventory and arriving trains send their units to the inventory.

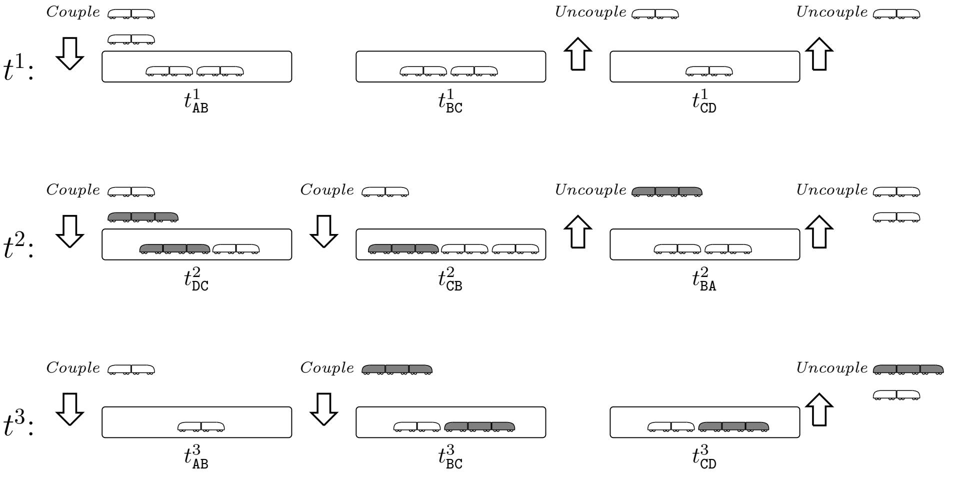


Figure 3.4: Assignment of compositions to trips and composition changes to connections.

Suppose two units with two carriages are available at station A, and one unit with two carriages and one unit with three carriages are available at station D at the beginning of the planning period. Further, we want one unit with two carriages to end at station A, and two units with two carriages and one unit with three carriages to end at station D. The problem is then how to assign the available rolling stock units to the trips in a way that respects the shunting possibilities as well as the rolling stock balances.

Figure 3.4 shows a possible result of the circulation generation step. Each row shows the compositions assigned to the trips in a train service and the composition changes that take place between trips. The units assigned to the first trip in a train service are coupled from the inventory and the units assigned to the last trip are uncoupled to the inventory. Between trips, units may be coupled or uncoupled.

In Section 3.4 we describe a model for circulation generation and discuss a number of extensions of the model. Later, in Section 3.5 we present an alternative model for the problem before returning to this example in Section 3.6 where we discuss the duty generation step.

3.4 Composition Model

Here we present a model for solving the circulation generation step in the RSRP solution approach. The model is an extension of the model by Fioole et al. (2006). It is a MIP model based on a multi commodity network flow with additional features for modeling the restricted shunting possibilities. The core decisions concern which compositions to assign to the trips and which composition changes to apply to the connections.

The description of the model is structured as follows. We first set up a basic model for circulation generation and then explain how each rescheduling option is taken into account.

3.4.1 Variables

The model incorporates a number of variables for the assignment of compositions to trips and the choice of composition changes at connections. The following variables are used in the model:

- $X_{t,p} \in \{0,1\}$ denotes whether composition $p \in \eta(t)$ is used for trip $t \in \mathcal{T}$.
- $Z_{c,q} \in \{0,1\}$ denotes whether composition change $q \in \varrho(c)$ is used for connection $c \in \mathcal{C}$.
- $I_{c,m} \in \mathbb{Z}_+$ denotes the number of units of type $m \in \mathcal{M}$ in the inventory at station $s(c)$ at time $\tau^+(c)$.
- $I^0_{s,m}$ and $I^\infty_{s,m} \in \mathbb{Z}_+$ denote the number of units of type $m \in \mathcal{M}$ at station $s \in \mathcal{S}$ at the beginning of the planning period and at the end of the planning period respectively.
- $C_{c,m}$ and $U_{c,m} \in \mathbb{Z}_+$ denote the number of units of type $m \in \mathcal{M}$ that are coupled and uncoupled respectively during connection c.

The integrality of some variables can be relaxed in the model since their integrality is ensured by the constraints of the model. We discuss this shortly.

3.4.2 Mathematical model

The basic model for the circulation generation process is formulated as follows.

$$\min f(X, Z, I) \tag{3.1}$$

subject to

$$\sum_{p\in\eta(t)} X_{t,p} = 1 \qquad \forall t\in\mathcal{T} \tag{3.2}$$

$$X_{t,p} = \sum_{q\in\varrho(\sigma(t)):p_{q,t}=p} Z_{\sigma(t),q} \qquad \forall t\in\mathcal{T}, p\in\eta(t) \tag{3.3}$$

$$X_{t,p} = \sum_{q\in\varrho(\pi(t)):p'_{q,t}=p} Z_{\pi(t),q} \qquad \forall t\in\mathcal{T}, p\in\eta(t) \tag{3.4}$$

$$C_{c,m} = \sum_{q\in\varrho(c)} \beta_{q,m} Z_{c,q} \qquad \forall c\in\mathcal{C}, m\in\mathcal{M} \tag{3.5}$$

$$U_{c,m} = \sum_{q\in\varrho(c)} \alpha_{q,m} Z_{c,q} \qquad \forall c\in\mathcal{C}, m\in\mathcal{M} \tag{3.6}$$

$$I_{c,m} = I^0_{s(c),m} - \sum_{\substack{c'\in\mathcal{C}:s(c')=s(c),\\ \tau^+(c')\le\tau^+(c)}} C_{c',m} + \sum_{\substack{c'\in\mathcal{C}:s(c')=s(c),\\ \tau^-(c')\le\tau^+(c)}} U_{c',m} \qquad \forall c\in\mathcal{C}, m\in\mathcal{M} \tag{3.7}$$

$$I^\infty_{s,m} = I^0_{s,m} - \sum_{\substack{c\in\mathcal{C}:\\ s(c)=s}} C_{c,m} + \sum_{\substack{c\in\mathcal{C}:\\ s(c)=s}} U_{c,m} \qquad \forall s\in\mathcal{S}, m\in\mathcal{M} \tag{3.8}$$

$$I^0_{s,m} = i^0_{s,m} \qquad \forall s\in\mathcal{S}, m\in\mathcal{M} \tag{3.9}$$

$$I^\infty_{s,m} = i^\infty_{s,m} \qquad \forall s\in\mathcal{S}, m\in\mathcal{M} \tag{3.10}$$

$$X_{t,p} \in \{0,1\} \qquad \forall t\in\mathcal{T}, p\in\eta(t) \tag{3.11}$$

$$C_{c,m}, U_{c,m}, I_{c,m} \in \mathbb{R}_+ \qquad \forall c\in\mathcal{C}, m\in\mathcal{M} \tag{3.12}$$

$$I^0_{s,m}, I^\infty_{s,m} \in \mathbb{R}_+ \qquad \forall s\in\mathcal{S}, m\in\mathcal{M} \tag{3.13}$$

$$Z_{c,q} \in \mathbb{R}_+ \qquad \forall c\in\mathcal{C}, q\in\varrho(c) \tag{3.14}$$

Constraints (3.2) specify that for each trip exactly one allowed composition should be chosen. Constraints (3.3) constitute flow conservation constraints for the allowed compositions and the composition changes for the succeeding connection. Thus for a trip t an allowed composition, $p\in\eta(t)$, is assigned if and only if an allowed composition change $q\in\varrho(\sigma(t))$ is chosen for the successor connection, and that composition change must have p assigned to incoming trip t. In the same way constraints (3.4) ensure flow conservation for trips and their predecessor connection.

Constraints (3.5) state that the number of units of type m coupled at connection c equals the number of units of that type coupled at the composition change assigned to that connection. Constraints (3.6) work similarly for uncoupling.

Constraints (3.7) state that at the time of coupling at connection c, the number of units of type m available at station $s(c)$ equals the number of units available at the same station at the beginning of the planning period, $I^0_{s(c),m}$, minus the number of units coupled at other connections before this one plus the number of units uncoupled to the inventory of the station.

Constraints (3.8) add up the number of units at a station at the end of the planning period. Constraints (3.9) and (3.10) denote that the number of units of each type at each station at the beginning and at the end of the planning period is fixed.

Constraints (3.11) state the integrality of the composition variables. Constraints (3.12) – (3.14) specify the non-negativity of the remaining variables. According to Fioole et al. (2006) these variables can be chosen to be continuous.

The core of the model involves the decisions on what compositions are assigned to each trip. The model then links the compositions on consecutive trips through the possible composition changes at the stations and the units in the inventory there.

Another way to view the model is to consider the *composition graph*. In this graph compositions are represented by nodes and the possible composition changes are represented by arcs between compositions on consecutive trips. Figure 3.5 shows an example composition graph for a chain of trips. The chain starts in Amsterdam (Asd). In Alkmaar (Amr) it is possible to couple a unit to the front of the train or to uncouple a unit from the rear of the train. The train is turned in Den Helder (Hdr). An allowed assignment of compositions to trips for a train is a resource-constrained path in the graph. The path is resource-constrained due to the availability of units in the inventory. Note that splitting and combining is represented by hyper-arcs with two heads or two tails respectively. Therefore an allowed assignment is even more complex than just a path for train services that involve splitting or combining.

3.4.3 Objective function

The objective (3.1) is a function of the compositions assigned to the trips and the composition changes used at the connections. By applying appropriate penalties for these variables, it is possible to account for the objectives of RSRP. For some objectives we need to extend the basic Composition Model with more variables and constraints.

We describe in the following sections how this is performed. In Section 3.4.4 we describe how to account for the efficiency related objective which here refers to carriage kilometers. Thereafter, we add the service related objectives in Section 3.4.5.

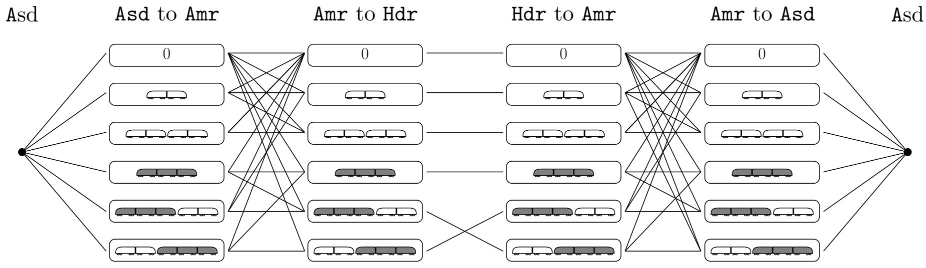


Figure 3.5: Composition graph for train service from Amsterdam through Alkmaar to Den Helder and back.

These consist of cancellations of trips and seat shortage kilometers. Lastly, we describe the process related objectives in Section 3.4.6. These objectives relate to the changes to the shunting plans, the off-balances, and the conductors.

3.4.4 Objectives for the efficiency perspective

The efficiency of the rolling stock schedule in the modified situation is measured by the number of carriage kilometers. This is taken into account in the objective function as the number of carriage kilometers is a linear function of the $X_{t,p}$-variables. Let $\texttt{km}(t)$ be the length of trip t in kilometers and let $\texttt{carr}(p)$ be the number of carriages in composition p. Then the carriage kilometers are taken into account with weight v by adding the following term to the objective function.

$$\sum_{t \in \mathcal{T}} \sum_{p \in \eta(t)} v \cdot \texttt{carr}(p) \cdot \texttt{km}(t) \cdot X_{t,p} \tag{3.15}$$

3.4.5 Objectives for the service perspective

The service perspective of the rolling stock schedule in the modified situation consists of two main elements; the number of canceled trips and the seat shortages compared to demand. We here show how to take these elements into account in the objective function of the Composition Model. Further, we suggest an alternative approach to assigning adequate capacity in the modified situation when insufficient information on passenger demand is available.

Cancellation of trips

In the rescheduling process it is an option to cancel trips due to rolling stock shortages. This is modeled in the following way. Canceling a trip is equivalent to assigning no rolling stock to it. Therefore, for a trip $t \in \mathcal{T}$ that may be canceled, we add a special *empty composition* '0' to the set of allowed compositions $\eta(t)$. Furthermore we add the appropriate composition changes to the set of possible composition changes $\varrho(\pi(t))$ and $\varrho(\sigma(t))$ at the predecessor and successor connections. We then impose a penalty of $w_{t,0}$ for assigning the empty composition to trip t by adding term (3.16) to the objective function.

$$\sum_{t \in \mathcal{T}} w_{t,0} X_{t,0} \tag{3.16}$$

An important note on this method for modeling cancellations is that the composition of the train is split into units and sent to the inventory before the cancellation. From there the units may be coupled to other trains without taking the order of units in the original composition into account. This may not always be implementable in practice.

Seat shortage kilometers

A seat shortage kilometer denotes one kilometer of train ride with a deficit of one seat compared to the forecasted passenger demand. Seat shortage kilometers can be directly accounted for as a linear function of the variables $X_{t,p}$. If d_t is the forecasted passenger demand on trip t then the expected seat shortage on the trip is equal to $\max\{d_t - \mathtt{cap}(p), 0\}$ when assigning composition $p \in \eta(t)$ with capacity $\mathtt{cap}(p)$. The number of seat shortage kilometers is thus $\max\{d_t - \mathtt{cap}(p), 0\} \cdot \mathtt{km}(t)$ where $\mathtt{km}(t)$ is the length of the trip in kilometers. We add the term (3.17) to the objective function.

$$\sum_{t \in \mathcal{T}} \sum_{p \in \eta(t)} w \cdot \max\{d_t - \mathtt{cap}(p), 0\} \cdot \mathtt{km}(t) \cdot X_{t,p} \tag{3.17}$$

Here w is the weight of the seat shortage kilometers in the objective. We note that this linear cost may be inappropriate for penalizing large seat shortages. In such a case a non-linear measure may be more appropriate.

Minimizing the deviation in assigned capacity

If there is no accurate forecast of passenger demand available in the modified situation, planners often resolve to inserting seat capacity that is similar to the current situation. This has pragmatic reasons as the planned rolling stock circulation is optimized with respect to matching capacity and demand. Therefore it is generally preferable to assign similar capacity to the trains in the rolling stock rescheduling phase if inadequate forecasts are available for the modified situation.

This approach is only applicable for trips that exist in both the original and modified timetable. For such trips let $p^*(t)$ be the composition assigned to trip $t \in \mathcal{T}$ in the current rolling stock schedule and let $w_{t,p}$ be the penalty for assigning composition $p \in \eta(t)$ to trip t. The penalty represents the difference in capacity between p and $p^*(t)$. Then we add the term (3.18) to the objective function.

$$\sum_{t \in \mathcal{T}} \sum_{p \in \eta(t)} w_{t,p} X_{t,p} \tag{3.18}$$

There are several natural candidates for the penalty such as the weighted absolute difference in the number of seats, $w_{t,p} = w|\texttt{cap}(p) - \texttt{cap}(p^*(t))|$, or the weighted difference in the number of seats squared, $w_{t,p} = w\left(\texttt{cap}(p) - \texttt{cap}(p^*(t))\right)^2$, where w is an appropriate weight.

3.4.6 Objectives for the process perspective

The process perspectives of the objectives for RSRP can be accounted for in the objective function for the Composition Model by adding a few variables and constraints. We here describe how to minimize the changes to the shunting plans, the number of off-balances, and the implicit changes to the conductor schedules.

Changing the current shunting operations

As explained earlier, it is an important goal of the current rescheduling process to minimize the changes to the current shunting operations. Consider a connection $c \in \mathcal{C}$. For each allowed composition change $q \in \varrho(c)$ it can be derived whether it requires any shunting, and if it does, which types of units are coupled or uncoupled, and to or from which positions. By comparing the allowed composition changes in the modified situation to the shunting operations in the current rolling stock schedules it can be determined whether q incurs a new shunting movement or changes the nature of an existing shunting movement (e.g. coupling instead of uncoupling). It is then possible to penalize the composition changes that incur larger modifications to the planned shunting operations by adding the term (3.19) to the objective function.

$$\sum_{c \in \mathcal{C}} \sum_{q \in \varrho(c)} \gamma_{c,q} Z_{c,q} \tag{3.19}$$

Here $\gamma_{c,q}$ denotes the penalty for using composition change $q \in \varrho(c)$ at connection $c \in \mathcal{C}$. The value of $\gamma_{c,q}$ depends on the implied changes to the shunting operations. Recall that the possible changes to the shunting plans and their general consequences are explained in Section 2.6.1 and illustrated in Figure 2.10. In general, changes to the shunting plans that are more complicated are penalized more strictly.

Deviations in the target inventories

In order to gain more flexibility in the rescheduling process we allow for some units to end their duty at a different station than originally planned. Generally it is not a problem that a unit ends its duty at a different station, but the number of units of the same type at a station must be the same at the end of the planning period. In other words, planners prefer having no off-balances. However, off-balances may be necessary to provide sufficient capacity on some trains or due to restricted shunting possibilities.

In the basic composition model we use a fixed target inventory of $i^{\infty}_{s,m}$ at station s for units of type m. Here we relax this fixed target by adding additional variables to the model and replacing constraints (3.10). Let variables $D^{\infty,+}_{s,m}$ and $D^{\infty,-}_{s,m} \in \mathbb{R}_+$ denote the number of units of type $m \in \mathcal{M}$ that are in excess and shortage respectively at station $s \in \mathcal{S}$ at the end of the planning period. The variables are added to the model with the constraints (3.20).

$$I^{\infty}_{s,m} = i^{\infty}_{s,m} + D^{\infty,+}_{s,m} - D^{\infty,-}_{s,m} \qquad \forall s \in \mathcal{S}, m \in \mathcal{M} \tag{3.20}$$

With the introduction of the deviation variables and constraints (3.20), we allow some flexibility in the target inventories by penalizing the variables in the objective function through the term (3.21).

$$\sum_{s\in\mathcal{S}} \sum_{m\in\mathcal{M}} \left(d^{\infty,+}_{s,m} D^{\infty,+}_{s,m} + d^{\infty,-}_{s,m} D^{\infty,-}_{s,m} \right) \tag{3.21}$$

Here $d^{\infty,+}_{s,m}$ denotes the cost of having an extra unit of type m at station s at the end of the planning period and similarly $d^{\infty,-}_{s,m}$ denotes the cost of having one less unit of type m at station s.

For some rolling stock types with similar characteristics, it is less urgent to have the correct number of units of each type at the end of the planning period. Although having exactly the planned number of units per type is preferable, it may suffice to aim at having the planned total number of units with similar characteristics available.

For each relevant subset of rolling stock types, $M \subseteq \mathcal{M}$, we define variables $A^{\infty,+}_{s,M}$ and $A^{\infty,-}_{s,M} \in \mathbb{R}_+$ that represent the cumulative deviation of units of type $m \in M$ at station s at the end of the planning period. The constraints (3.22) are added to the model.

$$\sum_{m\in M} I^{\infty}_{s,m} = \sum_{m\in M} i^{\infty}_{s,m} + A^{\infty,+}_{s,M} - A^{\infty,-}_{s,M} \qquad \forall s \in \mathcal{S} \tag{3.22}$$

The weighted number of cumulative deviations for a set M of rolling stock types with similar characteristics can be minimized by adding the term (3.23) to the objective function.

$$\sum_{s\in\mathcal{S}} \left(a^{\infty,+}_{s,M} A^{\infty,+}_{s,M} + a^{\infty,-}_{s,M} A^{\infty,-}_{s,M} \right) \tag{3.23}$$

Here the cost of cumulative deviations in the planned inventories at the end of the planning period is defined as $a^{\infty,+}_{s,M}$ and $a^{\infty,-}_{s,M}$ for each set of rolling stock types $M \subseteq \mathcal{M}$ with similar characteristics.

We note that if $d^{\infty,+}_{s,m}$ and $d^{\infty,-}_{s,m} \geq 0$ there exists an optimal solution with $D^{\infty,+}_{s,m} + D^{\infty,-}_{s,m} = |I^{\infty}_{s,m} - i^{\infty}_{s,m}|$ which implies $D^{\infty,+}_{s,m} = |I^{\infty}_{s,m} - i^{\infty}_{s,m}|$ or $D^{\infty,-}_{s,m} = |I^{\infty}_{s,m} - i^{\infty}_{s,m}|$. The same holds for the cumulative deviation variables.

Changing the number of conductors needed

The number of carriages in a train determines the number of conductors needed on the train. Longer trains require a higher number of conductors. It is preferable not to change the number of conductors needed on the trips as personnel is a resource that needs to be planned separately. Since each composition p needs a minimum number of conductors, the penalty for assigning a composition which needs more conductors to a trip can be incorporated in the objective function. This is achieved by adding the term $w \cdot \max\{0, \texttt{con}(p) - \texttt{con}(p^*(t))\}$ where $\texttt{con}(p)$ denotes the number of conductors needed to operate composition p.

Note that the actual rescheduling of the conductors is performed separately, the penalties above are only meant to promote solutions with less need for extra conductors.

3.4.7 Additional features

In addition to the objectives described in the preceding sections, there is a number of features that further increase the practical applicability of the Composition Model. First, we describe in Section 3.4.7.1 how to include the option of using flexible turning patterns at a terminal station. Second, in Section 3.4.7.2 we suggest an extension to the model that stimulates a duty structure that resembles that of the current duties. Third, in Section 3.4.7.3 we discuss how to ensure that at least one unit is assigned from the start station to the terminal station of a train service.

3.4.7.1 Flexible turning patterns at terminal stations

The described model uses connections between trips at the terminal stations to enforce the planned turning pattern. Upholding a rigid turning pattern makes sense in the planning phase as it eases the task of local planning by continuously repeating turning schemes. However, as discussed in Section 2.6.1, it may be too restrictive to uphold the planned turning patterns in the rescheduling process.

Flexibility in the turning patterns at a station is here added to the model by introducing the following changes to the basic model. The basic idea is to introduce variables that keep track of the compositions that are available at the platforms of the station. We call the available compositions the *pool of compositions* at the station. Whenever a train departs from the station it can use a composition from the pool of compositions, and whenever a train arrives at the station it may add a composition to the station's pool of compositions. We ensure that at any time no more compositions are stored at the platforms than the number of platforms available.

The idea is shown schematically in Figure 3.6. Here an arriving trip t may use its connection $c = \sigma(t)$ which is linked to departing trip t'. Or the arriving trip may send a composition to the pool of compositions while the departing trip takes a composition from the pool. Units coupled or uncoupled in either case are taken from or sent to the inventory as usual.

In this section we consider a specific station s where flexible turning patterns are used. For this fixed s we define two subsets of trips $\mathcal{T}_1, \mathcal{T}_2 \subseteq \mathcal{T}$, where $\mathcal{T}_1$ denotes the trips that arrive at the station s, i.e. $\mathcal{T}_1 = \{t \in \mathcal{T} \mid s(\sigma(t)) = s\}$, and similarly $\mathcal{T}_2$ denotes the trips that depart from the station; $\mathcal{T}_2 = \{t \in \mathcal{T} \mid s(\pi(t)) = s\}$. For a trip $t \in \mathcal{T}_1$ and a composition p, we denote by $\mathrm{UN}_{p,t}$ the set of allowed compositions that may result from uncoupling zero or more units from the composition p at the arrival station of trip $t \in \mathcal{T}_1$. In the same way, we denote by $\mathrm{CO}_{p,t}$ the set of compositions from which one can construct p by coupling zero or more units at the departure station of trip $t \in \mathcal{T}_2$.

We introduce variables $Y_{t,p,p'} \in \{0,1\}$ for the relevant incoming trips $t \in \mathcal{T}_1$, compositions $p \in \eta(t)$ and compositions after uncoupling, $p' \in \mathrm{UN}_{p,t}$. These variables denote that the train assigned to trip t arrives with composition p and leaves composition p' at the platform, the uncoupled units are sent to the inventory. In

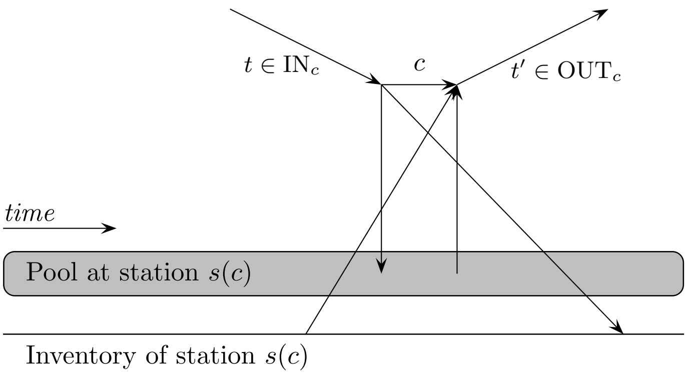


Figure 3.6: Extension of the chart in Figure 3.2. In addition to the inventory we use the pool to keep track of the compositions parked at the platforms.

the same way we introduce variables $W_{t,p,p'} \in \{0,1\}$ for the relevant outgoing trips $t \in \mathcal{T}_2$, compositions $p \in \eta(t)$ and compositions before coupling, $p' \in \text{CO}_{p,t}$. These variables denote that the train assigned to trip t departs with composition p and takes composition p' from the platform. Any units coupled are taken from the inventory.

For trips in $\mathcal{T}_1$ and $\mathcal{T}_2$ respectively, replace constraints (3.3) and (3.4) with constraints (3.24) and (3.25):

$$X_{t,p} = \sum_{q \in \varrho(\sigma(t)):p_{q,t}=p} Z_{\sigma(t),q} + \sum_{p' \in \text{UN}_{p,t}} Y_{t,p,p'} \qquad \forall t \in \mathcal{T}_1, p \in \eta(t) \tag{3.24}$$

$$X_{t,p} = \sum_{q \in \varrho(\pi(t)):p'_{q,t}=p} Z_{\pi(t),q} + \sum_{p' \in \text{CO}_{p,t}} W_{t,p,p'} \qquad \forall t \in \mathcal{T}_2, p \in \eta(t) \tag{3.25}$$

Constraints (3.24) ensure that for an incoming trip, either the original connection is used or part of the train is sent to the pool of compositions at the station. Similarly, for an outgoing trip, constraints (3.25) ensure that either the original connection is used or part of the composition is taken from the pool of compositions at the station.

As some trips do not use their predetermined connections but rather use the pool of compositions at station s, we also need to account for the coupled and uncoupled units in the variables $U_{c,m}$ and $C_{c,m}$ by extending constraints (3.5) and (3.6). We denote the connections that take place at the station in question by $\tilde{\mathcal{C}}$. We replace

(3.5) and (3.6) with (3.26) and (3.27) for those connections.

$$C_{c,m} = \sum_{q\in\varrho(c)} \beta_{q,m} Z_{c,q} + \sum_{t\in\mathrm{OUT}_c} \sum_{p\in\eta(t)} \sum_{p'\in\mathrm{CO}_{p,t}} \delta_m(p,p') W_{t,p,p'} \qquad \forall c \in \tilde{\mathcal{C}}, m \in \mathcal{M} \tag{3.26}$$

$$U_{c,m} = \sum_{q\in\varrho(c)} \alpha_{q,m} Z_{c,q} + \sum_{t\in\mathrm{IN}_c} \sum_{p\in\eta(t)} \sum_{p'\in\mathrm{UN}_{p,t}} \delta_m(p,p') Y_{t,p,p'} \qquad \forall c \in \tilde{\mathcal{C}}, m \in \mathcal{M} \tag{3.27}$$

Here the parameter $\delta_m(p_1,p_2) = \nu_m(p_1) - \nu_m(p_2)$ denotes the difference in the number of units of type m between composition p_1 and p_2. With the extensions in (3.26) and (3.27) we account for the units uncoupled from incoming trips whose composition is sent to the pool, and the units coupled to outgoing trips whose composition is taken from the pool.

To keep track of the number of compositions available at the platform we introduce variables $P_{c,p} \in \mathbb{Z}_+$ for $c \in \tilde{\mathcal{C}}$ who denote the number of trains with composition p that are parked at the platforms of station $s(c)$ at time $\tau_d(c)$. The constraints (3.28) are added to the model.

$$P_{c,p} = \sum_{\substack{c'\in\tilde{\mathcal{C}}:s(c')=s(c),\\ \tau_a(c')<\tau_d(c)}} \sum_{t\in\mathrm{IN}_{c'}} \sum_{p':p\in\mathrm{UN}_{p',t}} Y_{t,p',p} - \sum_{\substack{c'\in\tilde{\mathcal{C}}:s(c')=s(c),\\ \tau_d(c')<\tau_d(c)}} \sum_{t\in\mathrm{OUT}_{c'}} \sum_{p':p\in\mathrm{CO}_{p',t}} W_{t,p',p} \qquad \forall c \in \tilde{\mathcal{C}}, p \in \mathcal{P} \tag{3.28}$$

These constraints state that the number of compositions, that are like p, stored at the platforms of station $s(c)$ at the departure time of connection c is equal to the number of such compositions that have been deposited there up to time $\tau_d(c)$ minus the number that have been taken from there in the same period in time. It is here assumed that the pool is initially empty, but the constraints (3.28) can easily be extended to account for initial compositions.

At the end of the planning period there may be compositions parked at the platform. We assume that they are added to the inventory of the station. This is achieved by extending constraints (3.8) to (3.29) for the relevant station s.

$$\begin{aligned} I^{\infty}_{s,m} = I^{0}_{s,m} &- \sum_{\substack{c\in\mathcal{C}:\\ s(c)=s}} C_{c,m} + \sum_{\substack{c\in\mathcal{C}:\\ s(c)=s}} U_{c,m} \\ &+ \sum_{\substack{c\in\tilde{\mathcal{C}}:\\ s(c)=s}} \sum_{t\in\mathrm{IN}_c} \sum_{p\in\mathcal{P}} \sum_{p'\in\mathrm{UN}_{p,t}} \nu_m(p') Y_{t,p,p'} \\ &- \sum_{\substack{c\in\tilde{\mathcal{C}}:\\ s(c)=s}} \sum_{t\in\mathrm{OUT}_c} \sum_{p\in\mathcal{P}} \sum_{p'\in\mathrm{CO}_{p,t}} \nu_m(p') W_{t,p,p'} \qquad \forall m \in \mathcal{M} \end{aligned} \tag{3.29}$$

With this constraint the compositions left at the platforms are added to the inventory at the end of the planning period.

Finally we add constraints (3.30) to limit the number of compositions at the platforms at any time by the number of available platforms.

$$\sum_{p\in\mathcal{P}} P_{c,p} + \sum_{\substack{c'\in\mathcal{C}:s(c')=s(c),\\ \tau_a(c')\le\tau_d(c)\le\tau_d(c')}} \sum_{q\in\varrho(c')} Z_{c',q} \le \mathtt{plat}(s(c)) \qquad \forall c \in \tilde{\mathcal{C}} \tag{3.30}$$

Note that it suffices to require that the platform capacity is respected immediately before the time of departure since the departure of the train involved in c cannot increase the size of the pool.

The parameter $\mathtt{plat}(s)$ denotes the number of available platforms at station s. Constraints (3.30) ensure that the capacity of the station is not exceeded. The second term on the left hand side ensures that connections that keep the original pattern are also counted for the duration of the dwelling at the platform.

The following constraints state the domains of the variables.

$$P_{c,p} \in \mathbb{Z}_+ \qquad \forall c \in \tilde{\mathcal{C}}, p \in \mathcal{P} \tag{3.31}$$

$$Y_{t,p,p'} \in \{0,1\} \qquad \forall t \in \mathcal{T}_1, p \in \eta(t), p' \in \mathrm{UN}_{p,t} \tag{3.32}$$

$$W_{t,p,p'} \in \{0,1\} \qquad \forall t \in \mathcal{T}_2, p \in \eta(t), p' \in \mathrm{CO}_{p,t} \tag{3.33}$$

If the decision maker prefers to limit the use of flexible turning patterns, a term penalizing their use can be added to the objective function:

$$\sum_{c\in\tilde{\mathcal{C}}} \sum_{t\in\mathrm{IN}_c} \sum_{p\in\mathcal{P}} \sum_{p'\in UN_{p,t}} g_{t,p,p'} Y_{t,p,p'} + \sum_{c\in\tilde{\mathcal{C}}} \sum_{t\in\mathrm{OUT}_c} \sum_{p\in\mathcal{P}} \sum_{p'\in\mathrm{CO}_{p,t}} h_{t,p,p'} W_{t,p,p'}$$

Here $g_{t,p,p'}$ and $h_{t,p,p'}$ are penalty coefficients for particular trips and compositions. It is also possible to include penalties for changing the shunting plans in these parameters.

In Section 3.8.4.1 we perform computational tests of the described extension for flexible turning patterns.

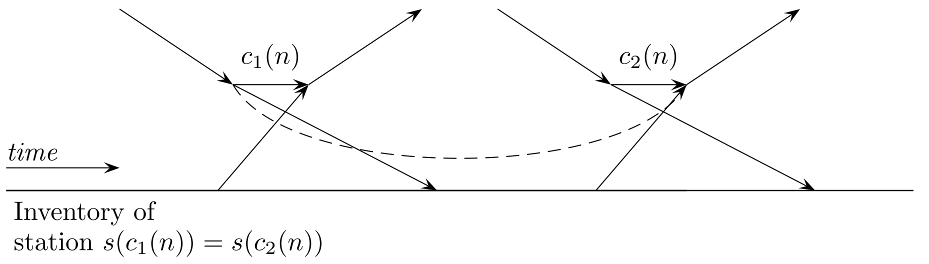


Figure 3.7: Extension of the chart in Figure 3.2. The dashed arc allows the explicit transfer of units between connections in a relation.

3.4.7.2 Keeping the current duty structure

In an attempt to keep some of the structure of the original rolling stock duties, we add the possibility of transferring units between non-connected trips. This is an implicit way to encourage the goal of minimizing the changes to the duties since the actual duties are constructed in the later duty generation phase. From the original set of duties we identify the pairs of connections where a unit was uncoupled at the first connection and coupled at the second connection. Let $\mathcal{N}$ denote the set of such uncoupling-coupling relations. Then the relation $n \in \mathcal{N}$ describes that $r(n)$ units of type $m(n)$ were uncoupled at connection $c_1(n)$ and coupled at connection $c_2(n)$.

The idea is shown schematically in Figure 3.7 where the dashed arc links two connections in an uncoupling-coupling relation. The relations are taken into account by adding variable S_n for each relation $n \in \mathcal{N}$ and replacing (3.5) and (3.6) with the following.

$$U_{c,m} + \sum_{\substack{n \in \mathcal{N}: \\ c_1(n)=c, m(n)=m}} S_n = \sum_{q \in \varrho(c)} \alpha_{q,m} Z_{c,q} \qquad \forall c \in \mathcal{C}, m \in \mathcal{M} \tag{3.34}$$

$$C_{c,m} + \sum_{\substack{n \in \mathcal{N}: \\ c_2(n)=c, m(n)=m}} S_n = \sum_{q \in \varrho(c)} \beta_{q,m} Z_{c,q} \qquad \forall c \in \mathcal{C}, m \in \mathcal{M} \tag{3.35}$$

$$S_n \leq r(n) \qquad \forall n \in \mathcal{N} \tag{3.36}$$

$$S_n \in \mathbb{Z}_+ \qquad \forall n \in \mathcal{N} \tag{3.37}$$

The relation variables thus provide the possibility of keeping an existing pair of first uncoupling and later coupling. The addition of the terms involving S_n in constraints (3.34) and (3.35) allows units to transfer directly between the connections in the relations. Constraints (3.36) ensure that the number of units that are transferred through a relation is limited by the number of units originally using the relation.

Keeping the original relations between uncoupling and later coupling of units can be encouraged by adding a penalty to the objective function for not keeping the uncoupling and later coupling of a specific type of unit intact. Add the following term (3.38) to the objective function.

$$\sum_{n \in \mathcal{N}} s_n \left(r(n) - S_n \right) \tag{3.38}$$

Since the objective is minimized, a penalty of s_n is used for not keeping the relation $n \in \mathcal{N}$.

We note that this approach does not explicitly keep track of the positions in the trains from which units are uncoupled and later coupled. This implies that if several units are uncoupled and later coupled in such a relation then their order may be perturbed between the involved connections. In Section 3.7 we introduce a model where we do keep track of the positions of units used in such relations. We there discuss the difference between the two approaches.

3.4.7.3 Continuity

In the timetable, it is announced that a train service goes all the way to the terminal station. This means that passengers expect to be able to board the train and go all the way to the announced destination without changing carriages underway. However, as we model train services as a series of trips, it may be possible for some sequences of trips to uncouple all the initially assigned units before reaching the terminal station. We therefore require a minimum number of units, say n, to be assigned to the entire sequence of trips. This requirement is referred to as the *continuity requirement* for a sequence of trips. We call a unit that goes all the way from departure station to arrival station a *through-going unit*.

Let $t_1, \ldots, t_k$ be the sequence of trips along the path from s_d to s_a, ordered by their occurrence in the path. Let $c_1, \ldots, c_{k-1}$ be the connections between these trips so that c_i is the connection between trips t_i and t_{i+1}. We keep track of the units that travel with this service from the departure station by adding a number of extra variables.

We note that Maróti (2006) suggests an extension of the Composition Model to account for the continuity requirements. The author keeps track of the indices of the left- and right-most through-going units in the compositions and adds constraints to ensure that the difference between the indices is at least n on trip t_k. The approach is, however, limited to simple line structures. Therefore we here suggest an alternative way to model the continuity constraints. For the instances we encountered during this study, continuity is always implied due to the structure of the involved lines. But until 2006 NS did have practical rolling stock scheduling instances where continuity was not necessarily implied.

The idea is to keep track of not only the compositions assigned to the trips, but also whether the unit in each position has been assigned to the train since trip t_1. Recall that a composition is a string of rolling stock types $m_k m_{k-1} \cdots m_1$. We introduce strings of 0s and 1s where 1 at a certain position denotes that the unit in that position has been in the train since trip t_1. Similarly, a 0 means that the unit has been coupled to the train later in the series of trips. The string 00110 thus indicates that the units in second and third positions have been in the train since trip t_1, whereas the units in positions one, four and five were coupled to the train later.

Let `Strings` be the set of all strings of 0s and 1s up to the appropriate length, and let $\texttt{Start} \subseteq \texttt{Strings}$ be the set of strings that only contains 1s. The set `Start` denotes the strings that may be used on trip t_1 since all units assigned to this trip of course have been assigned since the first trip in the sequence. Let $\texttt{End} \subseteq \texttt{Strings}$ be the set of strings with at least n 1s – one of these strings must be assigned to the last trip t_k.

Let the variables $H_{t_i,r} \in \{0,1\}$ denote that string $r \in \texttt{Strings}$ is assigned to trip t_i. Furthermore, for the connections $c_1, \ldots, c_{k-1}$ we introduce variables that link the strings between consecutive trips; let the variables $G_{c_i,r,r'} \in \{0,1\}$ denote that string r is assigned to trip t_i and r' is assigned to trip t_{i+1}. The idea is then that the variables $G_{c_i,r,r'}$ are enabled by the connection variables $Z_{c_i,q}$. We then add the following constraints for each continuity requirement.

$$H_{t_1,r} = \sum_{\substack{p \in \eta(t_1): \\ \texttt{len}(p)=\texttt{len}(r)}} X_{t_1,p} \qquad \forall r \in \texttt{Start} \tag{3.39}$$

$$H_{t_i,r} = \sum_{r' \in \texttt{Strings}} G_{c_i,r,r'} \qquad \forall i = 1, \ldots, k-1 \tag{3.40}$$

$$H_{t_{i+1},r'} = \sum_{r \in \texttt{Strings}} G_{c_i,r,r'} \qquad \forall i = 2, \ldots, k \tag{3.41}$$

$$G_{c_i,r,r'} \leq \sum_{(r,r') \in \texttt{Allow}(q)} Z_{c_i,q} \qquad \forall i = 1, \ldots, k-1, q \in \varrho(c_i) \tag{3.42}$$

$$\sum_{r \in \texttt{End}} H_{t_k,r} = 1 \tag{3.43}$$

Constraints (3.39) state that a string $r \in \texttt{Start}$ must be assigned to t_1 with the same length as the composition assigned to the trip. These constraints ensure that a string of 1s is assigned to t_i and the string has the same length as the composition assigned to the trip.

Constraints (3.40) denote that if string r is assigned to t_i then a pair of strings (r, r') is assigned to the succeeding connection c_i. Similarly, constraints (3.41) state that if string r' is assigned to t_{i+1} then a pair of strings (r, r') is assigned to the preceding connection.

Constraints (3.42) state that the choice of composition change at connection c_i enables the change of strings (r, r') between t_i and t_{i+1}. Here $\texttt{Allow}(q)$ denotes the set of pairs of strings (r, r') where string r is changed to r' through the composition change q.

3.5 Task Model

In this section we present an alternative model for the circulation generation. The model is based on the flow of rolling stock units through a network constructed from the potential rolling stock tasks.

The idea is to use unit capacity arcs associated with the trips to denote the potential tasks, and unit capacity arcs at the connections to denote the transfer between tasks on consecutive trips. Other unit capacity arcs denote the coupling and uncoupling to and from the inventory at the connections. A unit valued path in this network thus tracks the part of a duty from the inventory of one station to the inventory of another station. The interaction of the units is modeled through constraints. We here describe the network and formulate the mathematical model.

Task- and connection-arcs

Consider a trip $t \in \mathcal{T}$. The rolling stock units assigned to the trip will constitute a number of tasks. There is an upper bound on the length of the train assigned to t which implies that there are at most a certain number of tasks related to the trip. Let $\texttt{tasks}(t)$ be the set of possible tasks in the train that serves trip t. A task k is in position $\texttt{pos}(k)$ in the train that serves trip $\texttt{trip}(k)$. The tasks are filled from the front which means that a unit can only be assigned to task k if another unit is assigned to task k' with position $\texttt{pos}(k') = \texttt{pos}(k) - 1$ for $\texttt{pos}(k) \geq 2$. The tasks can

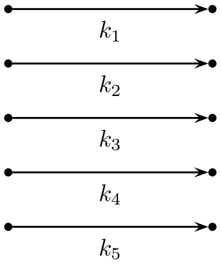


Figure 3.8: Arcs for all tasks $k_1, \ldots, k_5$ belonging to a given trip t.

be regarded as *task-arcs* in a network as shown in Figure 3.8. A flow value of 1 on such an arc indicates that a unit is assigned to the task. We must thus ensure that the arc for task k_i only carries flow if the arc for task k_{i-1} carries flow for $i \geq 2$.

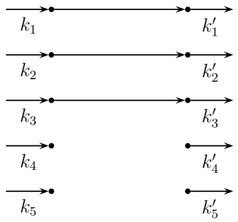
For the connections between trips we add a number of arcs to facilitate the flow of units between tasks: For connection $c \in \mathcal{C}$ we add a *connection-arc* from task k on an incoming trip to task k' on an outgoing trip if there exists a composition change $q \in \varrho(c)$ in which the unit assigned to task k continues with task k'.

Couple- and uncouple-arcs

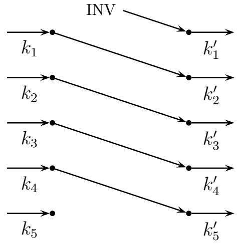
If there exists a composition change where a unit is coupled to task k of some outgoing trip we add a *couple-arc* from the inventory to task k, and similarly if there exists a composition change where a unit is uncoupled from task k of some incoming trip we add an *uncouple-arc* from task k to the inventory. The exact purpose of the arcs to and from the inventory will be explained shortly.

For connection $c \in \mathcal{C}$, each composition change $q \in \varrho(c)$ states exactly the value of the flow on the connection-arcs between the involved tasks and on the relevant couple- and uncouple-arcs. Figure 3.9 shows four composition changes at a connection between two trips. When a specific composition change is used for a connection then only a certain subset of the connection-, couple- and uncouple-arcs at the connection carry flow.

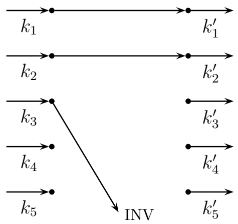
The principle is the same for connections that denote splitting or combining: A connection-arc connects a task in an incoming trip to a task in an outgoing trip. Figure 3.10 shows an example of a composition change in a splitting connection.



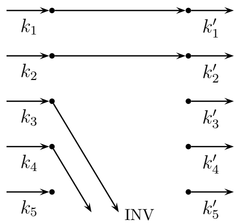
(a) q_1: Three through going units.



(b) q_2: Four through going units and one unit coupled in front.



(c) q_3: Two through going units and the unit in third position is uncoupled.



(d) q_4: Two through going units and the units in third and fourth positions are uncoupled.

Figure 3.9: Valid composition changes between trips t and t' where $\texttt{tasks}(t) = \{k_1, \ldots, k_5\}$ and $\texttt{tasks}(t') = \{k_1', \ldots, k_5'\}$. The shown connection-, couple- and uncouple-arcs carry a flow of value 1 if the corresponding composition change is used and all others have value 0.

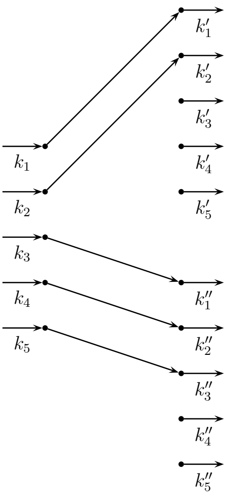


Figure 3.10: Composition change at a connection where a train is split. In this case the two front units continue on one trip and the three rear units continue on the other trip.

Inventory-arcs

To model the flow of units through the inventory we consider for each station $s \in \mathcal{S}$ the set of all relevant times at the station $\mathcal{R}(s)$. This is the set of all times that are either coupling time $\tau^+(c)$ or uncoupling time $\tau^-(c)$ for a connection c, or more formally:

$$\begin{aligned}\mathcal{R}(s) = &\big\{d \in \mathbb{R} \mid \exists c \in \mathcal{C} : \big(s(c) = s \text{ and } \tau^+(c) = d\big)\big\} \cup \\ &\big\{d \in \mathbb{R} \mid \exists c \in \mathcal{C} : \big(s(c) = s \text{ and } \tau^-(c) = d\big)\big\} \cup \{d_0, d_\infty\}\end{aligned}$$

where d_0 is an extra time that precedes all other times and is used for the initial inventory and d_∞ is an extra time that is after all other times and is used for the final inventory. For a relevant time $d \in \mathcal{R}(s)$, let $\texttt{prev}(d)$ and $\texttt{next}(d)$ denote the previous relevant and next relevant times at the station respectively. For each station $s \in \mathcal{S}$, we add an *inventory-arc* from d to $\texttt{next}(d)$ for all $d \in \mathcal{R}(s) \backslash \{d_\infty\}$. Furthermore the coupling- and uncoupling-arcs at the connections are linked to their corresponding relevant times.

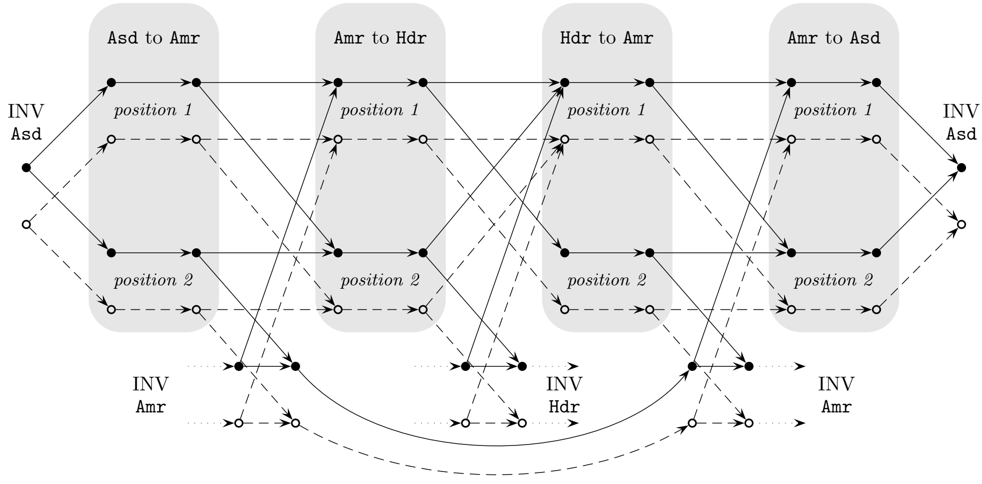


Figure 3.11: Network for the example in Figure 3.5. Solid arcs and nodes represent one rolling stock type while dashed arcs and non-solid nodes represent the other.

Arcs for multiple types

For each task it is possible to assign different types of rolling stock. To accommodate this we create a copy of each arc in the network for each rolling stock type $m \in \mathcal{M}$.

Example network

Figure 3.11 shows the resulting network for the example from Figure 3.5. The example involves two types of rolling stock units that may be combined into compositions of up to two units. Thus each trip has task-arcs for position one and two. The solid arcs and nodes belong to one type of unit whereas the dashed arcs and non-solid nodes belong to the other type. Each gray box corresponds to a trip with four task-arcs – one for each type and position. The arcs between trips represent coupling and uncoupling of units to and from the inventory, and the transfer of units between tasks in connected trips.

At station Alkmaar (`Amr`) we can couple a unit to the front of a train or uncouple a unit from the rear. This is modeled by couple-arcs from the inventory to the front position of the outgoing train and uncouple-arcs from the second position of the incoming train to the inventory. In the inventory there are inventory-arcs from each relevant time to the next relevant time. At station Den Helder (`Hdr`) the train is turned and units may be coupled to or uncoupled from the rear end of the train. The connection-, couple- and uncouple-arcs are added according to the shunting rules. A flow in this network with 0/1 value on task-, connection-, couple- and uncouple-arcs, and integer value on the inventory-arcs represents a solution to the circulation generation problem if there are no 'holes' in the compositions and the flow respects the possible shunting operations in the connections. The feasibility of a flow will be explained in more detail in the following sections.

3.5.1 Variables

An integral flow in the above described network represents the assignment of rolling stock types to tasks. To formalize the network flow mathematically we introduce the following variables.

- $T_{m,k} \in \{0,1\}$ denotes whether a rolling stock unit of type $m \in \mathcal{M}$ is assigned to task k for $\texttt{trip}(k) \in \mathcal{T}$.
- $R_{m,k,k'} \in \{0,1\}$ denotes whether a rolling stock unit of type $m \in \mathcal{M}$ assigned to task k is transferred to task k' in a connection.
- $U_{m,k} \in \{0,1\}$ denotes whether a unit of type $m \in \mathcal{M}$ is uncoupled after doing task k for $\texttt{trip}(k) \in \mathcal{T}$.
- $C_{m,k} \in \{0,1\}$ denotes whether a unit of type $m \in \mathcal{M}$ is coupled to do task k for $\texttt{trip}(k) \in \mathcal{T}$.
- $I_{s,m,d} \in \mathbb{Z}_+$ denotes how many units of type $m \in \mathcal{M}$ are in the inventory at station $s \in \mathcal{S}$ from relevant time $d \in \mathcal{R}(s) \backslash \{d_\infty\}$ to the next relevant time $\texttt{next}(d) \in \mathcal{R}(s)$.

Note that the use of coupling and uncoupling variables differs from the Composition Model. Here the variables denote coupling or uncoupling from or to a specific position rather than the number of units coupled or uncoupled at the connection.

3.5.2 Mathematical model

The core of the model is the manner in which the flow of rolling stock units is controlled at the connections. A connection $c \in \mathcal{C}$ involves a number, say n, connection-, coupling- and uncoupling-arcs. For a composition change $q \in \varrho(c)$ the values of the corresponding R-, U- and C-variables form a vector in $\{0,1\}^n$. Thus $\varrho(c)$ is associated with the set of allowed vectors $W \subseteq \{0,1\}^n$.

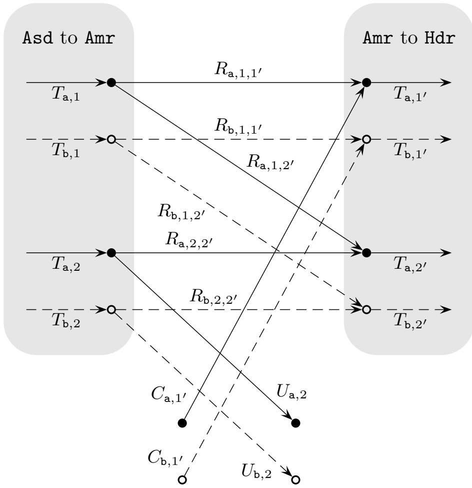


Figure 3.12: Part of the network from Figure 3.11. The arcs involved in the first connection at station Alkmaar (`Amr`) are labeled by their corresponding variables in the Task Model.

As an example we consider in Figure 3.12 the part of the network from Figure 3.11 that involves the connection at station Alkmaar (`Amr`) between the trip from `Asd` and the trip to `Hdr`. The task in the first position of the incoming trip is named 1 and the task in the second position is named 2. Similarly, the tasks for the outgoing trip are named $1'$ and $2'$. The two rolling stock types are named `a` and `b`. In the figure the arcs are labeled by their corresponding variables. In total, the connection involves $n = 10$ arcs.

The set of composition changes for the example, $\varrho(c)$, contains $|\varrho(c)| = 11$ elements. Table 3.1 shows for each composition change the value of the variables when the composition change is used. Some composition changes state that the incoming units are assigned to the same positions on the outgoing trip, whereas other composition changes imply the coupling or uncoupling of one unit. If each composition change q_i is considered as a vector in $\{0,1\}^{10}$ then the set $W = \{q_1, \ldots, q_{11}\} \subseteq \{0,1\}^{10}$ is the set of allowed combinations of flow values on the involved arcs for this connection.

The convex hull of W expresses exactly the subset of $\{0,1\}^n$ that corresponds to the allowed composition changes. We are therefore interested in describing the convex hull of W by a number of linear inequalities. In general, the number of linear inequalities may be superpolynomial in the number of points $|W|$ and the dimension n. But for the instances we encountered we were able to enumerate the

q	$R_{a,1,1'}$	$R_{a,1,2'}$	$R_{b,1,1'}$	$R_{b,1,2'}$	$R_{a,2,2'}$	$R_{b,2,2'}$	$U_{a,2}$	$U_{b,2}$	$C_{a,1'}$	$C_{b,1'}$
q_1	1	0	0	0	0	0	0	0	0	0
q_2	0	0	1	0	0	0	0	0	0	0
q_3	1	0	0	0	1	0	0	0	0	0
q_4	1	0	0	0	0	1	0	0	0	0
q_5	0	0	1	0	1	0	0	0	0	0
q_6	0	1	0	0	0	0	0	0	1	0
q_7	0	1	0	0	0	0	0	0	0	1
q_8	0	0	0	1	0	0	0	0	1	0
q_9	1	0	0	0	0	0	1	0	0	0
q_{10}	1	0	0	0	0	0	0	1	0	0
q_{11}	0	0	1	0	0	0	1	0	0	0

Table 3.1: The allowed composition changes at a connection and the corresponding values of the involved connection-, couple- and uncouple-variables.

facets of the convex hull of the allowed settings of the involved variables in negligible computation time. We used a brute force enumeration algorithm implemented in the software package PORTA (see Christof and Löbel (2009)).

Let $H(c)$ be the set of inequalities that describe the convex hull of the allowed settings of the variables involved in connection c. Then for inequality $s \in H(c)$, $\alpha^s_{m,k,k'}$ is the coefficient for variable $R_{m,k,k'}$, $\beta^s_{m,k}$ is the coefficient for variable $U_{m,k}$, $\gamma^s_{m,k}$ is the coefficient for variable $C_{m,k}$, and ζ_s is the right hand side. The inequality thus has the format:

$$\sum_{m\in\mathcal{M}}\sum_{t\in\mathrm{IN}_c}\sum_{k\in\mathtt{tasks}(t)}\sum_{t'\in\mathrm{OUT}_c}\sum_{k'\in\mathtt{tasks}(t')}\alpha^s_{m,k,k'}R_{m,k,k'} + \sum_{m\in\mathcal{M}}\sum_{t\in\mathrm{IN}_c}\sum_{k\in\mathtt{tasks}(t)}\beta^s_{m,k}U_{m,k} + \sum_{m\in\mathcal{M}}\sum_{t\in\mathrm{OUT}_c}\sum_{k\in\mathtt{tasks}(t)}\gamma^s_{m,k}C_{m,k} \le \zeta_s$$

For the example in Table 3.1, the following 5 inequalities describe the allowed composition changes.

$$
\begin{aligned}
&R_{\mathtt{a},1,2'} + R_{\mathtt{b},1,2'} - C_{\mathtt{a},1'} - C_{\mathtt{b},1'} && = 0\\
&R_{\mathtt{a},1,1'} + R_{\mathtt{b},1,1'} + C_{\mathtt{a},1'} + C_{\mathtt{b},1'} && = 1\\
&- R_{\mathtt{b},1,2'} + C_{\mathtt{a},1'} && \geq 0\\
&R_{\mathtt{b},1,1'} + R_{\mathtt{b},2,2'} + U_{\mathtt{b},2} + C_{\mathtt{a},1'} + C_{\mathtt{b},1'} && \leq 1\\
&R_{\mathtt{a},2,2'} + R_{\mathtt{b},2,2'} + U_{\mathtt{a},2} + U_{\mathtt{b},2} + C_{\mathtt{a},1'} + C_{\mathtt{b},1'} && \leq 1
\end{aligned}
$$

The description of the convex hull of the allowed settings of the variables involved at each connection ensures that the flow of units in the network adheres to the allowed composition changes. We have now described the variables involved in the model and the inequalities that describe allowed values of variables at the connections. The Task Model is then expressed mathematically as follows:

$$\text{minimize } f(T, R, I) \tag{3.44}$$

subject to

$$T_{m,k} = U_{m,k} + \sum_{t' \in \mathrm{OUT}_{\sigma(t)}} \sum_{k' \in \mathtt{tasks}(t')} R_{m,k,k'} \qquad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathtt{tasks}(t) \tag{3.45}$$

$$T_{m,k} = C_{m,k} + \sum_{t' \in \mathrm{IN}_{\pi(t)}} \sum_{k' \in \mathtt{tasks}(t')} R_{m,k',k} \qquad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \mathtt{tasks}(t) \tag{3.46}$$

$$
\begin{aligned}
&\sum_{m \in \mathcal{M}} \sum_{t \in \mathrm{IN}_c} \sum_{k \in \mathtt{tasks}(t)} \sum_{t' \in \mathrm{OUT}_c} \sum_{k' \in \mathtt{tasks}(t')} \alpha^s_{m,k,k'} R_{m,k,k'}\\
&\quad + \sum_{m \in \mathcal{M}} \sum_{t \in \mathrm{IN}_c} \sum_{k \in \mathtt{tasks}(t)} \beta^s_{m,k} U_{m,k}\\
&\quad + \sum_{m \in \mathcal{M}} \sum_{t \in \mathrm{OUT}_c} \sum_{k \in \mathtt{tasks}(t)} \gamma^s_{m,k} C_{m,k} \leq \zeta_s \qquad \forall c \in \mathcal{C}, s \in H(c)
\end{aligned} \tag{3.47}
$$

$$
\begin{aligned}
&\sum_{\substack{c \in \mathcal{C}: s(c)=s,\\ \tau^-(c)=d}} \sum_{t \in \mathrm{IN}_c} \sum_{k \in \mathtt{tasks}(t)} U_{m,k} + I_{s,m,d} =\\
&\sum_{\substack{c \in \mathcal{C}: s(c)=s,\\ \tau^+(c)=d}} \sum_{t \in \mathrm{OUT}_c} \sum_{k \in \mathtt{tasks}(t)} C_{m,k} + I_{s,m,\mathtt{next}(d)} \qquad \forall s \in \mathcal{S}, d \in \mathcal{R}(s) \backslash \{d_\infty\}
\end{aligned} \tag{3.48}
$$

$$I_{s,m,d_0} = i^0_{s,m} \qquad \forall s \in \mathcal{S}, m \in \mathcal{M} \tag{3.49}$$

$$I_{s,m,\texttt{prev}(d_\infty)} = i^\infty_{s,m} \qquad \forall s \in \mathcal{S}, m \in \mathcal{M} \tag{3.50}$$

$$T_{m,k}, U_{m,k}, C_{m,k} \in \{0,1\} \qquad \forall m \in \mathcal{M}, t \in \mathcal{T}, k \in \texttt{tasks}(t) \tag{3.51}$$

$$R_{m,k,k'} \in \{0,1\} \qquad \forall m \in \mathcal{M}, c \in \mathcal{C}, t \in \mathrm{IN}_c, k \in \texttt{tasks}(t), t' \in \mathrm{OUT}_c, k' \in \texttt{tasks}(t') \tag{3.52}$$

$$I_{s,m,d} \in \mathbb{R}_+ \qquad \forall s \in \mathcal{S}, m \in \mathcal{M}, d \in \mathcal{R}(s) \backslash \{d_\infty\} \tag{3.53}$$

Constraints (3.45) and (3.46) are flow conservation constraints at the connections. Constraints (3.45) state that a unit of type m is assigned to task k only if it is either uncoupled or assigned to a task of an outgoing trip in the succeeding connection. Similarly constraints (3.46) state that a unit of type m is assigned to task k only if it is either coupled or assigned to a task of an incoming trip in the preceding connection.

Constraints (3.47) describe the allowed sets of variables for each connection $c \in \mathcal{C}$ through the set of inequalities $H(c)$.

Constraints (3.48) are flow conservation constraints for the inventories. The number of units of type m at a station at the relevant time d plus the units uncoupled at that time equals the number of units coupled at that time plus the number of available units at the next relevant time.

Constraints (3.49) ensure that the number of units of type m at station s at the beginning of the planning period is equal to the planned number of units available there. Note that variable I_{s,m,d_0} denotes the flow of units of type m at the first relevant time at station s. Similarly, constraints (3.50) state the same for the end of the planning period.

Constraints (3.51) – (3.53) denote the domains of the variables. We declared the variables $I_{s,m,d}$ continuous, but their integrality follows from the following proposition.

Proposition 1. *The variables $I_{s,m,d}$ are integer in any solution to the model.*

Proof. The first inventory arcs are integer due to constraints (3.49), and constraints (3.48) state that only integer values are added to or removed from the inventory flow. □

3.5.3 Reformulating the model

Since we intend to solve the Task Model with a general MIP solver it turns out to be advantageous for the solution procedure to introduce a number of higher level binary variables. We describe those variables and their accompanying constraints in this section.

The variables $T_{m,k}$ describe the rolling stock units assigned to each position in the train for trip t. For a trip $t \in \mathcal{T}$ we add a variable $X_{t,p} \in \{0,1\}$ for each possible composition $p \in \eta(t)$ along with the following constraints:

$$\sum_{p \in \eta(t)} X_{t,p} = 1 \qquad \forall t \in \mathcal{T} \tag{3.54}$$

$$\sum_{p \in \eta(t)} a_{p,m,k} X_{t,p} = T_{m,k} \qquad \forall t \in \mathcal{T}, m \in \mathcal{M}, k \in \texttt{tasks}(t) \tag{3.55}$$

$$X_{t,p} \in \{0,1\} \qquad \forall t \in \mathcal{T}, p \in \eta(t) \tag{3.56}$$

where the binary coefficient $a_{p,m,k}$ denotes whether a unit of type m is assigned to position $\texttt{pos}(k)$ in composition p. The addition of the $X_{t,p}$ variables adds the possibility of branching on higher level decision variables rather than on the low level task variables.

The following proposition states that the addition of the $X_{t,p}$ variables along with constraints (3.54) – (3.56) enables us to relax the integrality constraints of the $T_{m,k}$ variables.

Proposition 2. *The integrality of the $X_{t,p}$ variables implies the integrality of the $T_{m,k}$ variables.*

Proof. The proposition follows from the equality in constraint (3.55) and the integrality of the left-hand side. □

Similarly, for each connection $c \in \mathcal{C}$, we add variables $Z_{c,q} \in \{0,1\}$ to denote that composition change $q \in \varrho(c)$ is used at connection c. Let the binary coefficient $b_{q,m,k,k'}$ denote whether a unit of type m is transferred from task k of incoming trip $\texttt{trip}(k)$ to task k' in outgoing trip $\texttt{trip}(k')$ in the composition change q. In the same way, let $u_{q,m,k}$ ($g_{q,m,k}$) denote whether a unit of type m is uncoupled from (coupled to) task k in incoming (outgoing) trip $\texttt{trip}(k)$ in the composition change q. Consider the set of constraints (3.57) – (3.61).

$$\sum_{q \in \varrho(c)} Z_{c,q} = 1 \qquad \forall c \in \mathcal{C} \tag{3.57}$$

$$\sum_{q \in \varrho(c)} b_{q,m,k,k'} Z_{c,q} = R_{m,k,k'} \qquad \forall c \in \mathcal{C}, m \in \mathcal{M}, t \in \mathrm{IN}_c : t = \texttt{trip}(k), t \in \mathrm{OUT}_c : t = \texttt{trip}(k') \tag{3.58}$$

$$\sum_{q \in \varrho(c)} u_{q,m,k} Z_{c,q} = U_{m,k} \qquad \forall c \in \mathcal{C}, m \in \mathcal{M}, t \in \mathrm{IN}_c : t = \texttt{trip}(k) \tag{3.59}$$

$$\sum_{q \in \varrho(c)} g_{q,m,k} Z_{c,q} = C_{m,k} \qquad \forall c \in \mathcal{C}, m \in \mathcal{M}, t \in \mathrm{OUT}_c : t = \texttt{trip}(k) \tag{3.60}$$

$$Z_{c,q} \in \{0,1\} \qquad \forall c \in \mathcal{C}, q \in \varrho(c) \tag{3.61}$$

The addition of the constraints (3.57) – (3.61) leads to the following proposition.

Proposition 3. *The variables $R_{m,k,k'}$, $U_{m,k}$ and $C_{m,k}$ are integral in any solution to the model* (3.45), (3.46), (3.48) – (3.53), (3.57) – (3.61).

Proof. Similarly to the proof of Proposition 2 the integrality of the variables $R_{m,k,k'}$, $U_{m,k}$ and $C_{m,k}$ follows from the equalities in constraints (3.58) – (3.60) and the integrality of the left-hand sides. □

The following proposition states that we can exchange the constraints describing the convex hull of the allowed composition changes of a connection by the constraints (3.57) – (3.61).

Proposition 4. *A feasible solution to the linear relaxation of the model* (3.45) – (3.53) *corresponds to a feasible solution to the linear relaxation of the model* (3.45), (3.46), (3.48) – (3.53), (3.57) – (3.61), *and vice versa.*

Proof. Consider a solution to the linear relaxation of (3.45) – (3.53). Then we construct a solution to the reformulated model by first assigning the same values to the variables $T_{m,k}$, $R_{m,k,k'}$, $U_{m,k}$, $C_{m,k}$ and $I_{s,m,d}$. Only the variables $Z_{c,q}$ remain. Let $c \in \mathcal{C}$ be a connection and consider the connection-, uncoupling- and coupling-variables $R_{m,k,k'}$, $U_{m,k}$, $C_{m,k}$ associated with c. According to (3.47) their values form a vector in the convex hull of the allowed settings associated with the composition changes $\varrho(c)$. Therefore the values can be described as a convex combination of the parameters $b_{q,m,k,k'}$, $u_{q,m,k}$ and $g_{q,m,k}$ weighted by the variables $Z_{c,q}$.

Conversely, consider a solution to the linear relaxation of the model (3.45), (3.46), (3.48) – (3.53), (3.57) – (3.61). The variables $T_{m,k}$, $R_{m,k,k'}$, $U_{m,k}$, $C_{m,k}$ and $I_{s,m,d}$ form a feasible solution to the linear relaxation of the model (3.45) – (3.53) since the constraints (3.47) are valid inequalities for the extended description (3.57) – (3.61). □

The proof can be repeated easily for the special case where only integral solutions are considered.

Proposition 5. *An integral solution of the model* (3.45) – (3.53) *corresponds to an integral solution of the model* (3.45), (3.46), (3.48) – (3.53), (3.57) – (3.61), *and vice versa.*

Proposition 4 and Proposition 5 state that the introduction of new variables and a reformulation of the model does not lead to stronger LP bounds. However, the addition of the $Z_{c,q}$ variables and constraints (3.57) – (3.61) enables us to branch on composition changes rather than on the flow of units inside the connection.

We perform computational tests of the Task Model with the reformulated version in Section 3.8.

3.5.4 Objective function

The objective function is a function of the compositions assigned to trips in the modified timetable and the composition changes at the connections between trips. With the introduction of the X- and Z-variables described in Section 3.5.3 it is possible to take the RSRP objectives into account in a similar manner as in the Composition Model.

Objectives for the service perspective

The service perspectives are taken into account in the same way as described for the Composition Model in Section 3.4.5. It requires only the addition of a few extra variables to the model.

The number of canceled trips can be measured in the same way as in the Composition Model, namely by introducing a variable $X_{t,0}$ for trip $t \in \mathcal{T}$. This variable is then equal to 0 if and only if all corresponding variables $T_{m,k}$ for the trip are 0.

The number of seat shortage kilometers can be measured by the variables $X_{t,p}$ as the capacity of the units in composition p assigned to trip t is known, as well as the forecasted number of passengers.

Objectives for the efficiency perspective

The number of carriage kilometers in a rolling stock schedule is a linear function of the variables $T_{m,k}$. Each of these variables represents the assignment of a unit of type m to task k in the train of trip $t = \texttt{trip}(k)$, and the variable thus leads to a known number of carriage kilometers when assigned the value 1. Alternatively, the number of carriage kilometers can be measured by the variables $X_{t,p}$ as described in Section 3.4.4.

Objectives for the process perspective

The changes to the planned shunting operations can be measured by the variables $Z_{c,q}$ as described in Section 3.4.6. Furthermore, it is possible to penalize or encourage uncoupling from or coupling to certain positions in the trains by explicitly including the variables $C_{m,k}$ and $U_{m,k}$ in the objective function.

Off-balances are modeled using the same idea as for the Composition Model. We add deviation variables $D^{\infty,+}_{s,m}$ and $D^{\infty,-}_{s,m}$ to denote the number of units of type m in excess or shortage at the end of the planning period at station s. We then replace constraints (3.50) by (3.62) in the model.

$$I_{s,m,\texttt{prev}(d_\infty)} = i^{\infty}_{s,m} + D^{\infty,+}_{s,m} - D^{\infty,-}_{s,m} \qquad \forall s \in \mathcal{S}, m \in \mathcal{M} \qquad (3.62)$$

Off-balances can then be penalized in the objective function as in the Composition Model. Similarly, we can discourage cumulative deviations as described for the Composition Model.

The implicit changes to the schedules for the conductors are taken into account through the variables $X_{t,p}$ just as in the Composition Model. These variables imply the number of conductors needed on a train in the same way as described earlier.

3.5.5 Equivalence with the Composition Model

The Composition Model as defined in Section 3.4 and the Task Model as defined in Section 3.5 model the same problem. A feasible solution to the Composition Model corresponds to a feasible solution in the Task Model and vice versa. With the addition of composition variables $X_{t,p}$ and composition change variables $Z_{c,q}$ to the Task Model, described in Section 3.5.3, it is straightforward to translate a feasible solution from one model to the other.

The objective terms we suggested to account for various managerial aspects can be applied for both models as described in Sections 3.4.4 to 3.4.6, and 3.5.4. When using the same objective terms in both models, a feasible solution to one model corresponds to a feasible solution to the other model with the same objective value.

The major difference between the models is the underlying network. The two networks primarily differ in the way the relationship between compositions and composition changes is expressed. More specifically, in the Composition Model the relationship between compositions and composition changes is modeled through flow conservation constraints, while in the Task Model the relationship is modeled through tracking the units assigned to the positions in the trains.

3.6 Duty Generation

The output of the circulation generation step is an assignment of compositions to trips in the modified timetable and the choice of composition changes at the connections. The assignment is based on anonymous units that are only distinguished by their types. What is still missing is to derive rolling stock duties from the assignment. In particular, the circulation does not specify how to join uncoupled units with later coupled units. Determining how the sequences of tasks are joined is the main problem in the duty generation phase. We here elaborate on the duty generation and show that the problem in this phase amounts to covering the nodes in a certain directed graph by a given number of paths.

The first step is to derive the *fundamental duties* for each type of rolling stock from the composition assignment. A fundamental duty is the list of tasks performed by a rolling stock unit of a certain type from coupling to uncoupling. A fundamental duty thus starts when the unit is coupled to a composition at a station and ends when the unit is uncoupled from a composition at another station. We emphasize that determining the fundamental duties from a circulation does not involve any decisions – they are uniquely determined by the circulation.

The example assignment in Figure 3.4 results in the seven fundamental duties shown in the diagram in Figure 3.13. The fundamental duties are numbered and the corresponding rolling stock type is shown next to the number. The tasks contain information on which trip they belong to, through a train service number and departure/arrival stations, and the position in the composition assigned to the trip. Recall that position 1 denotes the front of the train.

Based on the fundamental duties we define the *duty graph*. It is a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ where the vertex set, $\mathcal{V}$, is the set of fundamental duties. In short, the

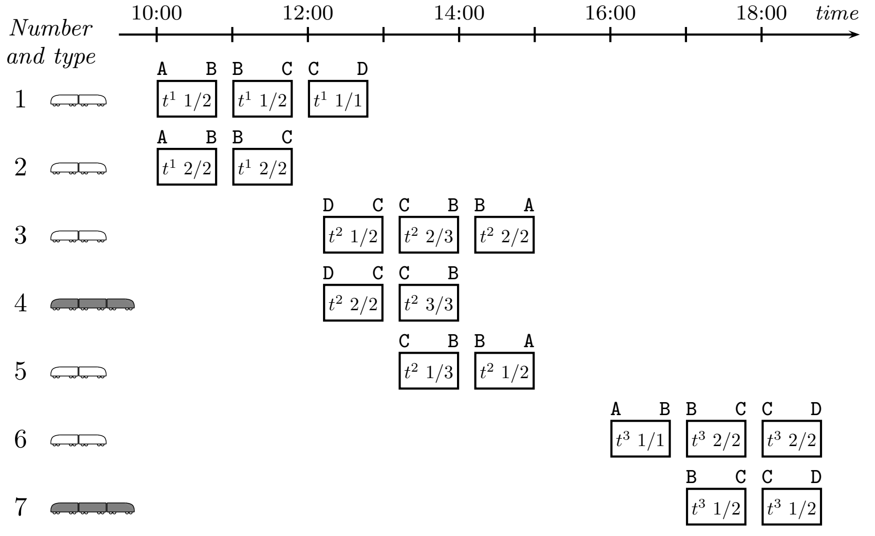


Figure 3.13: Fundamental duties for the circulation in Figure 3.4.

graph connects two vertices if it is possible for a rolling stock unit to perform the second fundamental duty right after the first.

More formally, we define the duty graph as follows. For a vertex $v \in \mathcal{V}$ we denote the corresponding rolling stock type by $m(v)$, and the departure and arrival time of the first and the last task by $d(v)$ and $a(v)$, respectively. Let $s(v)$ denote the station where the duty starts and let $e(v)$ be the station where the duty ends. Let $\delta(u, v)$ be the time needed to uncouple the unit from the train at station $e(u)$, send it to the inventory, retrieve it from the inventory and couple it to the train of v. In other words, let $\delta(u, v) = \tau^-(c) - a(u) + d(v) - \tau^+(c')$ where c is the successor connection of the last trip in u and c' is the predecessor connection of the first trip in v.

We then define the arc set as

$$\mathcal{A} = \{(u, v) \in \mathcal{V} \times \mathcal{V} \mid m(u) = m(v),\ e(u) = s(v) \text{ and } a(u) + \delta(u, v) \leq d(v)\}$$

which means there is an arc from u to v if the duties are served by the same rolling stock type, duty u ends at the same station as v starts, and there is enough time to perform the necessary shunting between the duties.

The duty graph for the fundamental duties in our example is shown in Figure 3.14. Clearly, a duty of a rolling stock unit naturally corresponds to a path in this graph. Moreover, the following observation holds.

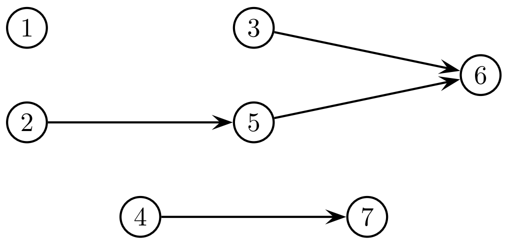


Figure 3.14: Duty graph for the fundamental duties in Figure 3.13.

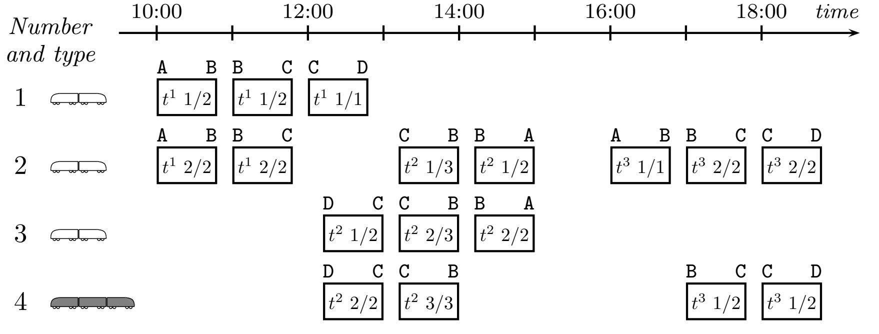


Figure 3.15: Final duties created from the fundamental duties in Figure 3.13.

Observation 1. *In the duty graph, any vertex associated with a given rolling stock type has only arcs to vertices associated with the same rolling stock type.*

Therefore the duty generation problem amounts to finding a path cover in the duty graph. A *path cover* in a graph is a set of vertex-disjoint paths such that each vertex is contained in exactly one of these paths. Here we also require that the number of paths for each rolling stock type is equal to the number of available units, and the path cover respects the rolling stock balance at the beginning and the end of the planning period. By Observation 1, the duty generation can be performed separately for each involved rolling stock type. We therefore can restrict ourselves to the case of a single rolling stock type without loss of generality.

For the running example a possible solution is to cover the duty graph with four vertex-disjoint paths, $\{1\}$, $\{2, 5, 6\}$, $\{3\}$, $\{4, 7\}$. This results in the duties shown in Figure 3.15.

With regards to the feasibility of the duty generation process, Proposition 6 states that whenever the Composition Model or the Task Model has a solution then the arising duty generation problem has a solution, too.

Proposition 6. *Let S be a solution to the Composition Model* (3.1) – (3.14) *or to the Task Model* (3.44) – (3.60)*, and let $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ be the duty graph based on S. Then there exists a path cover in $\mathcal{G}$ where the number of paths corresponds to the rolling stock availability.*

Proof. Let $\mathcal{G}$ be the duty graph constructed from a solution S to the circulation generation problem. Observation 1 states that we only need to consider one rolling stock type so let $m \in \mathcal{M}$ be an arbitrary rolling stock type.

The solution S immediately implies an integer network flow of rolling stock units of type m in the time-expanded network. Such a network flow can be decomposed to a set P of paths. We will show that P gives rise to a path cover in $\mathcal{G}$. Note that sources and sinks of the network flow are determined by the parameters $i^0_{s,m}$ and $i^\infty_{s,m}$, and therefore P respects the rolling stock availability.

Indeed, the solution S uniquely determines the fundamental duties, and P covers them exactly once. Moreover, by the definition of arcs in the duty graph, a unit may be assigned to fundamental duty v after u only if $a(u) + \delta(u, v) \leq d(v)$. This is equivalent to stating $\tau^-(c) \leq \tau^+(c')$ for the involved connections which is exactly the condition for an uncoupled unit to be ready for coupling according to constraints (3.7) (constraints (3.48) for the Task Model). Therefore each path of P corresponds to a path in $\mathcal{G}$. □

For the objective of the duty generation process we mentioned earlier that two elements are important: (i) The overlap between current and modified duties, i.e. the amount of work each current duty has in common with its corresponding modified duty. (ii) The number of times a modified duty contains an uncoupling from a train and later coupling to a train that also existed in an original duty.

When the rolling stock planners consider the original and modified duties they prefer that the modified duties resemble the original ones as much as possible. Therefore, the planners emphasize element (i) in the duty generation objective. However, from the perspective of the local planners, element (ii) is important since an intact pattern of uncoupling and later coupling implies that the storage of a unit during its idle-time possibly can stay intact. Fortunately, the two elements in the objective coincide to some extend since keeping the uncoupling/coupling pattern of a duty implies an overlap between the corresponding current and modified duty.

Next, we present two models for duty generation and explain how the objective is taken into account. The Duty Flow Model presented in Section 3.6.1 is based on a network flow in the duty graph where we encourage certain pairs of fundamental duties. The Duty Path Model presented in Section 3.6.2 is based on the path formulation of the same network flow. The path formulation allows us to apply a more sophisticated objective function.

3.6.1 Duty Flow Model

The idea behind the Duty Flow Model is to propose a simple and intuitive model based on a network flow in the duty graph. We require that the flow on each arc has at most value 1 and the flow through each node is exactly 1. A unit valued flow in the graph then corresponds to a set of disjoint paths that cover all vertices. Each path in this network flow covers one or more vertices and corresponds to a duty. Consider the following variables:

- $K_a \in \{0,1\}$ is defined for all $a = (u,v) \in \mathcal{A}$ and denotes whether vertices u and v are performed in succession by the same duty.

- $B_v \in \{0,1\}$ denotes whether vertex $v \in \mathcal{V}$ is the first vertex in a duty.

- $E_v \in \{0,1\}$ denotes whether vertex $v \in \mathcal{V}$ is the last vertex in a duty.

Define for each vertex the set of incoming arcs $\delta^{\texttt{in}}(v) \subseteq \mathcal{A} = \{a \in \mathcal{A} \mid \exists u \in \mathcal{V} : a = (u,v)\}$, and the set of outgoing arcs $\delta^{\texttt{out}}(v) \subseteq \mathcal{A} = \{a \in \mathcal{A} \mid \exists w \in \mathcal{V} : a = (v,w)\}$. The following mathematical model forms the basis of the duty generation.

$$\max \sum_{(u,v)\in\mathcal{A}} c_{u,v} K_{u,v} \tag{3.63}$$

subject to

$$\sum_{a\in\delta^{\texttt{out}}(v)} K_a + E_v = 1 \qquad \forall v \in \mathcal{V} \tag{3.64}$$

$$\sum_{a\in\delta^{\texttt{in}}(v)} K_a + B_v = 1 \qquad \forall v \in \mathcal{V} \tag{3.65}$$

$$\sum_{v \in \mathcal{V}: s(v)=s} B_v = I^{0*}_{s,m} \qquad \forall s \in \mathcal{S} \tag{3.66}$$

$$\sum_{v \in \mathcal{V}: e(v)=s} E_v = I^{\infty*}_{s,m} \qquad \forall s \in \mathcal{S} \tag{3.67}$$

$$B_v, E_v \in \{0,1\} \qquad \forall v \in \mathcal{V} \tag{3.68}$$

$$K_a \in \{0,1\} \qquad \forall a \in \mathcal{A} \tag{3.69}$$

Constraints (3.64) and (3.65) are flow conservation constraints. More specifically, constraints (3.64) state that the unit covering a vertex $v \in \mathcal{V}$ continues to cover another vertex or is the last vertex in the duty. In the same way constraints (3.65) ensure that a vertex $v \in \mathcal{V}$ is either preceded by another vertex or is the first vertex in a duty. Constraints (3.68) and (3.69) denote the domains of the variables.

We want the entire plan to be covered by a certain number of rolling stock units. The rolling stock units are positioned at certain stations at the beginning and end of the planning period. The positions of the rolling stock at the beginning of the planning period are given by the variables $I^0_{s,m}$ in the Composition Model (I_{s,m,d_0} in the Task Model). At the end of the planning period the positions are given by $I^\infty_{s,m}$ in the Composition Model ($I_{s,m,\texttt{prev}(d_\infty)}$ in the Task Model). Let $I^{0*}_{s,m}$ and $I^{\infty*}_{s,m}$ denote the value of the variables in the solution to the circulation generation problem. Constraints (3.66) and (3.67) ensure the correct number of starting and ending units at each station.

Some units may have no tasks to perform, they should then be assigned to an empty duty. However, an empty duty implies that the unit is stored at a station for the entire planning period, and the unit should thus be counted in the number of units starting and ending there. This is easily achieved by adding an appropriate number of dummy vertices v' to the set $\mathcal{V}$ where each vertex v' is connected to no other vertices and has $s(v') = e(v')$.

The objective function of the model is to maximize a bonus $c_{u,v}$ for using u and v consecutively in the same duty. These bonuses can help to achieve the overall rescheduling goal of creating duties that resemble the original duties. We restrict ourselves to the cost structure where $c_{u,v} = 1$ if a duty for a unit of the same type originally contained an uncoupling after the last task in u and a coupling before the first task of v. We set $c_{u,v} = \frac{1}{2}$ if a duty for a unit of a different type originally had the same property. We set $c_{u,v} = 0$ for the remaining arcs.

For the practical instances we encountered the duty graph is relatively small; it consists of up to around $n = |\mathcal{V}| = 200$ fundamental duties and around $m = |\mathcal{A}| = 300$ potential ways to connect them. We note that the Duty Flow Model (3.63) – (3.67) can be transformed to an instance of the minimum cost flow problem with unit arc capacities which is solved in $\mathcal{O}(m \log n(m{+}n \log n))$ time by applying the algorithm of Orlin (1988). We solved the models in neglectable time using a commercial-strength MIP solver.

One shortcoming of the flow based model is that the parameters in the objective function are based on whether pairs of vertices are consecutive in the duties. It is thus not possible to take into account the structure of an entire potential duty. We therefore introduce an alternative model in Section 3.6.2 where we take entire potential duties into account rather than just how to join pairs of fundamental duties.

3.6.2 Duty Path Model

As we observed in the Duty Flow Model we are only able to stimulate the joining of pairs of fundamental duties. If we were to consider entire potential duties instead we could stimulate the structure of entire duties instead. It therefore seems intuitive to enumerate all paths in the duty graph and select a covering based on how much the paths resemble the complete original duties. This is the idea behind the Duty Path Model presented in this section.

In fact, the Duty Path Model is a path based formulation of the Duty Flow Model, but with the addition of the matching with the original duties it is a multi-commodity flow model rather than a single-commodity flow model.

Let $\mathcal{P}$ denote the set of all paths in the duty graph. In the example from Figure 3.14 we have $\mathcal{P} = \{\{1\}, \{2\}, \{2,5\}, \{2,5,6\}, \{3\}, \{3,6\}, \{4\}, \{4,7\}, \{5\}, \{5,6\}, \{6\}, \{7\}\}$. Let $\mathcal{U}$ denote the set of original rolling stock duties. We then let variable $Q_{u,p} \in \{0,1\}$ denote whether original duty $u \in \mathcal{U}$ is assigned to path $p \in \mathcal{P}$ in the modified situation. Let parameter $h_{v,p} = 1$ if vertex v is contained in path p and 0 otherwise. Parameter $s_{u,p}$ denotes the similarity between the original duty $u \in \mathcal{U}$ and the duty implied by path $p \in \mathcal{P}$. We elaborate on the measure of similarity later in this section. We model the duty generation by the following path cover model.

$$\max \sum_{u \in \mathcal{U}} \sum_{p \in \mathcal{P}} s_{u,p} Q_{u,p} \tag{3.70}$$

subject to

$$\sum_{p \in \mathcal{P}} Q_{u,p} = 1 \qquad \forall u \in \mathcal{U} \tag{3.71}$$

$$\sum_{u\in\mathcal{U}}\sum_{p\in\mathcal{P}} h_{v,p}Q_{u,p} = 1 \qquad \forall v \in \mathcal{V} \tag{3.72}$$

$$Q_{u,p} \in \{0,1\} \qquad \forall u \in \mathcal{U}, p \in \mathcal{P} \tag{3.73}$$

Constraints (3.71) state that each unit is assigned exactly one path and constraints (3.72) ensure that each vertex is covered exactly once.

We note that it is possible to add dummy paths to denote empty duties for units that have no tasks to perform.

The parameter $s_{u,p}$ expresses the similarity between the original duty u and the potential new duty p. We use as similarity measure a number in the interval $[0,1]$ to denote how closely related the two duties are. Here we restrict ourselves to the following definition: If the duties start at the same station we add a value α to the measure and if the duties end at the same station we add a value β where $0 \leq \alpha, \beta \leq 1$ and $\alpha + \beta < 1$. Let m_1 and m_2 be the number of minutes of train tasks in the two duties respectively, and let c be the number of minutes of train task the two duties have in common. Then define m as the relative number of minutes of train tasks the two duties have in common by $m = \frac{c}{\max\{m_1,m_2\}}$, and add $(1-\alpha-\beta)m$ to the similarity measure. (In the special case where both duties are empty, i.e. $m_1 = m_2 = 0$, we define $m = 1$ if both duties are at the same station and $m = 0$ otherwise). This way we arrive at a number in the interval $[0,1]$:

$$s_{u,p} = \alpha + \beta + m\left(1 - (\alpha + \beta)\right)$$

The following proposition states that a similarity of 1 implies that the sets of tasks performed are identical.

Proposition 7. *If duties d_1 and d_2 have a similarity of 1 then d_1 and d_2 are identical.*

Proof. Consider two duties d_1 and d_2 with similarity 1, that is

$$s_{d_1,d_2} = (1 - \alpha + \beta)m + \alpha + \beta = 1$$

The definition implies

$$(1 - m)(\alpha + \beta - 1) = 0 \Rightarrow$$
$$m = 1 \text{ or } \alpha + \beta = 1$$

The definition of m implies that either the minutes of work performed by the duties, m_1 and m_2, are both zero or $\max\{m_1, m_2\} = c$. Since the number of minutes of work the duties have in common, c, is subject to $c \leq \min\{m_1, m_2\}$ we get that either the duties are identical or $\alpha + \beta = 1$. □

If we set $\alpha + \beta = 1$, the similarity is only measured by whether the start and end stations coincide. Tests of the model revealed that values of $\alpha = \beta = 0.05$ led to reasonable solutions.

For the solvability of the model we note that traditionally many multi-commodity flow problems require advanced algorithmic techniques to solve, but this problem is solved in a few seconds with a commercial MIP solver due to two properties. First, the size of practical instances is relatively small because the problem is decomposed by rolling stock types. The largest instances found have around $|\mathcal{P}| = 1500$ potential new duties. Second, with the proposed similarity measure, an original duty is usually only highly similar to at most one new duty and the LP relaxation is therefore quite strong.

The two proposed models both attempt to achieve the goal of creating duties that as far as possible resemble the original duties. However, there is no truly objective measure of this goal. This also means that there is no objective way to compare the outcome of the two proposed models. We did manage to get the outcome of both models evaluated by practitioners and even though this is a rather subjective comparison, the results of the path based model were preferable. This is mainly due to the use of duty similarities in the objective function which minimizes the visual changes in the presentation of the modified duties.

3.7 Simultaneous circulation and duty generation

The models described so far in this chapter solve RSRP by first generating a circulation and then creating a set of duties from the circulation. In this section we describe an integrated approach that integrates the construction of the duties in the circulation generation.

The idea is to extend the notion of the inventory in the Task Model. The inventory, as modeled in the Task Model in Section 3.5, accounts for the number of units of each type available at any time at each station. Instead we want to explicitly track which units are uncoupled from specific positions in arriving trips and later coupled to specific positions in departing trips at the same station. This is what is decided in the duty generation process. The proposed extension is modeled by exchanging the inventory arcs in the Task Model with dedicated arcs between each pair of tasks in an incoming trip and a later outgoing trip at the same station. The resulting model is called the Duty-Task Model.

We introduce a number of new variables. Let the variable $J_{m,k,k'} \in \{0,1\}$ denote the decision that a unit of type $m \in \mathcal{M}$ performs task k of trip $t = \texttt{trip}(k)$, is uncoupled and then later coupled to task k' of trip $t' = \texttt{trip}(k')$. This variable is only introduced for pairs (k, k') where the first trip t arrives at the departure station of trip t' with enough time to perform the necessary shunting movement. In formal terms it is required that the uncoupling and later coupling take place at the same station, i.e. $s(\sigma(t)) = s(\pi(t'))$. Furthermore it is required that there is enough time to perform the movements of the unit inside the station, i.e. $\tau^-(\sigma(t)) \leq \tau^+(\pi(t'))$.

Let variable $J^0_{m,k} \in \{0,1\}$ denote whether a unit of type m is coupled to task k of departing trip $\texttt{trip}(k)$ without being assigned to any trip earlier. Similarly, variable $J^\infty_{m,k} \in \{0,1\}$ denotes whether a unit of type m is uncoupled from task k of an arriving trip $\texttt{trip}(k)$ and is not used in any later tasks. Variable $J^{\texttt{idle}}_{s,m}$ denotes the number of units of type m that are available at station s but are neither coupled to a trip nor uncoupled from a trip. In other words, they are idle at the same station during the entire planning period.

The number of units of a certain type at a given station at the beginning of the planning period is equal to the number of units that are coupled to a trip departing from that station plus the number of units that are idle. Similarly, the number of units of a certain type at a given station at the end of the planning period is equal to the number of units that are uncoupled from a trip arriving at the station plus the number of units that are idle.

We construct the Duty-Task Model by replacing the constraints (3.48) – (3.50) and (3.53) in the Task Model with the constraints (3.74) – (3.80). The Duty-Task Model is thus described by (3.45) – (3.47), (3.51) – (3.55), (3.57) – (3.60), (3.74) – (3.80).

$$U_{m,k} = \sum_{\substack{t' \in \mathcal{T}: s(\sigma(t)) = s(\pi(t')), \\ \tau^-(\sigma(t)) \leq \tau^+(\pi(t'))}} \sum_{k' \in \texttt{tasks}(t')} J_{m,k,k'} + J^\infty_{m,k} \qquad \forall t \in \mathcal{T}, k \in \texttt{tasks}(t), m \in \mathcal{M} \tag{3.74}$$

$$C_{m,k} = \sum_{\substack{t' \in \mathcal{T}: s(\sigma(t')) = s(\pi(t)), \\ \tau^-(\sigma(t')) \leq \tau^+(\pi(t))}} \sum_{k' \in \texttt{tasks}(t')} J_{m,k',k} + J^0_{m,k} \qquad \forall t \in \mathcal{T}, k \in \texttt{tasks}(t), m \in \mathcal{M} \tag{3.75}$$

$$\sum_{t \in \mathcal{T}: s = s(\pi(t))} \sum_{k \in \texttt{tasks}(t)} J^0_{m,k} + J^{\texttt{idle}}_{s,m} = i^0_{s,m} \qquad \forall s \in \mathcal{S}, m \in \mathcal{M} \tag{3.76}$$

$$\sum_{t \in \mathcal{T}: s = s(\sigma(t))} \sum_{k \in \texttt{tasks}(t)} J^\infty_{m,k} + J^{\texttt{idle}}_{s,m} = i^\infty_{s,m} \qquad \forall s \in \mathcal{S}, m \in \mathcal{M} \tag{3.77}$$

$$J_{m,k,k'} \in \{0,1\} \qquad \forall t \in \mathcal{T}, k \in \texttt{tasks}(t), t' \in \mathcal{T}, k' \in \texttt{tasks}(t'), m \in \mathcal{M} \tag{3.78}$$

$$J^0_{m,k}\ ,\ J^\infty_{m,k} \in \{0,1\} \qquad \forall t \in \mathcal{T}, k \in \texttt{tasks}(t), t' \in \mathcal{T}, k' \in \texttt{tasks}(t'), m \in \mathcal{M} \tag{3.79}$$

$$J^{\texttt{idle}}_{s,m} \in \mathbb{Z}_+ \qquad s \in \mathcal{S}, m \in \mathcal{M} \tag{3.80}$$

Constraints (3.74) state that a unit of type m that is uncoupled from task k is either coupled to some other task k' in a later trip departing from the same station or stays at that station until the end of the planning period. If the unit is coupled to a later departing trip it must respect the repositioning time. In other words, the coupling time of $\texttt{trip}(k')$ must be no later than the uncoupling time of $\texttt{trip}(k)$.

Similarly, constraints (3.75) ensure that a unit of type m that is coupled to task k was either previously uncoupled from some task k' in a trip arriving at the same station or has stayed at that station since the beginning of the planning period.

Constraints (3.76) state that the units of type m at station s are either coupled to a departing trip or idle during the planning period. Similarly, constraints (3.77) describe that the units of type m that end at station s are either uncoupled from an arriving trip or have been idle the entire planning period.

The following proposition states that the Task Model is equivalent to the Duty-Task Model from the feasibility point of view.

Proposition 8. *The Task Model has a solution if and only if the Duty-Task Model has a solution. The solutions may use the same values for T-, R-, C- and U- variables.*

Proof. The proposition follows from the observation that the variables $J_{m,k,k'}$, $J^0_{m,k}$, $J^\infty_{m,k}$ and $J^{\texttt{idle}}_{s,m}$ in the Duty-Task Model represent a path decomposition of the variables $I_{s,m,d}$ in the Task Model. □

Constraints (3.78) – (3.80) state the domains of the variables. They are declared integral, but according to the following proposition we may declare them continuous.

Proposition 9. *Suppose we relax the integrality of the variables $J_{m,k,k'}$, $J^0_{m,k}$, $J^\infty_{m,k}$ and $J^{\texttt{idle}}_{s,m}$. Then for a given solution to the Duty-Task Model there exists an integral solution with an objective value that is at least as good.*

Proof. According to Proposition 2, Proposition 3 and Proposition 8 the variables $T_{m,k}$, $R_{m,k,k'}$, $U_{m,k}$ and $C_{m,k}$ are integral in a feasible solution. Consider the network flow problem induced by constraints (3.74) – (3.77) where $U_{m,k}$ and $i^0_{s,m}$ are sources, and $C_{m,k}$ and $i^\infty_{s,m}$ are sinks. Since the outflow of the sources and the inflow in the sinks are integer, there exists an integer valued flow on all J variables with an objective value that is at least as good. □

Comparison with the extended Composition Model

In Section 3.4.7.2 we introduced an extension of the Composition Model which allows us to encourage keeping some of the original duty structure in the circulation generation phase. We defined the set $\mathcal{N}$ of uncouple-couple relations where $n \in \mathcal{N}$ denotes that $r(n)$ units of type $m(n)$ were uncoupled at connection $c_1(n)$ and later coupled at connection $c_2(n)$.

The Duty-Task Model is more refined than the extended Composition Model since positions of units are explicitly considered – in the extended Composition Model only the pairs of connections at which uncoupling/coupling takes place are considered. The inventory variables in the Duty-Task Model allow the user to penalize or stimulate duty structures in a more detailed way since the paths of all units in the network are tracked.

As far as feasibility is concerned, the Duty-Task Model and the extended Composition Model are equivalent since a feasible solution in one model can be translated to a feasible solution in the other. We can apply the same cost structure in the Duty-Task Model as suggested for the extended Composition Model to make equivalent feasible solutions to the two models to have the same objective value. This is achieved by taking the above mentioned relations $\mathcal{N}$ into account in the Duty-Task Model – we add similar variables S_n for $n \in \mathcal{N}$ with $S_n \in \mathbb{Z}_+$, and constraints (3.81) and (3.82).

$$S_n \leq r(n) \tag{3.81}$$

$$S_n \leq \sum_{t \in \mathrm{OUT}_{c_1(n)}} \sum_{k \in \mathtt{tasks}(t)} \sum_{t' \in \mathrm{IN}_{c_2(n)}} \sum_{k' \in \mathtt{tasks}(t')} J_{m(n),k,k'} \tag{3.82}$$

Along with the constraints, the same term used in the extension of the Composition Model, (3.38), is added to the objective function. The constraints (3.81) and (3.82) ensure that the number of units of the appropriate type involved in a relation is at most the number of units actually transferred between the connections, and at most the number of units transferred in the original duties $r(n)$.

Name	# RS types	# units in current	# units in modified	# trips in current	# trips in modified	# comp.	type	remarks
S3000-D1-P1	2	11, 9	11, 9	471	270	7	Disruption	
S3000-D1-P2	2	11, 9	11, 9	471	270	7	Disruption	High demand
S3000-D2-P2	2	11, 9	11, 9	471	458	7	Disruption	High demand
S3000-D3-P2	2	11, 9	11, 9	471	455	7	Disruption	High demand
S3000-H1-P3	2	11, 9	11, 9	471	471	7	Holiday	-30% pass.
S3000-H2-P4*	3	10, 11, 11	10, 11, 11	471	471	111	Holiday	+10% pass. Types: 2,3,4
S3000-H3-P4*	3	10, 10, 7	10, 10, 7	471	471	30	Holiday	+10% pass. Types: 3,4,5
S3000-B1-P1	2	10, 11	9, 10	471	471	7	Breakdown	
S3000-B2-P1	2	10, 11	8, 9	471	471	7	Breakdown	
NO-D1-P1	2	71, 44	71, 44	853	823	30	Disruption	
NO-D2-P1	2	71, 44	71, 44	853	831	30	Disruption	
NO-H1-P4	2	71, 44	71, 44	853	853	30	Holiday	+10% pass.
NO-H1-P5	2	71, 44	71, 44	853	853	30	Holiday	+20% pass.
NO-B1-P1	2	71, 44	68, 42	853	853	30	Breakdown	
NO-B2-P1	2	71, 44	66, 41	853	853	30	Breakdown	

Table 3.2: Test instances. The table shows the number of rolling stock types, the number of available rolling stock units in the current and modified situation, the number of trips in the current and the modified timetable, the number of available compositions on each trip, the type of instance, and some remarks on the particularities of each instance.

3.8 Computational comparison of models

In this section we compare the described models for rolling stock rescheduling computationally. In particular, we compare the Composition Model, the Task Model, and the Duty-Task Model for the circulation generation. The comparison includes solvability and solution time of instances of varying size. Furthermore, we provide a comparison of the duties resulting from the two step procedure versus the duties resulting from the integrated Duty-Task Model. Finally, we examine the value of the flexible turning pattern extension.

3.8.1 Test instances

In order to test the models we use the set of instances described in Table 3.2. The instances in the table are (variants of) instances encountered in various applications of the Rolling Stock Rescheduling Problem from other chapters in this thesis. The name of an instance, such as "NO-H1-P5", consists of three parts, first the lines involved, second, the type of instance, and third, the passengers in the instance. The table contains information on the number of available rolling stock types, the number of rolling stock units in the current and the modified situation, the number of trips in the current and the modified situation, the number of different allowed compositions on each trip, and the type of instance.

The type of instance is either disruption, holiday, or breakdown. For a disruption, some of the trips are canceled in the modified timetable and the rolling stock has to be rescheduled accordingly. For a holiday instance, the number of passengers differs between the current and modified situation. In a breakdown instance less rolling stock units are available in the modified situation.

The instances involve either the 3000 line (named "S3000") or the Noord-Oost lines (named "NO"). The 3000 line is a simple line which consists of a total of 471 trips, whereas the Noord-Oost lines have a more complicated structure and involve both splitting and combining of trains. The Noord-Oost lines consist of 853 trips. The 3000 line is served by units of the type VIRM which are available in two types with 4 and 6 carriages, respectively. The possible combinations of these types result in 7 allowed compositions. The Noord-Oost lines are operated by the so called Koploper units which are available in lengths of 3 and 4 carriages. This results in 30 allowed compositions.

The rolling stock types of two instances are different from the rest. Instance "S3000-H2-P4*" is operated using artificial rolling stock types available in lengths of 2, 3 or 4 carriages which allows 111 compositions. Instance "S3000-H3-P4*" is operated using artificial rolling stock types consisting of 3, 4 or 5 carriages which allows 30 different compositions. The rolling stock types used in these instances have been artificially constructed and do not represent realistic sets of rolling stock types. The two sets of types have been constructed for the purpose of testing the effect of adding more types to the models.

The objective of the disruption-instances is primarily concerned with minimizing the changes to the shunting plans. The objective of the holiday- and breakdown-instances is more concerned with assigning enough capacity to the given demand while minimizing operational cost. The exact structure of the objective function, however, is not relevant in this context since the computational results are meant for comparing the performances of the models rather than exploring the results.

	Composition Model				Task Model				Duty-Task Model			
Name	Int.	Cont.	Cons.	Obj terms	Int.	Cont.	Cons.	Obj terms	Int.	Cont.	Cons.	Obj terms
S3000-D1-P1	1890	11880	8778	6734	9552	7006	10226	6734	9552	102401	11852	6734
S3000-D1-P2	1890	11880	8778	6776	9552	7006	10226	6776	9552	102401	11852	6776
S3000-D2-P2	2804	17604	13606	10652	13754	10550	15627	10652	13754	149955	18227	10652
S3000-D3-P2	2815	17646	13600	10679	13826	10591	15663	10679	13826	151880	18245	10679
S3000-H1-P3	2895	18254	14016	11294	14306	10871	16103	11294	14306	160416	18781	11294
S3000-H2-P4*	13803	228820	101527	189173	222914	34002	48784	189173	222914	1042743	60205	189173
S3000-H3-P4*	7002	65052	36916	49147	59146	21439	30971	49147	59146	418343	37142	49147
S3000-B1-P1	2895	18254	14016	11239	14306	10867	16099	11239	14306	160412	18777	11239
S3000-B2-P1	2895	18254	14016	11240	14306	10867	16099	11240	14306	160412	18777	11240
NO-D1-P1	8803	97380	44208	74669	90340	30684	43690	74669	90340	1239091	52268	74669
NO-D2-P1	8881	98521	44608	75648	91423	30941	44070	75648	91423	1270116	52738	75648
NO-H1-P4	9113	102585	45818	79169	95335	31765	45246	79169	95335	1367734	54168	79169
NO-H1-P5	9113	102585	45818	79169	95335	31765	45246	79169	95335	1367734	54168	79169
NO-B1-P1	9113	102585	45818	79169	95335	31765	45246	79169	95335	1367734	54168	79169
NO-B2-P1	9113	102585	45818	79169	95335	31765	45246	79169	95335	1367734	54168	79169

Table 3.3: The size of the models for the instances. The models considered are the Composition Model, the Task Model, and the Duty-Task Model. The four columns for each model denote the number of integer variables, the number of continuous variables, the number of constraints, and the number of terms in the objective function.

3.8.2 Dimensions and solvability of the models

For each model, the number of integer and continuous variables, the number of constraints, and the number of terms in the objective function are shown in Table 3.3. The Composition Model and the Task Model have comparable numbers of variables and constraints in all instances although the majority of variables are continuous in the Composition Model while the opposite is the case for the Task Model.

The Duty-Task Model and the Task Model have the same number of integer variables, but the number of continuous variables is much larger in the Duty-Task Model due to the enlarged description of the inventory.

All models have the same number of terms in the objective function since the objective is linear in the composition and transition variables which are expressed in the same manner in all three models.

We solved the models using CPLEX 11.0 on an Intel Core 2 duo 3.33 GHz desktop computer with 3 GB of RAM. Table 3.4 shows information on the solvability of the models. First, the table shows the LP relaxation of each model. Furthermore, the time at which a solution is found with a provable 5% optimality gap is shown as well as the final solution and the computation time of the final solution. We allowed a maximum of 2 hours of computation time.

	Composition Model				Task Model				Duty-Task Model			
Name	Root LP	Time for 5%	Final	Time	Root LP	Time for 5%	Final	Time	Root LP	Time for 5%	Final	Time
S3000-D1-P1	93224.3	3.8	93575.0	3.8	91737.0	5.2	93575.0	9.2	91737.0	43.1	93575.0	112.0
S3000-D1-P2	94783.8	6.1	95602.0	8.7	92383.1	22.8	95602.0	94.5	92383.1	166.5	95602.0	228.1
S3000-D2-P2	109974.0	0.1	110360.0	0.1	107882.0	0.2	110360.0	0.2	107882.0	0.7	110360.0	0.7
S3000-D3-P2	132893.7	4.3	135318.0	4.9	126746.8	13.0	135318.0	16.2	126746.8	71.3	135318.0	160.6
S3000-H1-P3	75904.5	5.0	76149.0	5.1	75407.1	9.0	76149.0	25.1	75407.1	55.4	76149.0	98.5
S3000-H2-P4*	80530.5	-	-	7200.0	78632.0	-	-	†	78632.0	-	-	†
S3000-H3-P4*	88665.5	3139.3	88976.0	3500.2	87586.7	-	-	7200.0	87586.7	-	-	†
S3000-B1-P1	84588.0	0.8	84627.0	0.8	84588.0	1.5	84627.0	1.5	84588.0	17.0	84627.0	17.0
S3000-B2-P1	88278.0	5.0	88527.0	5.0	88278.0	5.6	88527.0	9.4	88278.0	265.8	88527.0	267.0
NO-D1-P1	434133.5	168.8	437379.0	181.6	414438.2	1451.3	437556.0	7200.0	414438.2	-	-	†
NO-D2-P1	401694.7	1.2	401755.0	1.9	400906.7	2.9	401755.0	2.9	400906.7	9.3	401755.0	11.9
NO-H1-P4	527313.9	4239.2	530698.0	5536.1	512075.7	-	-	7200.0	512075.7	-	-	†
NO-H1-P5	799839.8	4661.9	806537.0	7200.0	775563.8	-	-	7200.0	775563.8	-	-	†
NO-B1-P1	403625.9	547.8	404419.0	603.1	397302.8	4864.9	405641.0	7200.0	397302.8	-	-	†
NO-B2-P1	428818.4	544.7	429798.0	626.9	417940.9	-	-	7200.0	417940.9	-	-	†

Table 3.4: The LP relaxation, solution value and solution time for each model for all instances. The table also shows the amount of time spent to find a solution with a proven 5% optimality gap. Solution times marked with † denote that the MIP solver ran out of memory.

We applied a number of speed-up techniques to the MIP solving procedure for the models. For the Composition Model we applied a technique discussed by Maróti (2006) for reducing the number of Integer variables. For the Task Model and the Duty-Task Model, we investigated the use of branching priorities and parameter tuning of the MIP solver although the investigation was not exhaustive. Our attempts, however, failed to consistently improve the performance of the MIP solving procedure for the models.

We observe that the LP relaxation of the Composition Model is generally stronger than the LP relaxation of the other models. This indicates that the Composition Model is likely to be easier to solve by an LP relaxation based branch-and-bound algorithm. Indeed, we observe that for all instances the Composition Model is solved faster than the Task Model which is again faster to solve than the Duty-Task Model. For some of the instances the MIP solver ran out of memory. This is particularly the case for the Duty-Task Model when solving the "NO"-instances.

Furthermore, we observe that the models for the disruption instances (except "NO-D1-P1") are solved relatively fast. Short computation times are crucial for instances involving disruptions since they are solved in a time-critical environment. There is, however, no guarantee on the expected computation time of an instance, but in Chapter 4 we explore decomposition methods to decrease the solution time of rolling stock rescheduling during disruptions.

The two instances "`S3000-H2-P4*`" and "`S3000-H3-P4*`" that use artificial sets of rolling stock types are difficult to solve with all models. A closer examination of the branch-and-bound procedure of CPLEX reveals that the large number of available compositions with similar contributions to the objective function hampers the algorithm's ability to prune the MIP node-tree.

We conclude that the Composition Model performs best for the circulation generation procedure and we use therefore this model in all applications of the rolling stock rescheduling problem in this thesis.

3.8.3 Duty generation

In Section 3.6 we introduced two models for the duty generation problem, i.e. the process of translating a rolling stock circulation to a set of rolling stock duties. Furthermore, in Section 3.7 we proposed an extension of the Task Model, called the Duty-Task Model, which integrates circulation and duty generation. The goal of the duty generation is to create a set of rolling stock duties for the modified situation that resemble the current duties as much as possible.

Concerning computation times we note that the Duty Flow Model is solved in a fraction of a second for all the instances we have encountered, while the Duty Path Model takes up to a few seconds to solve complex instances. In either case the computation time of the duty generation step is neglectable.

The Duty Flow Model and the Duty Path Model use different schemes to minimize the changes to the duties in that the Duty Path Model allows the user to penalize changes at a more detailed level. The choice of the duty generation model depends on the preferences of the decision maker; the computation time is not a limiting factor in this choice. It is, however, out of the scope of this thesis to fully investigate the practitioners' preferences for the duty generation objectives. We did manage to get a qualitative comparison by a practitioner of the two models on some instances. For this comparison we used the visualization tool described in Section 6.3.4. The practitioner preferred the Duty Path Model as it seems to provide visually fewer changes to the duties.

We investigated the relationship between keeping the uncoupling/coupling patterns, and the similarity of original and modified duties. Recall from Section 3.6.2 that we defined the similarity of an original duty and a potential new duty as a number between 0 and 1. The similarity between the current duties performed by a

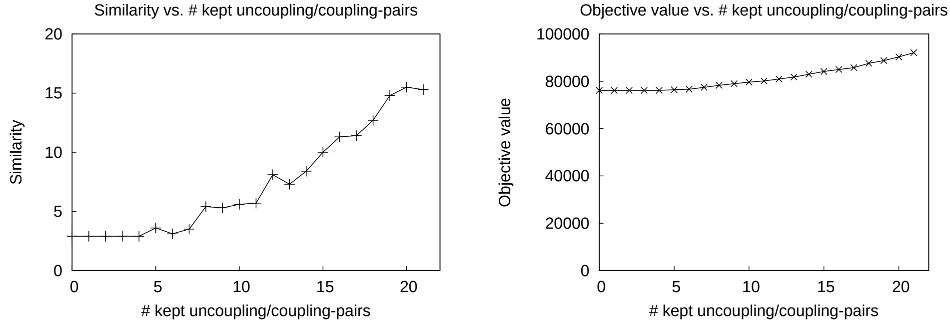


Figure 3.16: Experiments with the Duty-Task Model for instance `S3000-H1-P3` with varying number of intact uncoupling/coupling pairs. *Left*: Cumulative similarity versus the number of intact uncoupling/coupling pairs. *Right*: Objective function of the circulation generation process versus number of intact uncoupling/coupling pairs.

set of rolling stock units and the corresponding modified duties is defined as the sum of the similarity measure between the current and modified duty for each unit.

In the Duty-Task Model we directly model the uncoupling/coupling pattern of the rolling stock units and we can therefore control how many pairs of uncoupling and later coupling from the current duties we want to keep in the modified duties. In this study we limit ourselves to the instance `S3000-H1-P3`. The current rolling stock schedule of the instance consists of 20 duties with a total of 21 pairs of underway uncoupling and later coupling of units. We used the Duty-Task Model to keep between 0 and 21 of the current uncoupling/coupling pairs and observed the similarity of the duties and the objective function of the circulation generation process. The left diagram in Figure 3.16 shows the relationship between the number of kept uncoupling/coupling pairs and the similarity between current and modified duties. The right diagram in Figure 3.16 shows the objective value of the circulation generation problem versus the number of kept uncoupling/coupling pairs.

Clearly, keeping more uncoupling/coupling pairs results in modified duties that are more similar to the current duties. At the same time the rescheduling costs increase as we attempt to keep more uncoupling/coupling pairs. An analysis of the modified duties reveals that this is due to the fact that there are 30% less passengers in the modified situation of instance `S3000-H1-P3` and therefore there is a significant potential for saving carriage-kilometers. The more similar the modified duties are to the current ones, the less carriage-kilometers are saved.

The duty generation process thus reveals another trade-off in the rolling stock rescheduling process: Less changes to the rolling stock duties comes at the cost of higher circulation rescheduling costs.

3.8.4 Computational test of additional features

In this section we perform computational tests of some of the additional features described in Section 3.4.7. In particular, the extensions that allow flexible turning patterns at terminal stations show interesting implications for practical applications.

3.8.4.1 Flexible turning patterns

In Section 3.4.7.1 we introduced an extension of the Composition Model to allow flexibility in the turning patterns at terminal stations. We implemented this feature and tested it on a number of disruption instances from Table 3.2.

The results are shown in Table 3.5. For the instances `S3000-D1-P2`, `S3000-D3-P2` and `NO-D1-P1` we report the reference solutions with computation time and objective value when no turning pattern flexibility is used. By introducing turning pattern flexibility at certain stations we constructed several solutions with varying numbers of trains that turn differently than planned. Further, we created some solutions where some trains not only turned differently, but also uncoupled or coupled one or more units. In the table we report the objective value measured by the same function as the reference solution, and the number of trains that turn differently at each terminal station. We also report the number of carriage and seat shortage kilometers in the solutions as they are major contributors to the objective value.

Instance `S3000-D1-P2` concerns the rolling stock schedules for a weekday on the 3000 line. In this instance the infrastructure north of Alkmaar is blocked for the entire day. We introduced turning pattern flexibility at the southern terminal station of Nijmegen (Nm) and constructed a number of solutions. We observe that by turning some trains out of their usual turning pattern we are able to decrease the objective value.

In instance `S3000-D3-P2` there is a disruption between Schagen (Sgn) and Alkmaar (Amr) on the 3000 line for four hours on a weekday. We constructed solutions that use flexible turnings at either the southern terminal station of Nijmegen (Nm), the northern terminal station Den Helder (Hdr), or both. We observe that we are able to significantly reduce the other rescheduling costs by turning more flexibly – especially at Nm. A more detailed inspection of the solutions reveals that the reduction

Instance	Remark	obj. value	carr KM	seat KM	DT Nm	DTS Nm	DT Hdr	DTS Hdr	Time
S3000-D1-P2	reference	95602	82116	3686	–	–	–	–	8.7
	flex Nm	95016	79840	4256	3	0	–	–	10.5
	flex Nm	92898	78112	4976	5	0	–	–	5.0
	flex Nm	92208	78672	4256	7	0	–	–	7.2
	flex Nm	91294	80272	2162	12	0	–	–	3.9
	flex Nm	92678	78112	4256	5	2	–	–	6.2
S3000-D3-P2	reference	135318	93048	28410	–	–	–	–	4.9
	flex Nm	126241	90464	25618	4	0	–	–	1.0
	flex Nm	124873	92660	21534	4	1	–	–	3.2
	flex Nm	123460	91628	19174	7	2	–	–	11.0
	flex Hdr	129684	92260	28544	–	–	2	0	4.1
	flex Hdr	128384	92260	28544	–	–	4	0	5.5
	flex Hdr	126508	91896	24262	–	–	8	0	0.7
	flex Hdr	126061	93296	24262	–	–	7	1	0.8
	flex Nm, Hdr	123396	90624	26522	7	0	2	0	1.3
	flex Nm, Hdr	122846	90064	23322	9	0	4	0	1.3

Instance	Remark	obj. value	carr KM	seat KM	DT Rtd	DTS Rtd	DT Gn	DTS Gn	Time
NO-D1-P1	reference	437379	357140	67543	–	–	–	–	181.6
	flex Rtd	432623	353604	64933	2	0	–	–	1286.6
	flex Rtd	431579	353212	63451	6	0	–	–	227.3
	flex Rtd	430830	352564	63451	5	1	–	–	368.9
	flex Gn	436960	358056	63879	–	–	3	0	194.8
	flex Gn	435745	357496	57357	–	–	2	2	213.3
	flex Rtd, Gn	430712	354762	61244	7	0	3	0	469.1

Table 3.5: Experiments with disruption instances using flexible turning patterns at certain stations. For each instance we report the reference solution without flexible turnings and a number of solutions with flexible turnings at one or more stations. The column abbreviations are: carr KM = carriage kilometers, seat KM = seat shortage kilometers, DT Nm = different turning at station Nm, DTS Nm = different turning with shunting at station Nm.

in objective value primarily comes from the need for less new shunting operations, less carriage kilometers and much less seat shortage kilometers.

The instance `NO-D1-P1` involves a disruption on the Noord-Oost lines which is a collection of lines that share rolling stock. To test the extension to the model we have chosen to apply turning pattern flexibility at the stations Rotterdam (Rtd) and Groningen (Gn). We observe that the application of flexible turnings at especially Rtd results in a significant reduction of the objective value.

We note that the computation times for the instances involving the 3000 line with turning pattern flexibility are comparable to the computation time for the reference solutions. For the instances involving the Noord-Oost lines the computation times are consistently longer than for the reference solution – for some instances the increase in computation time is marginal, but for one solution the computation time is seven times as long.

3.9 Conclusion and future research

In this chapter we formalized the Rolling Stock Rescheduling Problem. It is the problem of rescheduling a set of rolling stock duties to a modified situation where the timetable, the amount of available rolling stock or the passenger demand is adapted. We utilized a two step procedure to first generate a circulation of rolling stock units and then translate the circulation to a set of modified rolling stock duties.

We formulated two mathematical models for the circulation generation problem; the Composition Model which is based on a multi-commodity flow model constrained by the assignment of allowed compositions to trips and composition changes to connections; and the Task Model which is based on the flow of rolling stock units through the tasks associated with the trips in the modified timetable. Computational tests indicate that the Composition Model generally performs better than the Task Model for the circulation generation process.

For the Duty generation process we developed two models: the basic Duty Flow Model based on a single commodity network flow in the duty graph, and the Duty Path Model which is based on a covering of the duty graph with paths. The second model fully contains the first and further allows the decision maker to construct modified duties based on their similarity with the corresponding original duties. Computational tests show that both models can be solved by an industrial strength MIP solver within a few seconds.

Considering the expressive power and computational tests of the studied models, we decided to utilize the Composition Model for the circulation generation process and the Duty Path Model for the duty generation process. These models are used for all further applications of the Rolling Stock Rescheduling Problem in this thesis.

Future research

The models investigated in this chapter for the circulation generation step utilize a simplified model for the shunting and storage of idle rolling stock units. The shunting and storage is modeled through the inventory which keeps track of the number of units of each type that is available for coupling at a station at any given time.

The inventory abstraction is generally acceptable for rolling stock rescheduling applications where storage capacity in depots is not a severely limiting factor. This is usually the case during the day where most rolling stock is in use and most of the storage capacity is not in use. However, shunt yard capacity is more binding at stations with small depots and during night time at larger stations. The rolling stock duties constructed by the methods in this chapter may be awkward in rare cases since the actual topology of the shunt yards is not directly taken into account. It could be interesting to investigate the integration of stronger models for the shunting and storage of units in such situations. We mentioned earlier that Lentink (2006) describes a range of models for the shunting processes at NS. We suggest to investigate whether these models could be (partially) integrated in the rolling stock rescheduling process.

Furthermore, the applied inventory model assumes that trains are always split up into separate rolling stock units when sent to the inventory. However, in practice it may be preferable to keep units in their arriving composition in the inventory so that they can be utilized later without changing the composition. We therefore suggest to investigate inventory models that incorporate the option of keeping units in compositions rather than as separate units.

Another issue is the structure of the resulting rolling stock duties. For the rolling stock rescheduling applications considered in this thesis we consider the duty generation objective which states that the modified duties should resemble the current ones as much as possible. This is a rather soft requirement that does not incur any hard constraints on the structure of the duties. Hard constraints in this context are issues such as refueling and maintenance constraints which require each duty to periodically pass by refueling and maintenance facilities. These issues are generally more binding in railway systems operated by diesel units, and for long distance or international train services where units operate far away from their designated maintenance facilities.

It is not possible to explicitly incorporate hard constraints on the duty structure in the two-step method proposed in this chapter. The reason is that units are not considered individually in the circulation generation phase, but rather at an aggregated level. This means that any constraints that require some duties to visit certain facilities may render the duty generation models in their current state infeasible.

For railway settings with hard constraints on the duty structure it could be fruitful to investigate models that track units more explicitly already when constructing the circulation. This could be achieved by considering the path of units in the network used by the Task Model. Recall that units are aggregated into types in the Task Model, but the units are modeled by a 0/1-flow along task arcs. A path in this network thus represents a duty. It could therefore be interesting to investigate a path formulation variant of the task model where duties are explicitly expressed as paths. To solve such a model one should likely utilize a branch-and-price approach.

Chapter 4

A Framework for Disruption Management in Passenger Railways

The uncertainty characterizing disruptions in passenger railways is related to the state of the system, the availability of resources, and further developments in the system's state.

In this chapter we discuss disruptions and disruption management in passenger railways, and develop a framework for modeling the dynamics of passenger railway disruptions.

Structure of the chapter

We first discuss common causes of disruptions and how they impact the railway system. We then describe a generic model for disruptions in the railway system that incorporates the uncertainty of the situation by gradually updating the available information on the state of the system. Next, we incorporate this disruption model in a framework for the rolling stock rescheduling stage in the disruption management process, and test the concept on instances derived from real-life schedules from NS. The findings are based on Nielsen and Maróti (2009) and Nielsen et al. (2010).

The notions of disruptions and disruption management discussed in this chapter come from observations from practice combined with the disruption management applications described in the literature by Yu and Qi (2004), Filar et al. (2001), Kohl et al. (2007) and Clausen et al. (2010). These studies describe frameworks for disruption management in the airline industry and other large-scale logistic settings. While Section 2.3 described the organizations involved in the overall operational process, this section focuses more on the process of actually handling disruptions.

4.1 Disruptions in a passenger railway system

A disruption in a passenger railway system is an event or a series of events caused by external or internal factors that leads to substantial deviations from planned operations. In either case, new restrictions are imposed on the railway system. These restrictions may have local or regional effects and may involve railway resources, infrastructure components or both.

External factors

External factors leading to disruptions over which the railway operator or railway system have no control include weather conditions and accidents involving other traffic.

Weather conditions such as blizzards and storms may reduce the maximum speed of trains or even prevent trains from running in certain regions. Weather-related train cancellations are usually imposed for safety reasons, both for travelers and personnel but also to prevent rolling stock and infrastructure components from sustaining further damage. Weather conditions are characterized by being literally impossible to prevent, they are difficult to predict and they strike indiscriminately. There may be up to a few days warning before severe weather conditions occur, but information on locality and severity is usually characterized by significant uncertainty.

Accidents with other traffic are a frequent cause of disruptions. In such a case, an investigation of the circumstances of the collision must be conducted by the authorities. Such accidents and their investigation usually prevent train travel on the involved tracks. The investigation itself may take hours, but if the infrastructure or rolling stock has been damaged the effects may disrupt train traffic even longer. Also, the rolling stock is likely to need repairs after being involved in a collision.

Power outages and cell phone network failures disrupt train traffic in whole regions depending on the area affected by the power outage or the failing local cell phone network. Of course, a railway operator that primarily operates electrical rolling stock seems especially vulnerable to power outages, but an operator that uses diesel units is unlikely to be better off in this situation as signaling systems and switches are out of function. Power outages can to some extend be prevented by the use of emergency power systems, but such systems are very costly to install and operate. A failing cell phone network prevents the communication between dispatchers and crew, and the tracking and tracing of the crew resources. This disrupts the execution of many tasks. In either case it is uncertain how long the disruption lasts. In particular, the railway operator has little influence on the repairs and the further development of the disruption.

Internal factors

Internal factors leading to disruptions in the railway system are generally related to the availability of resources, including rolling stock breakdowns and crew shortages.

A *rolling stock breakdown* disrupts the train traffic in several ways. First, the broken down rolling stock blocks part of the infrastructure which decreases the capacity of the system until the broken down rolling stock is removed. Second, broken down rolling stock cannot be assigned to any train services before it has been repaired. Therefore a rolling stock breakdown decreases the overall resource availability of the system. Third, passengers may be stuck in a defective train. A response to the breakdown thus requires proper handling of the passengers as well.

Crew shortages are a frequent cause of disruptions. A train is generally not allowed to drive without the proper number of conductors and of course a driver. The reasons for crew shortages are numerous and may be direct, such as due to illness, or caused by other incidents, such as a disrupted train carrying crew for other trains. In more severe cases crew shortages are on a regional or national scale.

Another common cause of disruptions in the railway system is *malfunctioning infrastructure*. This category includes rails, switches, signals, drawbridges, road-intersection barriers, overhead electricity wires and central computer systems. If some infrastructure component is failing it usually requires manual repair which may take hours or even days depending on the location and the severity of the failure. A failing computer system may suspend train traffic if signals and switches do not function properly.

4.2 Features of disruptions

Regardless of the cause of a disruption it has an impact on the railway system. That impact is generally in the form of a change in the system settings, a change in resource availability, or both. We here discuss some overall aspects of the impact of disruptions.

4.2.1 Change in system environment

A disruption may cause a change in the environment in which the system operates. The effect may be decreased maximum speed of trains or even prevent trains from running on certain parts of the network.

Closing a station (or part thereof) temporarily is another example of a change in the system environment that affects the system's ability to operate. In such a situation it may be impossible to accommodate all the trains with planned stops at the involved station.

A further change in the system environment is a deviation in demand. In the passenger railway setting this refers to the flow of passengers which may deviate significantly from the forecasted traveling patterns under certain conditions. Especially a blockage in the system is likely to heavily affect demand for capacity on alternative routes.

4.2.2 Change in resource availability

Disruptions usually involve a change in resource availability. This is the case when rolling stock breaks down or when crew calls in sick. The response to a change in resource availability is to replan the current operations to apply only the available resources which may include giving up some of the planned services.

A change in the availability of resources may be the cause of the disruption or a consequence of the disruption. As mentioned, the disruption may indeed be caused by rolling stock breakdowns, or a lack of crew, or a blockage may prevent rolling stock and crew from being at the right place. Additionally, some resources may be taken out of service to undergo preventive maintenance or to build up a buffer of additional resources for an upcoming event. For example, some rolling stock may be removed from circulation if a snowstorm is expected to ensure that a certain amount of functional rolling stock is available when the weather conditions have settled.

4.2.3 Predictability of events

The events that lead to disruptions in the railway system are generally hard to predict and if an event is anticipated, the consequences are even harder to forecast. Some events are completely spontaneous and strike without warning. This is generally the case with collisions with other traffic and power outages.

Other events are predictable to a certain extent such as adverse weather conditions. Another example is during a flu epidemic where a given share of the population is sick. Then we would expect a similar share of the personnel to be sick too. However the exact shortage of staff is not known beforehand as well as whether the sick share of the personnel includes key personnel that cannot be replaced on the short term.

4.2.4 Time-critical environment

The decision making process associated with disruption management exists in a time-critical environment. After the occurrence of a disruption, the first decisions on how to adapt the system must be made within a few minutes. Those early decisions mostly relate to which handling scenario to use and how to shift to the situation described in the scenario. The handling scenario then describes where and when to turn trains in case of infrastructure blockages and which train services to give up if needed.

For the rolling stock decisions during a disruption, the major issue is how to assign capacity to trains. These decisions are less pressing than the timetable related decisions right at the occurrence of the disruption since it is possible to more or less keep the rolling stock assigned to their current trains even if those trains are over- or undercapacitated. Therefore the time window for the rolling stock decisions is somewhat longer than for the timetable decisions.

The available time window for decision making during disruptions necessitates some specific requirements for the design of a system for assisting the dispatchers in the disruption management process. A decision support system that cannot suggest a solution or at least a partial solution within a few minutes is of limited value in this time-critical environment.

4.3 Uncertainty

One of the main features of disruptions is the accompanying uncertainty. We have already mentioned one level of uncertainty, namely the predictability of the events that lead to disruptions. But also the uncertainty of the impact is an important characteristics of the disruption.

4.3.1 Uncertainty of impact

Generally it is difficult, if not impossible, to forecast the exact impact of a disruption when it occurs. However, for some localized disruptions we know from experience what the likely outcome of the situation is.

Consider the example of a failing switch at a busy railway node. First, a problem with the switch is detected. Then, a repair crew is dispatched to assess the state of the problem and repair it. When assessing the damage the repair crew determines how much time is needed before the switch is again functional. However, this estimate of the time needed for repair may be given as an interval such as two to four hours, adding to the uncertainty of the situation.

Another example is the occurrence of accidents. In such an event, the police must conduct an investigation. Independently of the outcome of the investigation the involved infrastructure is blocked for an uncertain amount of time. From experience the railway operators know that the length of the blockage has a certain distribution, and can therefore react according to a likely or expected outcome, but still the impact is uncertain.

4.3.2 Uncertainty in the state of the system

In addition to the uncertainty of the impact of the disruption, some details of the state of the system may be uncertain as well. For example, exact information on the positions of all rolling stock and crew may not be available in real time. Also, since traffic control and rolling stock dispatching are conducted by different organizations the intention of each dispatcher may not be clear to the others. This is largely due to the locality of information or even miscommunication. In recent years, the application of improved tracking and tracing systems have enhanced the dispatchers' ability to assess the state of the system.

Further, dispatchers have to query their information system to access better information on the state of the system which means that it takes some time to gain an overview of the situation needed to make an informed decision. This delay in decision making due to queries to the system is significant in a time-critical situation such as during a disruption.

4.3.3 Uncertainty in available countermeasures

In a disrupted situation the dispatchers attempt to react by implementing some countermeasures. However, any decisions on changes to the pre-set allocation of resources have to be implemented in practice by the involved personnel. But the dispatcher may not have full information on the workload of that personnel or may not have insight into the exact work methods of all involved organizations. A dispatcher may come up with a possible solution to the problem at hand that seems elegant, but he may not know if it is implementable. Fully evaluating all options in a given situation is time consuming and therefore generally not possible.

In the rolling stock context, available options for rescheduling are linked to the options for changing local operations. In particular, some rescheduling options may involve coupling or uncoupling of different rolling stock units. Since the exact details of these operations are locally planned and performed, it takes time to evaluate if the proposed solution is practically implementable. For drivers and conductors the crew members themselves may reject a proposed solution if it deviates too much from the original schedule.

Instead of investigating the options for and consequences of all immediate countermeasures, the dispatcher relies on his experience for evaluating which countermeasures are most likely to be accepted by all parties.

4.4 A model for disruptions in rolling stock schedules

To develop a decision support system for disruption management we need a way to represent the occurrence of the disruption, the development of the situation, and the accompanying uncertainty. We here briefly discuss how disruptions are modeled in other contexts and how uncertainty can be represented, and we motivate an approach to modeling railway disruptions inspired by current practice.

In several studies of disruption management in various time-critical settings, the associated problems are described as pure *recovery* problems. That is, the problem is to reassign the available resources to a new situation. Teodorovic and Stojkovic (1990), Thengvall et al. (2001), Bard et al. (2001) offer examples of recovery models in various settings in the airline industry. Strictly speaking, such models consider a plan for normal operations and a disruption that poses new constraints thus rendering the current plan infeasible. The problem is then to adapt the plans to take the new constraints into account while minimizing a number of metrics such as operational cost, delays and cancellations.

In general, recovery models consider the disruption as an event that leads to a new situation and do not explicitly consider the dynamics of the disruption. A disruption is namely a developing situation where the knowledge of the state of the system only gradually becomes available. Consider the case of a failing infrastructure component; at first the repair crew may estimate it takes 1.5 hours to repair. However, after one hour the repair crew conclude that their first estimate was too optimistic and provide a new estimate that the total repair time is likely to be one hour longer. This suggests considering the disruption as an online problem rather than a one-time recovery problem.

Extending the example from Section 2.6, a disruption could develop as sketched in Figure 4.1: The train services are planned according to the timetable shown as a time-space diagram in Figure 4.1(a). At time 12:45 an infrastructure malfunction is reported at a point between Alkmaar (Amr) and Schagen (Sgn). The repairs are expected to be done at time 14:15 whereafter the tracks are again opened for train traffic.

According to the precomputed handling scenario the trains are turned at the closest major stations on either side of the blocked section of infrastructure. For this line these are Alkmaar (Amr) and Schagen (Sgn). This turning pattern and the expected duration of the blockage results in the timetable shown as a time-space diagram in Figure 4.1(b).

At time 13:45 it turns out that the problems with the infrastructure are more complicated than first assumed. The repair crew provides a new estimate of the duration of the blockage; it will now last until 15:15. This leads to a new update of the timetable as shown in Figure 4.1(c).

The arrival and departure times are shown as fixed in the updated timetables in the figure. However, in the realization of the timetable departure and arrival times are likely to deviate from those times announced. But from the point of view of the rolling stock such small disturbances have little influence on the feasibility of the rolling stock schedules. Furthermore, if one or more trains are heavily delayed that information can be given as a new timetable update.

Note that the paths of some of the trains in the time-space diagram are changed both when the disruption occurs and at the moment information about the duration of the disruption is updated. For the train following the bold path in the diagram it means that any planned composition changes along the path will have to be rescheduled after the occurrence of the disruption. But when it becomes clear that the blockage lasts longer than first estimated those composition changes have to be rescheduled again.

(a)
(b)
(c)

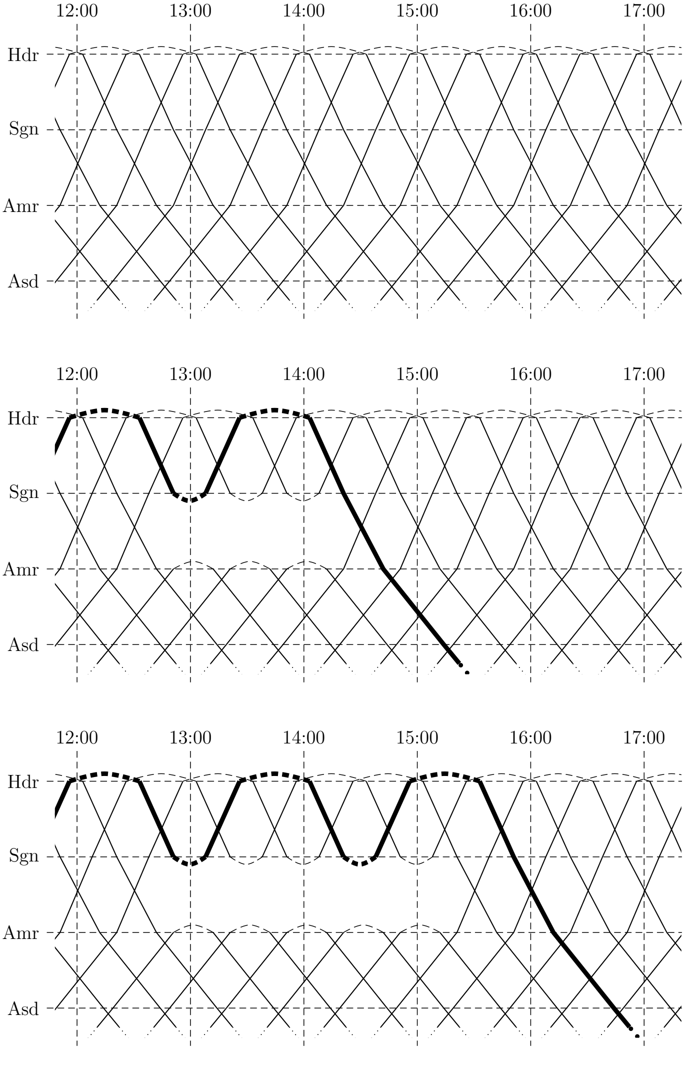


Figure 4.1: (*a*) Time-space diagram for part of the 3000 line. (*b*) Adapted timetable after the occurrence of a disruption between Schagen (Sgn) and Alkmaar (Amr). (*c*) Adapted timetable according to updated information on the disruption.

The above example shows how a disruption may evolve. First, a blockage is detected and the impact of the disruption is estimated. Then, a number of countermeasures is taken, i.e. the timetable is updated and the planned composition changes of some trains are revised. Later still, the situation develops and the blockage is now estimated to last one hour longer. In Section 4.4.1 we discuss different ways to model disruptions while taking their uncertain nature into account.

4.4.1 Scenario-based modeling

One way to accommodate the uncertainty of a disrupted situation is through *scenario-based modeling* where the concept of scenarios is used to model each possible state of the system. If the repair crew in the example in Figure 4.1 concluded that it would take between one and three hours to repair the infrastructure malfunction, then it would result in five scenarios: One scenario assumes the repairs are carried out in one hour and cancel and turn two trains in each direction. Another scenario assumes the repairs take between one hour and one hour and thirty minutes leading to three trains canceled and turned in each direction like in Figure 4.1(b), and so on. In general this results in numerous scenarios.

Classically, two classes of approaches to decision making under uncertainty rely on the concept of scenarios; *Robust Optimization* (see Kouvelis and Yu (1997) and Ben-Tal et al. (2009)) and *Stochastic Programming* (see Ruszczyński and Shapiro (2003)). Robust Optimization attempts to find a solution that in some sense is feasible in all scenarios while minimizing the cost of the worst case behavior of the solution. Approaches based on Stochastic Programming, on the other hand, rely on probability distributions on the outcome of the scenarios, and then the expected outcome is optimized according to the known probability distributions.

There are both pros and cons for such scenario-based approaches. Robust Optimization offers a guaranteed lower bound on the quality of the performance in the given scenarios whereas Stochastic Programming offers an optimal expected performance. There is, however, a number of problems with the applicability of such approaches in a real-time setting; first of all it may be difficult to enumerate all scenarios in a more involved disruption than the example above. Second, even a limited number of scenarios may result in computationally challenging models which is detrimental in a time-critical setting. Third, it is difficult to estimate the probability distributions of the scenarios in practice as needed for the Stochastic Programming approach.

Although scenario-based modeling offers a range of tools for dealing with uncertainty we choose to follow a different paradigm for incorporating the uncertainty of disruptions in the decision making process. We elaborate on this in the following section.

4.4.2 Online combinatorial problem

Rather than trying to describe the possible outcomes at the occurrence of a disruption by enumerating scenarios we describe the uncertainty of the situation as an online combinatorial optimization problem (see Borodin and El-Yaniv (1998)). The occurrence of a disruption and the development of the situation is then described by the following elements.

- The undisrupted scenario $\mathcal{S}_0$.
- A finite list of updates $\langle t_1, \mathcal{S}_1 \rangle, \ldots, \langle t_n, \mathcal{S}_n \rangle$.

Here, the scenario $\mathcal{S}_0$ describes the state of all parameters in the system in the undisrupted situation. An element in the list is a pair consisting of a time instant t_i and a scenario $\mathcal{S}_i$ that describes the system at time t_i. For the time instants we assume that $t_1 < \cdots < t_n$. The element $\langle t_i, \mathcal{S}_i \rangle$ denotes that at time t_i the system changes to $\mathcal{S}_i$. The list thus represents the uncertainty of the development of the situation.

The task is then at time instant t_i to replan the resources in the system so that the plan is feasible under the constraints described by scenario $\mathcal{S}_i$. At this time there is no knowledge of future changes in the system, i.e. the rest of the list $\langle t_{i+1}, \mathcal{S}_{i+1} \rangle$, $\ldots$, $\langle t_n, \mathcal{S}_n \rangle$ is not known. Also, the number of updates n is not known.

In the context of disruption management of rolling stock in a passenger railway system, the scenario $\mathcal{S}_i$ describes the updated timetable, the rolling stock availability, passenger demand, and all parameters on shunting possibilities. The disruption in Figure 4.1 can be modeled by letting $\mathcal{S}_0$ describe the ordinary timetable of Figure 4.1(a) and letting the list consist of two updates $\langle t_1, \mathcal{S}_1 \rangle$, $\langle t_2, \mathcal{S}_2 \rangle$.

The first element is $\langle t_1, \mathcal{S}_1 \rangle$ where $t_1 = 12{:}45$ and $\mathcal{S}_1$ describes the timetable update in Figure 4.1(b). Similarly, the second element in the list is $\langle t_2, \mathcal{S}_2 \rangle$ where $t_2 = 13{:}45$ and $\mathcal{S}_2$ describes the timetable update in Figure 4.1(c).

This way of modeling the progression of a disrupted situation is known in practice as a wait-and-see approach (see Wets (2002)). In this approach, no assumptions are made on the probability distribution of the occurrence of the possible scenarios. Below we discuss some theoretical properties of the problem.

Theoretical competitiveness properties

The following definition by Karlin et al. (1988) measures the performance of an online algorithm by comparing it to the optimal offline algorithm, OPT, which knows the entire input sequence in advance.

Definition 1. *Let $\mathcal{I}$ be the set of all instances for an online minimization problem P. The competitive ratio of an online algorithm $\mathcal{A}$ for P is defined as the smallest number k such that for some constant α,*

$$\mathcal{A}(I) \leq k \cdot \mathrm{OPT}(I) + \alpha \quad \forall I \in \mathcal{I}.$$

An online algorithm $\mathcal{A}$ is said to be k-competitive if the competitive ratio is k. If the algorithm is k-competitive for some constant k, the algorithm is said to be competitive.

We shortly note the following proposition for the online RSRP.

Proposition 10. *No online algorithm for the online RSRP is competitive.*

Proof. Consider a timetable with three trips from station `A` to station `B` where one rolling stock unit is assigned to each trip. If the timetable is changed so that the second trip is canceled then the three rolling stock units must be assigned to the first and the third trip to avoid off-balances. If at least one unit is assigned to the third trip, then a new timetable update with a cancellation of that trip will lead to off-balances. If no unit was assigned to the third trip it leads to a canceled trip due to rolling stock shortage. The offline algorithm, which knows the timetable updates beforehand, will distribute the rolling stock units on the trips that are not canceled, thus leading to a solution with no off-balances or cancellations due to lack of rolling stock. To overcome the additive term α in Definition 1, the construction can be copied an arbitrary number of times. □

We note that Proposition 10 holds for randomized online algorithms too, as either the expected number of off-balances or the expected number of cancellations due to rolling stock shortage will be positive.

The timetable and the updates used in the above informal proof of Proposition 10 are rather unrealistic and only serve a theoretical purpose. Nevertheless, the result does show that there exist instances where any online algorithm may be arbitrarily worse than the optimal offline algorithm.

4.5 Framework for disruption management of passenger railway rolling stock

In this section we describe a framework for disruption management of passenger railway rolling stock. The framework is derived from observations from practice combined with the model for disruptions from Section 4.4.2. We first elaborate on the observations from practice.

4.5.1 Concepts

From observations of current practice and through conversations with planners and dispatchers we have derived a number of concepts in disruption management of passenger railway rolling stock. We here describe those observations and discuss how they are integrated in the framework.

Deviation from planned operations

The assignment of rolling stock in the operations is determined through the planning process described in Chapter 2. The resulting plans are highly optimized with respect to a number of criteria including service and operational cost. In a disrupted situation the operations will have to deviate from the planned operations, but such deviations incur additional costs in the operations.

The costs associated with deviating from the planned operations may be hard to identify, let alone quantify, in a time-critical setting, but they generally refer to either (*i*) direct monetary costs such as fuel, overtime payment for crew, and wear and tear of equipment, (*ii*) indirect monetary costs due to loss of goodwill of passengers and business partners, or (*iii*) destabilization of the system since deviations from the operational plans have to be communicated to and agreed upon by several parties.

It is therefore generally desirable that the rescheduled plans balance staying close to the original plans and conforming to the disrupted situation. This trade-off can be modeled by introducing costs for deviating from the original plans.

Return to the planned situation

Continuing the line of thought from the previous observation, it is clearly desirable to return to the undisrupted situation in a timely manner. The existing operational plans for the normal situation are, as mentioned, highly optimized. So returning to a situation where those plans can be operated is desirable.

Two issues speak against returning to the planned situation as quickly as possible. First, changing the operational plans is costly with respect to the same cost metrics mentioned above. The costs for communicating changes apply when changing already adapted plans back to the original plans. Second, as long as there is significant uncertainty in the system it may not pay off to invest a lot of effort in getting back to the planned operations. It only becomes relevant when the parameters describing the system have returned to the normal state or at least close enough to the normal state.

Time window	*Description*
Disruption time window	Expected duration of the disruption
Recovery time window	When the operations should be back to normal
Allocation time window (for a resource)	Time at which changes in resource assignment must be communicated

Table 4.1: Time windows in passenger railway disruption management.

The desire to return to the planned situation can be taken into account by increasing deviation costs in the recovery period after the situation has normalized.

Multicriteria decision making

Rescheduling the rolling stock during and after a disruption is a multicriteria decision making problem. On one hand we want to provide a service level that is as high as possible, but on the other hand we have to consider operational cost and deviation cost as well.

Not all criteria have the same priority at different stages in the process. Just after the occurrence of a disruption, in the shift from the undisrupted situation to a disrupted state, operational and deviation cost are of little importance. But after the situation has settled and there is less uncertainty in the development in the situation the operational cost and deviation cost gain relatively more importance.

Time windows

Dispatchers work with a number of time windows in the disruption management process, the most important of which are shown in Table 4.1. The *disruption time window* refers to the time at which the parameters of the system are expected to return to their normal state. It could for example be the time at which the repairs of a malfunctioning infrastructure component are expected to be over, or the time at which a broken down rolling stock unit is expected to be back in operation. Note that if multiple issues contribute to a complex disruption the dispatchers may work with more than one disruption time window. This way the dispatchers keep track of when the different involved parts of the system return to their original state. Also note that the disruption time window may very well change as the situation evolves due to the inherent uncertainty of disruptions.

The decision makers usually decide on a point in time where the normal operations should be restored. We call it the *recovery time window*. This time window is set sufficiently long after the disruption time window(s) to allow for a smooth convergence from the disrupted operations to the normal operations.

For the rolling stock we may distinguish two different recovery time windows; one for the time at which we want to have the same assignment of capacity to trains as in the normal plan, and one for the time at which we also want the amount of idle rolling stock units at the stations to be the same as in the normal plan. The latter is usually set at the night before the following weekday where all rolling stock is needed for the morning peak hours. This concept is known as *rebalancing the rolling stock*. In weekends less rolling stock is needed so the issue is less pressing, but still sufficient rolling stock is needed for the trains departing the following morning.

The recovery time window represents a trade-off between running an inefficient plan and the effort for returning to the normal solution. Clearly, a sooner recovery time window results in a potentially more costly effort for returning to the originally planned operations.

The third type of time window that we consider denotes the time period ahead in time in which a feasible solution must exist. For example, train crew would like to know their schedule in the disrupted situation some time ahead, and local planners would like to know about upcoming composition changes so they can prepare for them and allocate the necessary personnel to the shunting tasks. We call them *allocation time windows* and note that different types of railway resources have different allocation time windows; the decisions on rolling stock are more involved than decisions on the insertion of crew because of the adaptation of local plans. In either case, the following rule of thumb applies; the longer one waits before communicating a decision to the affected parties the greater the risk that it incurs delays in its execution.

The time windows described above are derived from practice and fit well into the model for disruptions in Section 4.4.2. The disruption time window, which denotes the time at which the parameters of the system are expected to be back to normal, can be incorporated in each scenario $\mathcal{S}_i$ since the time window is itself a parameter that describes the system. The recovery time window is decided by the decision makers based on the current scenario $\mathcal{S}_i$ and an estimate of the time needed to restore the original plans. The allocation time windows are more based on rules of thumb and internal company agreements. The allocation time windows may be less binding at certain stages of the process, especially right after the occurrence of the disruption as the state of the system is very uncertain.

Partial solutions

In current railway practice the decision makers utilize the concept of *partial solutions* to deal with the uncertainty of the development of a disruption and the fact that only very limited time is available for creating solutions. A partial solution is a solution that leaves part of the problem unsolved or that violates some of the less binding constraints. It is possible to work with partial solutions because only the most immediate decisions strictly need to be communicated and executed, i.e. those decisions that are within the allocation time window.

For the rolling stock a partial solution could be for example an assignment of rolling stock to trains that is feasible for the next two hours but leaves the allocation for later trains open. Or a partial solution could be a solution that solves only the issue of capacity allocation and leaves a number of off-balances to be solved through repositioning-trips during the night. Such examples widely occur as long as there is still a substantial amount of uncertainty in the system.

Multiple solutions

In practice it is desirable to generate multiple solutions to explore the trade-off between the multiple objectives, but in an environment where only limited time is available for creating solutions there is very little time for generating multiple solutions. This is also due to the limited integration of decision support tools in the disruption management process.

During a period of time with many rolling stock break downs due to abnormal weather conditions, we observed planners creating different rolling stock schedules by varying the number of units in reserve. This allowed the decision makers to consider a number of 'what-if' scenarios before choosing to implement one of the solutions.

4.5.2 Approach based on rolling horizon

We now combine the model for disruptions in Section 4.4.2 with the observations from practice in Section 4.5.1 to form a generic framework for disruption management of rolling stock in passenger railways.

The framework presented in this section does not include a specific algorithm for rescheduling the rolling stock. In principle, any model for rolling stock rescheduling can be used as the core of the solution approach. The framework only states how to decompose the problem into a series of subproblems without losing the overall structure of the problem.

The approach is based on the concept of a rolling horizon and works by considering only those trips that are within a certain time horizon from the time at which rescheduling takes place. Whenever new information on the state of the system becomes available, the rolling stock is rescheduled for a further time horizon. We note that Törnquist (2007) uses a similar approach for another railway application, namely traffic disturbance management.

To formalize the idea of rescheduling with a rolling horizon, let h be a *horizon parameter*, i.e. how far ahead we wish to take the current information into account. At time instant t_i when the information $\langle t_i, \mathcal{S}_i, \rangle$ becomes known, only the trips that are within the horizon are taken into account. More specifically, the trips in the timetable described by $\mathcal{S}_i$ that depart no later than $t_i + h$ are rescheduled and the remaining ones are ignored. If a new information update $\langle t_{i+1}, \mathcal{S}_{i+1} \rangle$ arrives, then the rolling stock is rescheduled accordingly for the time interval from t_{i+1} to $t_{i+1} + h$.

For a time interval without any information update, the current scenario is still the best available forecast on the development of the disruption. Still, a feasible plan only exists from the last update point until the end of the horizon. Therefore we introduce a new parameter p, called the *update parameter*, which denotes how often the circulation should be updated when the available information does not change. Then if no new information arrives for p minutes we create a dummy information update $\langle t_i + p, \mathcal{S}_i \rangle$ that uses the same scenario and we reschedule accordingly.

The idea of rolling horizon rescheduling is shown in Figure 4.2: A time-space diagram is shown for a timetable with trips between stations `A` to `D`. At time t_1 a disruption occurs which leads to some cancellations of trips. The rolling stock is rescheduled to serve the updated timetable until $t_1 + h$. Since no new information arrives for some time, rescheduling is performed again at time $t_1 + p$ with a horizon $t_1 + p + h$. However shortly later, at time t_2 new information becomes available and again the rolling stock is rescheduled until $t_2 + h$. Then no new information arrives for some time which means that rescheduling is performed again at time $t_2 + p$ and so on. This process continues until the end of the recovery time window where the state of the system is back to normal.

Note that if h and p are chosen sufficiently large it corresponds to trusting that the most recent information $\mathcal{S}_i$ received at time t_i is correct and we therefore reschedule the rolling stock until the end of the recovery time window.

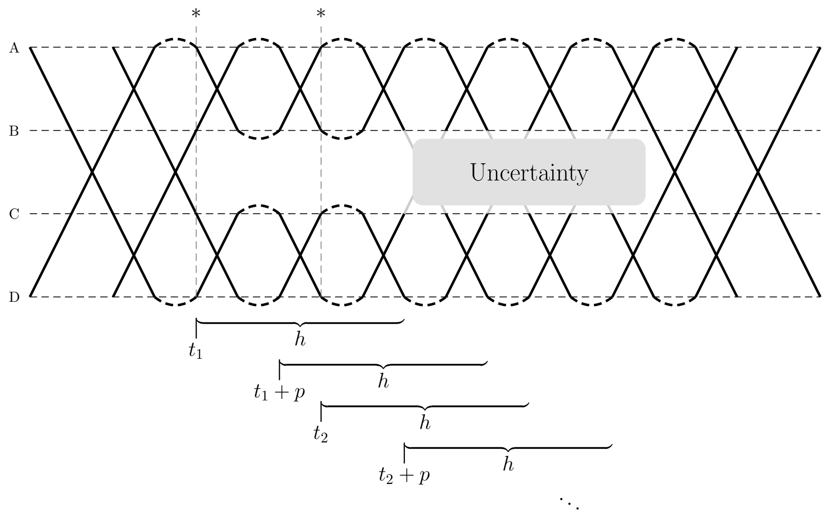


Figure 4.2: Rescheduling using a rolling horizon. When new information is available, the circulation is rescheduled h time steps ahead. $(*)$ marks the time instants where new information arrives.

Realistic values of h are around two to three hours as that allows dispatchers to react to the most immediate conflicts and still respect the allocation time windows. Parameter p should be somewhat smaller than h to allow a smooth roll of the horizon with sufficient overlap. It could be for example one hour if h is two hours.

4.6 Applying the framework to disruption management of rolling stock at NS

We now apply the framework developed in Section 4.5 to a concrete application in the disruption management process for rolling stock at NS. We assume that the development of a disruption is given as a series of scenarios (as described in Section 4.4.2) that describe the timetable after it is adapted to the current knowledge of the disruption.

Further we assume that an offline algorithm $\mathcal{A}$ is available which, given the input on the original plans, is able to reschedule the rolling stock according to the new state of the system. This $\mathcal{A}$ is thus able to solve an offline problem instance. The algorithm takes the service-, operational and deviation costs into account by applying penalties to the following elements.

- Canceling trips is penalized according to an estimate of the impact of the cancellation on the service.
- Assigning less capacity than needed according to the estimated number of passengers is penalized.
- Operational costs are taken into account by penalizing the number of kilometers of train ride performed by carriages.
- The deviation costs associated with changing local plans are accounted for by penalizing changes to the shunting plans.
- The deviation costs associated with off-balances at the end of the planning period are taken into account by penalizing the number of occurring off-balances.

It is then possible to use $\mathcal{A}$ to trade-off the involved objectives by applying the proper penalties to the potential assignment of the rolling stock in the timetable.

We use the offline algorithm $\mathcal{A}$ to reschedule the rolling stock in the time interval $[t_i; t_i + h]$ when the information $\langle t_i, \mathcal{S}_i, \rangle$ becomes known. We discuss next how to take the beforementioned objectives into account.

4.6.1 Accounting for the objectives when rescheduling with rolling horizon

The objectives for cancellation of trips can be taken into account by penalizing the assignment of no rolling stock to a trip. Other costs associated with the assignment of compositions to trips can be accounted for by applying the appropriate penalties to the assignment of rolling stock to trips. This is of course only possible for trips within the horizon but the rolling stock will be rescheduled for all trips as time progresses and the horizon rolls on. Similarly, the penalties for changes to the shunting plans can be taken into account when using the horizon.

The end-of-day off-balances are a different matter. In a generic iteration, the rolling stock circulation model is applied for the time interval $[t_i; t_i+h]$. Therefore the off-balances cannot be taken into account explicitly with a rolling horizon approach. We now propose a heuristic approach to account for off-balances.

Heuristic for accounting for off-balances with rolling horizon

Suppose the deviation costs associated with an off-balance of one unit of type m at station s at the end of the recovery window is $\beta_{s,m}$. But when the rolling stock is rescheduled at time t_i, we only reschedule the rolling stock until time $t_i + h$. This means that the end of the recovery time window may not be in the horizon yet. However, we can use $\mathcal{A}$ to measure the deviation of the rolling stock inventories at time $t_i + h$. The question is which target inventories are to apply. We propose to use the original (undisrupted) rolling stock assignment as a guideline.

One can explicitly compute the intermediate inventories $i^t_{s,m}$ according to the original plans: these numbers describe how many units of type m are located at station s at time instant t. Here $i^t_{s,m}$ is determined by counting the number of units of type m in the inventory of station s in the original plan and adding to these numbers the units of type m assigned to trips departing before t and arriving at s at time no earlier than t. We declare the values $i^{t_i+h}_{s,m}$ to be the target inventories when planning for the time interval $[t_i; t_i + h]$.

This guideline may be fairly inaccurate for early time instants t_i, and intuitively it becomes more and more precise as t_i approaches the end of the recovery time window. Therefore the importance of the off-balances in the objective function of $\mathcal{A}$ should increase with t_i. We achieve this by multiplying the penalty $\beta_{s,m}$ by a parameter $\varrho(t)$ where the parameter $\varrho(t)$ depends on the time of rescheduling t.

There are several ways to increase $\varrho(t)$ over time. We have restricted this study to a particular class of functions that increase the relative importance of the intermediate inventories from 0 to 1 over time. The function used is a scaled logistic curve. We introduce two parameters a and b that guide when the intermediate inventories start to be taken into account and when they are taken into account with full weight. The function is as follows

$$\varrho(t) = \begin{cases} 0, & t < a \\ \dfrac{f(t) - f(a)}{f(b) - f(a)}, & a \leq t < b \\ 1, & t \geq b \end{cases}$$

where

$$f(t) = \frac{1}{1 + e^{-\left(t - \frac{b-a}{2}\right)}}$$

The function is scaled so that it maps the interval $[a, b]$ onto the interval $[0, 1]$. The curve is sketched in Figure 4.3.

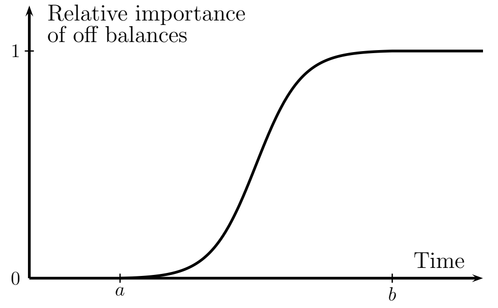


Figure 4.3: Increasing the relative importance of off-balances by a scaled logistic function.

4.7 Computational results

In this section we present computational results using the framework from Section 4.6. As the algorithm $\mathcal{A}$ we use the extended Composition Model described in Section 3.4. Recall that the Composition Model is a MIP-based model that reschedules the rolling stock to a modified timetable and allows the user to apply penalties to the options mentioned in Section 4.6.

4.7.1 Test instances

To analyze the relationship between parameter settings and solutions we use a realistic set of instances of NS. The instances are divided into four groups each concerning a particular disruption in the network. In this section we use one of the groups of instances to analyze the parameters while the results of the remaining results are presented in Appendix C.

The group of instances used in this section involve a single disruption. The difference between the instances is the time at which information becomes available and the accuracy of that information. The instances all have the same optimal offline solution. The disruption involves the so-called Noord-Oost lines, a system of interconnected lines with a closed rolling stock circulation (see Figure 4.4). The Noord-Oost lines form the most challenging cases for rolling stock scheduling at NS as they have a complicated structure such as underway splitting or combining of trains. We note that Fioole et al. (2006) report computation times of several hours

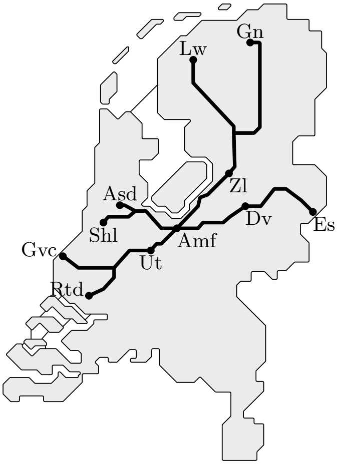


Figure 4.4: The Noord-Oost lines.

for tactical and operational scheduling problems on the Noord-Oost lines when the planning horizon is an entire day.

The disruption occurs at 12:00 between Utrecht (Ut) and Amersfoort (Amf). As a consequence, no trains can run between the two stations and all trains are turned in Utrecht and Amersfoort according to the handling scenario similar to the example for the 3000 line in Figure 4.1. The disruption turns out to last 2.5 hours. However, the actual length of the disruption is not known at the occurrence of the disruption but only an estimated length is available. An instance consists of the original timetable with the original circulation of rolling stock and a list of scenario updates that become available at certain times. Table 4.2 shows the updates that are available in each instance. The scenario $\mathcal{S}_{t_1-t_2}$ denotes the scenario in which the timetable for the Noord-Oost lines has had all trips canceled between Utrecht and Amersfoort in the time interval from t_1 to t_2. For example, in instance #6, the disruption is first estimated to last 1.5 hours, but at time 13:00 the estimated length of the disruption is changed to 2.5 hours in total.

Objective parameters

The overall objective is to balance the objectives; (*i*) changes to the shunting plans, (*ii*) off-balances, (*iii*) canceled trips, (*iv*) adequate seat capacity and (*v*) operational cost. Seat capacity is measured by the shortage of seats multiplied by the number of kilometers of train travel, and operational cost is measured by the number of

Instance	Timetable updates
#1	$\langle$12:00, $\mathcal{S}_{12:00-13:00}\rangle$, $\langle$12:30, $\mathcal{S}_{12:00-14:30}\rangle$
#2	$\langle$12:00, $\mathcal{S}_{12:00-13:00}\rangle$, $\langle$13:00, $\mathcal{S}_{12:00-13:30}\rangle$, $\langle$13:30, $\mathcal{S}_{12:00-14:30}\rangle$
#3	$\langle$12:00, $\mathcal{S}_{12:00-13:00}\rangle$, $\langle$13:00, $\mathcal{S}_{12:00-14:00}\rangle$, $\langle$13:30, $\mathcal{S}_{12:00-14:30}\rangle$
#4	$\langle$12:00, $\mathcal{S}_{12:00-13:00}\rangle$, $\langle$13:00, $\mathcal{S}_{12:00-14:00}\rangle$, $\langle$14:00, $\mathcal{S}_{12:00-14:30}\rangle$
#5	$\langle$12:00, $\mathcal{S}_{12:00-13:00}\rangle$, $\langle$13:00, $\mathcal{S}_{12:00-14:30}\rangle$
#6	$\langle$12:00, $\mathcal{S}_{12:00-13:30}\rangle$, $\langle$13:00, $\mathcal{S}_{12:00-14:30}\rangle$
#7	$\langle$12:00, $\mathcal{S}_{12:00-13:30}\rangle$, $\langle$13:00, $\mathcal{S}_{12:00-14:00}\rangle$, $\langle$13:30, $\mathcal{S}_{12:00-14:30}\rangle$
#8	$\langle$12:00, $\mathcal{S}_{12:00-13:30}\rangle$, $\langle$13:00, $\mathcal{S}_{12:00-14:00}\rangle$, $\langle$14:00, $\mathcal{S}_{12:00-14:30}\rangle$
#9	$\langle$12:00, $\mathcal{S}_{12:00-13:30}\rangle$, $\langle$13:30, $\mathcal{S}_{12:00-14:30}\rangle$
#10	$\langle$12:00, $\mathcal{S}_{12:00-14:00}\rangle$, $\langle$13:00, $\mathcal{S}_{12:00-14:30}\rangle$
#11	$\langle$12:00, $\mathcal{S}_{12:00-14:00}\rangle$, $\langle$13:30, $\mathcal{S}_{12:00-14:30}\rangle$
#12	$\langle$12:00, $\mathcal{S}_{12:00-14:00}\rangle$, $\langle$14:00, $\mathcal{S}_{12:00-14:30}\rangle$
#13	$\langle$12:00, $\mathcal{S}_{12:00-14:30}\rangle$

Table 4.2: Instances with a disruption on the Noord-Oost lines between Utrecht and Amersfoort.

kilometers each carriage travels. We here only analyze the first three objectives, but a study of the trade-off with the last two objectives is found in Appendix C.

Based on discussions with dispatchers we constructed a set of objective parameters. The parameters are used to explore the effect of the rolling horizon. The cost of canceling a trip is set to 10,000 which outweighs all other objective parameters. This reflects the practitioners' preference to use train cancellations only as a very last option. Minor changes to already planned shunting operations cost either 1, 2 or 5 depending on their nature, while new shunting operations cost 100 and off-balances cost 200. The same objective function is used for each test instance. With this set of weights the objective is focused on the local feasibility of the rolling stock schedules and off-balances while passenger capacity and operational costs are not explicitly taken into account.

4.7.2 Exploring the parameters

In this section we investigate different settings of the parameters of the rolling horizon approach. We then discuss their consequences and implications for practical use.

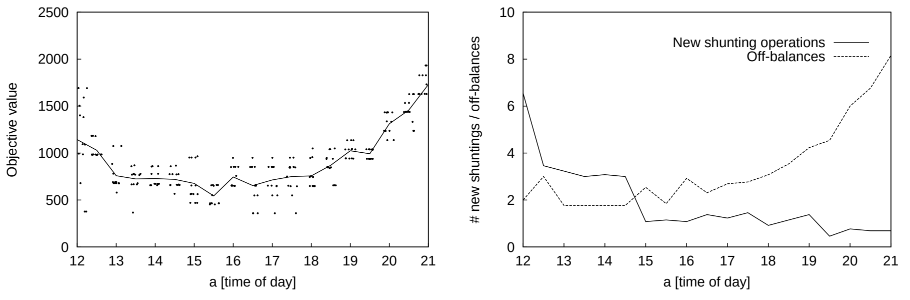


Figure 4.5: *Left*: Objective function of each instance as a function of parameter a. The line represents the average. *Right*: Average number of new shunting operations and off-balances.

Parameters for the intermediate inventories

As mentioned in Section 4.6, we handle off-balances heuristically by increasing the importance of balancing the intermediate inventories. The parameters a and b determine when the importance of the intermediate inventories starts to be taken into account and when it is taken into account with full weight, respectively.

In the first set of experiments, we set the horizon to $h = 3$ hours and the update parameter to $p = 1$ hour. We then applied the rolling horizon framework with a at different values between 12:00 and 21:00 with 30 minutes between values. The parameter b which controls when off-balances in the intermediate inventories are taken into account with full weight is set to 6 hours after a in all tests.

The left diagram of Figure 4.5 shows the objective values of all instances as a function of the parameter a. Each dot represents the outcome of one instance with the given parameters. The average outcome over all instances is plotted as a line in the diagram as well. It turned out that none of the obtained solutions requires cancellations of trips. In the right diagram of Figure 4.5 the average number of new shunting operations and off-balances are shown as they are the major contributors to the objective function.

We note that when the balancing is initiated rather late, more off-balances remain in the solutions. However, when balancing is initiated rather early, more new shunting operations are introduced without resolving more off-balances. The best balanced solutions seem to be found with a between 15:00 and 17:00.

Instance	New shunting operations	Weighted number of minor changes to shunting plans	off-balances	Objective
1	2	60	2	660
2	2	57	2	657
3	1	52	2	552
4	0	65	2	465
5	0	65	2	465
6	0	65	2	465
7	2	53	1	453
8	0	65	2	465
9	2	53	1	453
10	2	60	2	660
11	2	52	2	652
12	0	65	2	465
13	2	60	2	660
(∗)	0-2	22-26	11-13	–
offline	0	37	0	37

Table 4.3: Results of all instances with parameters $h = 3$ hours, $p = 1$ hour, $a =$ 18:30 and $b =$ 0:30. (∗) shows the results of all instances when rescheduling with the same horizon parameters except that off-balances are not taken into account.

The characteristics of all results with a specific choice of parameters are shown in Table 4.3. The parameters are $h = 3$ hours, $p = 1$ hour, $a = $ 15:30 and $b = $ 21:30. In all instances 1 or 2 off-balances occur and up to 2 new shunting operations are introduced. In the optimal offline solution it is possible to reschedule the rolling stock in such a way that no off-balances occur and no new shunting operations are introduced. However, when the balancing heuristic is not used and off-balances are not taken into account in the rescheduling, it results in 11-13 off-balances. The large reduction in off-balances comes at a cost of some minor changes to the existing shunting operations. In real-life railway practice a reduction in potential off-balances of this magnitude would be an excellent result. This suggests that the heuristic guidance of rolling stock to the target inventories works well in these instances.

We observe that parameter a represents a trade-off between off-balances and changes to the shunting plans. The tests have shown that a being around 15:00 to 17:00 yields balanced results for the Noord-Oost lines. However, the tests in Appendix C involving other lines have shown that the best choice of parameter a depends on the structure of the involved lines. For shorter lines the balancing can be initiated later since here the stations of the line are closer together.

Length of horizon

A set of experiments has been conducted to explore the relationship between the length of the horizon and the solution quality. The length of the horizon represents how far ahead in the current timetable trips are taken into account. Naturally, looking further ahead gives a larger solution space and thus potentially better solutions. However, this comes at a cost of computation time and there is no guarantee to obtain better solutions due to the online nature of the problem.

The instances have been tested with values for the horizon parameter h between 2 and 5 hours with 15 minutes between values. Based on earlier results we set parameter a to 17:00 and b to 23:00. The update parameter p is set to 1 hour. The objective values have been plotted as a function of horizon length in Figure 4.6.

There is a relationship between solution quality and horizon length; though longer horizons do not yield strictly better solutions, there is a tendency toward getting better results when using longer horizons. Also we barely experience any improvement for horizons longer than 3:30 which is in line with experience from practice.

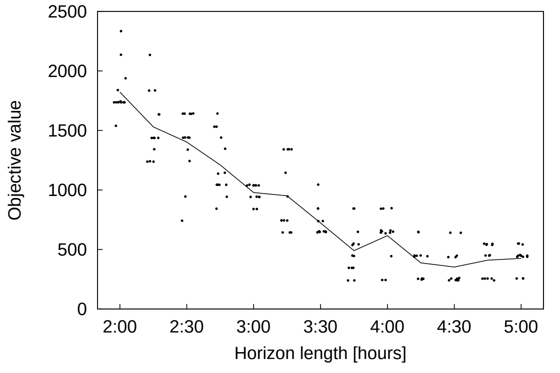


Figure 4.6: Objective function of each instance as a function of the parameter h. The line represents the average.

Update parameter

The parameter p determines how often the circulation is rescheduled when no updated information arrives. This way it controls the roll of the horizon. Revising the situation more often potentially yields better results as non-executed decisions can be revised when later trips come into the horizon. However, in a real-life setting the frequency of updates is limited by the time it takes to discuss possibilities with local planners and to communicate the changes to the involved parties.

We have tested the instances for values of p from 30 to 120 minutes with 15 minutes between values. The remaining parameters are fixed at $h = 4$ hours, $a =$ 17:00 and $b =$ 23:00. The objective values are plotted as a function of parameter p in Figure 4.7. It shows that updating more often potentially yields somewhat better solutions on average. The effect does not strictly increase with the value of p though.

4.7.3 Computation times

The tests have been conducted on a Pentium 4 2.8GHz desktop with 1 GB of RAM using CPLEX 10.1. As discussed earlier, the size of the solution space depends on the horizon length so the relationship between computation time and horizon length is explored. The computation times are presented in Figure 4.8 as a function of horizon length. Values of h from 2 hours to 5 hours were tested with 15 minutes between values. As expected, the computation time of the model can be somewhat unpredictable. This is clearly visible in the diagram where some computations take

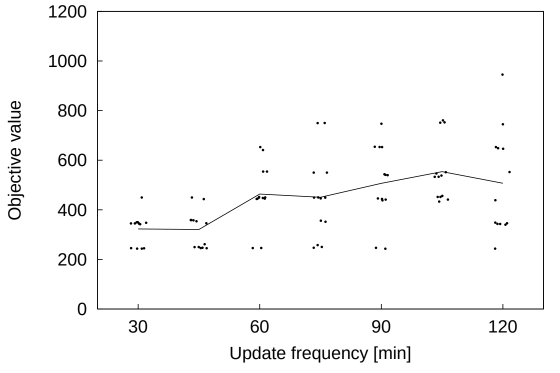


Figure 4.7: Objective function of each instance as a function of the update parameter p. The line represents the average.

up to one minute while other computations with the same parameters only take a few seconds. There is a significant correspondence between the longest experienced computation times and the length of the horizon. When using a horizon of more than 3.5 hours, unpredictably long computation times may occur. We note that when setting h large enough to reschedule the rest of the day, computation times of up to 10 minutes occurred, which is too long for the purpose of rescheduling in real time. Dispatchers would perform several runs of the model to explore different scenarios in case of a disruption. Thus fast computation times are imperative for usage in a disruption management system.

4.8 Conclusions

In this chapter we developed a framework for real-time disruption management of passenger railway rolling stock. We model disruptions in a pure wait-and-see fashion that utilizes the fact that decisions can be rescheduled rather close to execution. The main assumption in the framework is that timetable updates are explicitly given. The goal is to adjust the original rolling stock schedules for the updated timetables, taking various objectives on service, efficiency and deviations cost into account.

In order to deal with uncertainties (and in order to reduce the problem size), we proposed a rolling horizon framework as a solution approach. In this framework we only consider rolling stock decisions within a certain horizon of the time of reschedul-

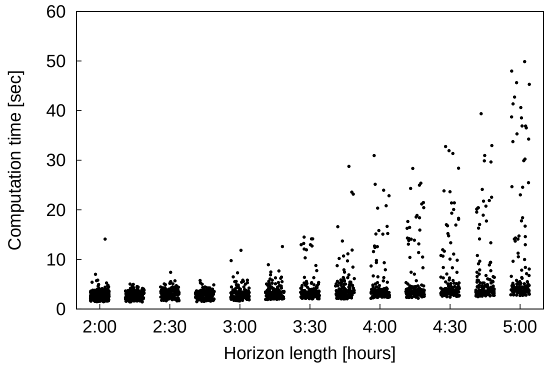


Figure 4.8: Computation time of each step of all instances as a function of parameter h.

ing. The schedules are then revised as the situation progresses and more accurate information becomes available.

The abstract framework needs to be adapted for the concrete specifications of real-life railway scheduling problems. We applied it to the rolling stock settings of NS by using the Composition Model described in Chapter 3. We used the model to penalize cancellations of trains, modifications of the shunting processes, carriage kilometers, seat shortage kilometers as well as the end-of-day rolling stock off-balances.

We note that the rolling horizon approach takes the off-balances into account in a heuristic way. Based on the undisrupted rolling stock circulation, we define target inventories during the recovery period, and we minimize the deviations from these targets. Arguably, the heuristic target becomes more accurate as the current horizon approaches the end of the recovery period. Therefore the objective coefficients of the off-balances increase as the horizon shifts ahead in time.

The performance of the rolling horizon approach and the heuristic for balancing the inventories at the end of the recovery depends on a number of parameters. These parameters include the length of the horizon and parameters that control the relative importance of the off-balances compared to the other criteria. Though the best values are instance dependent, the test results indicate that starting the balancing too early results in many changes to the shunting plans without much effect on the final balances. However, starting too late leaves little time for balancing the rolling stock and leads to many off-balances. The best time depends on the structure of the involved lines.

The length of the horizon offers a trade-off between solution quality and computation time, though longer horizons do not offer strictly better solutions, mainly because of the uncertainty related to the disruptions.

The tests show that the method can be used to reschedule the rolling stock during a disruption with minor effects for the shunting plans. At the same time the number of off-balances can be reduced by a few changes to the planned shunting operations. The obtained solutions may have a large relative gap to the optimal offline solution, but the values themselves are quite appealing in practice. This along with the short computation times indicates that the approach is a good candidate for the core of a decision support system for rolling stock rescheduling.

More tests need to be conducted to establish the best choice of the parameters for different classes of instances. Also, it would be useful to verify the model in a dynamic and stochastic environment either by simulation or in real life. Furthermore, it would be useful to investigate stochastic variants of the involved models.

We tested the disruption management approach on a number of disruption instances constructed from real-life rolling stock schedules in a simulated environment. However, testing the methods in a real-life setting would require a deep integration with the existing information systems to access the required data in a timely manner, but the development of the necessary interfaces with current systems is out of the scope of this thesis. Instead, we had the opportunity to test the methods in a real-life setting on a number of cases from the short-term planning department at NS. We discuss the experiences from that study in Chapter 6.

Chapter 5

Rolling Stock Rescheduling with Dynamic Passenger Flows

In the rolling stock rescheduling applications studied so far in this thesis, we have either treated passenger behavior as static whose influence on the system is unchanged in a disrupted situation, or we treated passenger behavior as a given input.

There are indeed situations where these assumptions on passengers are appropriate. When the changes to the system caused by a disruption are only light or moderate, we may assume that passenger demand is unchanged. Additionally, if suitable historical data is available we may assume that passenger behavior is given as input. In other situations, however, we may expect the flow of passengers to change significantly.

The changed demand for capacity may be alleviated by rescheduling the rolling stock, thus transferring capacity from low demand trains to high demand trains. Any attempt to reschedule the rolling stock implies a balance between the rescheduling effort and the corresponding service level. In this chapter we present a model for passenger flows during disruptions and we describe an iterative heuristic for optimizing the rolling stock to the disrupted passenger flows.

5.1 Models for passenger flows in various logistic settings

The dynamics of passenger flows have been studied in several logistic settings such as railways, airlines and transit networks, and with different focus such as line planning, revenue management and disruption management. We here list a number of references to studies in the literature, discuss their methods and assumptions, and compare them to the problem studied in this chapter.

Bratu and Barnhart (2006) study disruption management at a major airline company. They present MIP models that incorporate decisions on aircraft, crew and passenger recovery with the options of postponing or canceling flight legs. The objective is to simultaneously minimize operating costs, estimated passenger delay and disruption costs. The operator is assumed to have full control of the passenger flow in the network and can thus decide how passengers are matched with available capacity.

The assumptions on operator control of the assignment of passengers may be realistic in some railway settings. Consider a railway system where a seat reservation is required to board a train. In that case the operator does have control of the matching of passengers to trains. This contrasts the situation in the railway network of NS where the operator decides how to assign capacity to the timetable, but passengers decide how to utilize the available capacity.

Dumas and Soumis (2008) present a model for the passenger flow in airline networks given data on the demand between pairs of origin and destination and the temporal distribution of bookings. Furthermore, the authors assume knowledge on the spill of passengers between itineraries, i.e. if a booking for a certain set of flights is rejected they know the proportion of passengers who would attempt to book a certain alternative set of flights.

Dumas et al. (2009) use the above passenger flow model for the fleet assignment problem. This is the problem of assigning a fleet of aircraft to a set of flights while maximizing expected revenue. In earlier models for the fleet assignment problem fixed demands for itineraries are assumed given (see Hane et al. (1995) and Abara (1989)), and the problem can be expressed by a large MIP model. But since Dumas et al. (2009) use a complex non-linear passenger flow model, they separate the fleet assignment decisions and the computation of the revenue from the resulting passenger flow. The computations are performed in an iterative approach where in each iteration a fleet assignment is computed and the passenger flow model is used to evaluate the quality of the current assignment and to estimate the impact of changing a decision in the next iteration. The iterative approach presented by Dumas et al. (2009) is used as an inspiration for the solution approach presented in this chapter. We do, however, utilize another model for the passenger flow and the context is significantly different; revenue management in an airline versus disruption management in a railway network.

Passenger flow modeling in urban transit networks is a well studied field. Oppenheim (1995) gives a comprehensive overview and discusses the behavioral approach in which travelers are assumed to make travel choices which are "best" for them. In this approach passengers can be aggregated by temporal and spatial origin and destination. The simplifications obtained through this aggregation comes at the cost of information on the individual's traveling choices.

The delay management problem considers the decision of deliberately delaying the departure of a vehicle to allow passengers on delayed arriving vehicles to maintain their connections. The problem is related to the problem studied in this chapter as both problems share the goal of minimizing passenger delays. However, vehicle capacities are generally not considered in the models in literature. We refer to Ginkel and Schöbel (2007), Giovanni et al. (2007), and Schachtebeck (2010) for an overview of delay management.

Passenger flows are also considered in the strategic railway planning phase when the line planning is conducted. In this setting the passenger flow is considered aggregated by origin/destination pairs over the entire day. Line planning then amounts to determining service lines that balance operational cost and service according to the forecasted passenger flows. We give a number of references to literature on line planning in Section 2.7.

5.2 Framework for rolling stock rescheduling with dynamic passenger flows

In this chapter we investigate a certain class of disruptions where passenger behavior may influence the performance of the system considerably. The problem arises from two observations. The first observation is that in case of a blockage in the railway network passengers will attempt to get to their destinations by alternative traveling routes. The second observation is that passengers who change their routes lead to an altered demand for capacity on the trains serving the alternative routes. In fact, the demand for capacity on the alternative routes may be so large that the trains cannot accommodate all passengers leading to further delays and instability of the system.

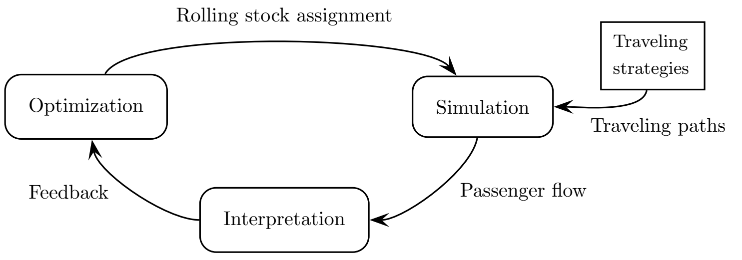


Figure 5.1: Iterative procedure for solving the rolling stock rescheduling problem with dynamic passenger flows.

The problem of rescheduling the rolling stock while accounting for the dynamics of passenger flows can be stated as the following generic optimization problem.

$$\min c(x) + d(y)$$

subject to

$$x \in \mathcal{X}(y)$$
$$y = f(x) \in \mathcal{Y}$$

Here $\mathcal{Y}$ is the set of feasible passenger flows. $\mathcal{X}$ is the set of rolling stock assignments to the timetable and $\mathcal{X}(y) \subseteq \mathcal{X}$ is the set of rolling stock assignments that are feasible for the passenger flow y. The function $f : \mathcal{X} \to \mathcal{Y}$ returns the anticipated passenger flow for a given assignment of rolling stock $x \in \mathcal{X}$. The cost function consists of two terms; the function $c : \mathcal{X} \to \mathbb{R}$ that gives the system related cost of a rolling stock assignment, and the function $d : \mathcal{Y} \to \mathbb{R}$ that gives the service related cost of a passenger flow.

We suggest an approach for solving the above model by iteratively rescheduling the rolling stock, simulating the passenger flows, and interpreting the flows to give an optimization direction for the next iteration. The approach is sketched in Figure 5.1. The framework is fully modular, which means it is possible to exchange any component in the procedure if we want to test a different set of underlying assumptions on the system.

Given a feasible rolling stock assignment, the anticipated passenger flow is computed by means of simulation. We introduce a model for passengers based on a multi commodity flow in an intuitive graph in Section 5.3. The model of passenger behavior follows a set of assumptions on passenger behavior. The assumptions relate to the traveling strategy applied by the passengers and to the interaction of the passengers when competing for capacity in the trains. We describe a simulation algorithm that implements this set of assumptions on passenger behavior in Section 5.4.

Based on the anticipated passenger flows computed in the simulation step, a feedback mechanism creates an optimization function for the next iteration. This function penalizes the assignment of rolling stock with too little capacity compared to the demand specified by the anticipated passenger flows. In Section 5.5 we discuss the design of such a feedback mechanism.

The optimization step involves assigning the rolling stock to the trains while taking the system related costs into account as well as the service related costs given by the feedback mechanism. In Section 5.6 we describe how to adapt the Composition Model introduced in Chapter 3 to this purpose.

Starting the iterative approach

We start the iterative procedure by performing the optimization step. Since no feedback from earlier iterations is available we only use the goals related to the system as objective in this first step. Based on the resulting assignment of capacity we perform the simulation step and compute the feedback for the next round of optimization.

Stopping criterion

We do not claim any guarantee on the performance of the heuristic iterative approach and there is therefore no natural stopping criterion. However, since the approach may be used in a time critical environment we could terminate the process when a time limit is reached or perform a fixed number of iterations. Alternatively, we could continue until no improvement is found during a number of iterations.

In our tests we limit ourselves to the simple stopping criterion of using a fixed number of iterations. In this way an acceptable number of solutions is visited within reasonable computation time.

5.3 Modeling the passenger flow

The model for passenger behavior relies on the following assumptions on individual passengers. A passenger enters the railway system at a specific time and wants to travel from an origin station to a destination station. If the passenger does not reach his destination within a certain time interval, we assume the passenger leaves the system and either gives up the travel or pursues another mode of transportation. We call the last time instant at which a passenger will accept to arrive the *deadline* of the passenger. Furthermore, each passenger has a *traveling strategy* for how to travel in the network. This strategy decides which trains he attempts to board given the available information on the state of the system.

The deadline represents the time at which a passenger leaves the system to either give up the intended travel or find a different mode of transportation. The deadline thereby implies that passengers are not willing to wait endlessly to get to their destination. When a passenger leaves the system due to the deadline it represents both a direct monetary cost for the operator as the passenger may be eligible for compensation, and an indirect monetary cost through the loss of goodwill.

Rather than modeling every single passenger in the system, we aggregate passengers with the same characteristics into *passenger groups*. Let $\mathcal{P}$ be the set of passenger groups. A passenger group $p \in \mathcal{P}$ has size n_p, and enters the system at time τ_p at origin station o_p. The group has destination station d_p and deadline $\tilde{t}_p$. The group uses traveling strategy S_p which is described in further detail in Section 5.3.2.

5.3.1 The passenger graph

The passengers travel in the time expanded graph defined by the timetable. We define the *passenger graph*, $G = (V, A)$, in the following way. Let $\mathcal{S}$ be the set of stations and let $\mathcal{T}$ be the set of trips. We add a node to the graph for each departure or arrival of a train at a station. Then a node denotes a station $s \in \mathcal{S}$ at a specific time τ. The set of nodes is thus defined as the set V:

$$V = \{(s, \tau) \mid \text{a train departs from or arrives at station } s \in \mathcal{S} \text{ at time } \tau\}$$

The set of arcs consists of two kinds of arcs, the *trip arcs* and the *time arcs*. A trip arc denotes a train traveling from one station to another whereas a time arc exists between every pair of consecutive nodes at the stations to denote waiting time at the station:

$$\begin{aligned}
A = &\{(u, v) \in V \times V \mid \text{a train departs at time } \tau \text{ from station } s \text{ where} \\
&\quad u = (s, \tau) \text{ and arrives at station } s' \text{ at time } \tau' \text{ where } v = (s', \tau')\} \\
&\cup \{(u, v) \in V \times V \mid u = (s, \tau), v = (s, \tau') \text{ where there does not exist a} \\
&\quad \text{node } w = (s, \tilde{\tau}) \text{ with } \tau < \tilde{\tau} < \tau'\}
\end{aligned}$$

The trains have a limited capacity depending on the rolling stock assigned to them. We define the capacity of a trip arc as the maximum number of passengers the train can accommodate according to the safety regulations. The capacity of a time arc is set to be infinite.

5.3.2 Traveling strategy

The traveling of a passenger constitutes a path in the passenger graph G. The traveling strategy S_p of a passenger group $p \in \mathcal{P}$ determines how the passengers in the group want to travel. It is a function $S_p : V \rightarrow \texttt{Paths}(G)$ from the nodes to the set $\texttt{Paths}(G)$ of directed paths in G. The path given by $S_p(v)$ denotes the preferred traveling path in the network when the group p is situated at node v. The definition of the traveling strategy S_p allows the passenger group to dynamically change its path in the event of a disruption i.e. a change in the structure of the graph. Note that S_p could be generalized to return a distribution of the passengers on a set of paths. But we limit this study to the special case where the traveling strategy implies that all passengers in the group prefer the same path.

The trip arcs in the graph have finite capacity so it may not be possible for all passengers in a group to travel with the same arc. This causes the passenger groups to interact when boarding the train as they compete for the scarce capacity. Once a passenger is in a train he occupies the capacity until he leaves the train and does therefore not have to compete with other passengers who attempt to board the train later. The assumptions on the interaction of the passengers are discussed in Section 5.4.1. Due to the interaction, the traveling of the passengers in a group p implies a network flow in the passenger graph rather than just a path. The flow originates from a source node $(o_p, \tau_p) \in V$ and flows to a set of nodes later in time. The combination of the flows of all the passenger groups constitutes a multi commodity flow in the passenger graph.

5.3.3 The quality of a passenger flow

The quality of the passenger flow is in this thesis measured by two criteria;

- Whether passengers are delayed compared to their expected arrival time.
- Whether passengers arrive at all at their destination station within their set deadline.

There are also other possible measures such as whether trains are overcrowded, but in this study we limit ourselves to the two above criteria i.e. delays and arrival within the deadline. More specifically, we define the *inconvenience* of a passenger as the number of minutes of delay experienced by the passenger plus a penalty for reaching the deadline. This penalty may then depend on the time of day, and whether the passenger can reach the destination by other means of transportation. In our experiments we penalize a passenger leaving the system by the number of minutes between the expected arrival time with the intended traveling path and the deadline. We define the overall service objective as the sum of the inconvenience experienced by each passenger.

5.4 Simulating the passenger flow

We designed a deterministic simulation algorithm to calculate the anticipated passenger flow. In this section we account for our assumptions on passenger behavior and then we present the simulation algorithm.

5.4.1 Assumptions

For the simulation of passenger behavior, we make assumptions on three key elements: (*i*) What information is available to the passengers? (*ii*) Which traveling strategy do passengers apply and how do they use the available information? (*iii*) How do passengers interact? We here discuss the assumptions we applied to these aspects and their implications on the practical applicability of our approach. We emphasize that the approach is modular which allows us to replace the strategy and interaction rules by any other set of assumptions.

Information

We assume that passengers know the timetable, i.e. they know the departure and arrival times of all trains. In addition, they know, at the occurrence of a disruption, which trains are canceled. Note that passengers do not know anything about cancellations before the disruption occurs. Furthermore, they do not know anything about the utilization of the capacity of the trains in the network either, and therefore they do not have any knowledge of whether they will be able to board trains on their traveling path.

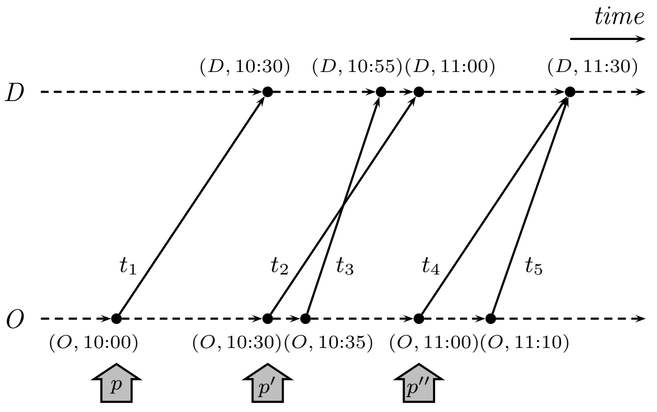


Figure 5.2: Part of a passenger graph with four trips from station O to station D. The gray arrows below the graph show when the three example passenger groups enter the system.

The assumption that all passengers know all departure and arrival times of all trains is a reasonable assumption since they are published in the timetable. However, when a disruption occurs there may be some uncertainty about the exact departure times of trains affected by the disruption. In this case it may be overly optimistic to assume knowledge of the exact event times. The assumption that passengers do not have any knowledge about the availability of capacity in the trains before attempting to board is realistic in a railway system where passengers cannot reserve capacity beforehand, and capacity is used on a first come first serve basis.

Strategy

We assume that the passengers in group p want to get to their destination as fast as possible. This means that they prefer traveling via a path in the passenger graph from the node (o_p, τ_p) to a node (d_p, τ) with smallest possible τ. If several such paths exist, the passengers prefer the path that involves the smallest number of transfers from one train to another. And again, if several such paths exist the passengers prefer the path with the earliest departure time.

We illustrate the assumptions by an example in Figure 5.2. The passenger graph is based on five trips $t_1, \ldots, t_5$ from station O to station D of which trips t_1, t_2 and t_4 have half an hour of traveling time whereas trips t_3 and t_5 take only 20 minutes. The nodes are named (s, τ) representing station s at time τ. Trip arcs are represented by solid arcs and time arcs are represented by dashed arcs.

A passenger group p traveling from origin $o_p = O$ to destination $d_p = D$ starting at time $\tau_p = 10{:}00$ will thus prefer traveling with trip arc t_1. A passenger group p' also traveling from origin $o_{p'} = O$ to destination $d_{p'} = D$ starting at time $\tau_{p'} = 10{:}30$ will prefer waiting on time arc $(O, 10{:}30)(O, 10{:}35)$ and then travel with trip arc t_3 rather than travel with t_2. On the other hand, a passenger group p'' also traveling from origin $o_{p''} = O$ to destination $d_{p''} = D$ starting at time $\tau_{p''} = 11{:}00$ will travel with trip arc t_4 rather than wait for t_5 as they have the same arrival time.

It is also part of the traveling strategy of a passenger group to react to the possibility that they may not reach their destination within the deadline. This is achieved by assuming that if a passenger group p is about to board a train that arrives after the deadline $\tilde{t}_p$, the passenger group will choose to leave the system rather than board the train. This behavior represents a passenger group choosing to either give up their intended travel or finding other means of transportation.

The assumed traveling strategy does not apply to all passengers in practice, some passengers may indeed choose for a later arrival time if it incurs fewer transfers, but in general passengers are expected to go for the earlier arrival time.

For the passengers leaving the system when they cannot reach their destination within their deadline, we recognize that passengers may have very different behaviors on when and where to leave the system. However, we would not expect passengers to stay in the system forever either. Modeling the behavior of leaving the system through a deadline thus enables us to track passengers who are severely affected by the state of the railway system and thereby represent a significant loss of goodwill. Note that the introduction of a deadline is not a limitation to the modeling power of the approach since, if the deadline is set sufficiently high, passengers will never leave the system at other stations than their destination.

Interaction

Passengers interact when attempting to board a train, in the sense that they compete for the limited capacity available in the train. When more passengers attempt to board a train than the available capacity allows for, only a portion of the passengers will enter the train. We assume that the number of passengers from each group who board a train is relative to the size of the group. Suppose passenger groups $p_1, \ldots, p_k$ attempt to board a trip arc a at node v with capacity c. Group p_i has size n_{p_i} and the combined size of the passenger groups is thus

$$n = \sum_{i=1}^{k} n_{p_i}$$

If $n > c$ then $f_{p_i} = c \cdot n_{p_i}/n$ passengers from group p_i board the train. The number of people from a group boarding a train may be fractional but that is not problematic since the train capacities are several hundreds and the contribution of fractional flows are therefore neglectable. Therefore we choose not to implement any tie breaking rule.

If not all passengers from group p_i are able to board a departing train then the boarded passengers as well as the rejected passengers will continue their journey according to their traveling strategy. The boarded passengers will constitute a flow of value f_{p_i} on arc a whereas the remaining $n_{p_i} - f_{p_i}$ passengers from the group will stay at node v and attempt to travel to d_{p_i} according to their strategy S_{p_i}. When some passengers from a passenger group p_i are rejected for boarding it is equivalent to splitting the group into two groups with the same characteristics except that one is situated at the arrival node of the arc with size f_{p_i} and the other at the departure node with size $n_{p_i} - f_{p_i}$.

Passengers who are already in the train and who wish to travel further do not participate in the above mentioned boarding procedure. This implies that when a train arrives at a station some of its capacity may already be in use.

Train capacities

In addition to the above assumptions on passenger behavior, we make one simplifying assumption on the train capacities. We assume that train capacities are only adapted at terminal stations and never underway. This implies that in any rolling stock assignment a train always has the same capacity on two consecutive trips. This further implies that there is enough capacity in a departing train to accommodate passengers who are already in the train and wish to travel further with the same train.

We note that this assumption is not a limitation of the approach since we can model trains with underway capacity adaptation by introducing parallel arcs in the passenger graph: Suppose a train uncouples a unit at an intermediate station, then parallel arcs can be used until that station. One set of arcs has capacity corresponding to the uncoupled unit and the other set of arcs has capacity corresponding to the remaining part of the train.

5.4.2 Simulation algorithm

We here describe a simulation algorithm that incorporates the above assumptions on traveling strategies and interactions between passengers. To the best of our knowledge there is no simulation algorithm in literature that allow us to incorporate exactly these assumptions.

Suppose the set of passenger groups is given by $\mathcal{P}$ and we are given a passenger graph $G = (V, A)$. For the purpose of simulating the passenger flow under the above assumptions we introduce a four-tuple (p, v, n, a) called a *container*. A container denotes n passengers from passenger group p that are positioned at node v and last traveled by trip arc a. The arc a thus indicates that the passengers are already in the train that performs a and do not participate in the boarding procedure if they wish to continue with the same train. If the passengers have not yet traveled with any arc, the arc a in the container is set to a dummy value ϕ.

Algorithm 5.1 generates the passenger flow under the given assumptions. In short, the algorithm works by performing a time sweep over all trip arcs in the passenger graph and moving the containers through the graph according to the assumptions on traveling strategies and interaction. The algorithm returns a function $f : A \times \mathcal{P} \to \mathbb{R}$ that maps the arcs and passenger groups to the size of the flow of the particular group on a certain arc.

In the preprocessing in lines 1 – 3 a set S of containers is initialized to hold each passenger group at its origin station at the time it enters the network. The set of arcs is then ordered by departure time. If several arcs depart at the same time their ordering is arbitrary. The function f is set to 0 for all pairs $(a, p) \in A \times \mathcal{P}$.

In line 4 we iterate over all arcs to apply the boarding procedure at each departure. In line 5 we denote the departure and arrival nodes of the arc a by (s, τ) and (r, σ), and in line 6 we denote the predecessor arc of a by $a_{\texttt{pred}}$. The predecessor arc is serviced by the train immediately before the trip arc a. This notion is necessary to determine which passengers are already in the train as they do not have to participate in the boarding procedure. Then the subset $S' \subseteq S$ is extracted holding the containers $g = (p, (s, \tau), n, b)$ that are situated at the departure node (s, τ) and want to travel with the trip in question. Note that this is determined by the traveling strategy S_p of the passengers in group p. Further, S'' is the subset of S' that is already in the train before it departs.

Next, in line 9, the free capacity of the arc a is calculated by subtracting from the total capacity $\texttt{cap}(a)$ the capacity taken up by passengers already in the train. Passengers already in the train are identified by having last traveled by arc $a_{\texttt{pred}}$. Note that the remaining capacity c is always non-negative as the capacity of the train is unchanged according to our assumptions, i.e. $\texttt{cap}(a_{\texttt{pred}}) = \texttt{cap}(a)$.

Input: Passenger graph $G = (V, A)$, set of passenger groups $\mathcal{P}$
Output: Flow function $f : A \times \mathcal{P} \to \mathbb{R}$

1 $S := \{(p, (o_p, \tau_p), n_p, \phi) \mid p \in \mathcal{P}\}$
2 $L :=$ An ordering of A by departure time
3 Let $f(a, p) := 0$ for all $a \in A$, $p \in P$
4 **foreach** $a \in L$ **do**
5   Let (s, τ) and (r, σ) be the departure and arrival nodes of a respectively
6   Let $a_{\mathtt{pred}}$ be the predecessor of a
7     $S' = \{g = (p, (s, \tau), n, b) \in S \mid g$ wants to travel with $a\}$
8     $S'' = \{g = (p, (s, \tau), n, b) \in S' \mid g$ is already in the train (i.e. $b = a_{\mathtt{pred}}$)$\}$
9     $c := \mathtt{cap}(a) - \sum_{g \in S''} n(g)$
10   **foreach** $g \in S'$ **do**
11   **if** $b = a_{\mathtt{pred}}$ **then**
12             $u := n(g)$
13   **else**
14             $u := \min\{n(g), c \cdot n(g) / \sum_{g' \in S' \setminus S''} n(g')\}$
15         $S := S \setminus \{g\}$
16   **if** $u < n$ **then**
17   Let (s, τ') be the next departure node at station s
18             $S := S \cup \{(p(g), (s, \tau'), n - u, \phi)\}$
19   **if** $r \neq d_p$ **then**
20             $S := S \cup \{(p(g), (r, \sigma), u, a)\}$
21         $f(a, p) := f(a, p) + u$
22 **return** f

Algorithm 5.1: Simulation algorithm for the passenger flow.

In the loop starting from line 10 each relevant container $g \in S'$ is handled. First the number of passengers u who obtain capacity on the train is calculated. If the passengers are not in the train (line 11) they participate in the boarding procedure (line 14) or else they all fit in the train (line 12).

The container is then removed from the set S, and replaced by new containers as follows. If not all passengers are assigned to the arc (i.e. $u < n$) then the $n - u$ rejected passengers from the group remain at the station at node (s, τ') defined as the departure node of the next arc departing from this station (lines 16 – 18). The number of passengers u that were accepted on a will reappear at the arrival node (r, σ) unless it is the destination station of the passengers (lines 19 – 20). Finally the flow function is updated in line 21.

Example

We return to the example in Figure 5.2. Let each of the trains assigned to the five trips have a capacity of 100, and let $\mathcal{P} = \{p_1, p_2\}$ consist of two passenger groups with origin $o_{p_1} = o_{p_2} = O$, destination $d_{p_1} = d_{p_2} = D$, size $n_{p_1} = n_{p_2} = 100$, and deadline sufficiently large. The groups enter the system at time $\tau_{p_1} = 10{:}30$ and $\tau_{p_2} = 10{:}35$, respectively. According to the traveling strategies S_{p_1} and S_{p_2} both groups prefer traveling by the path in the passenger graph that uses trip t_3. In the simulation the containers are initialized to

$$S = \{(p_1, (O, 10{:}30), 100, \phi), (p_2, (O, 10{:}35), 100, \phi)\}$$

and after processing the outgoing arcs of node $(O, 10{:}30)$ the set contains

$$S = \{(p_1, (O, 10{:}35), 100, \phi), (p_2, (O, 10{:}35), 100, \phi)\}.$$

When processing the arc of trip t_3 both groups will attempt to board the train and according to the assumptions on the boarding procedure 50 passengers from each group will be assigned to the arc thus setting $f(t_3, p_1) = f(t_3, p_2) = 50$. New containers are inserted at the next relevant node thus resulting in the following set of containers:

$$S = \{(p_1, (O, 11{:}00), 50, \phi), (p_2, (O, 11{:}00), 50, \phi)\}.$$

When processing the arc of trip t_4 the remaining passengers will be able to board, thus setting $f(t_4, p_1) = f(t_4, p_2) = 50$.

We can calculate the delays of the involved passenger groups using the flow f. In the above example 50 passengers from each of the groups p_1 and p_2 suffered a delay of 35 minutes compared to their initially intended journeys which results in a total delay of $2 \cdot 50 \cdot 35 = 3500$ delay minutes. We observe that the scarce capacity in the trains combined with the greedy traveling strategy of passengers may lead to delays. In our experiments we choose to penalize delay minutes uniformly although one may argue that longer delays are worse than several small delays.

5.4.3 Implementation issues

In addition to the generic assumptions on passenger behavior implemented in the simulation algorithm 5.1 there are several other issues to be taken into account. Those are primarily related to assumptions on special cases and implementation issues.

The traveling strategy of the passengers is implemented as a shortest path algorithm in the passenger graph. The shortest path search can be performed in linear time since the passenger graph is acyclic. In the implementation we store the desired traveling path with the container to avoid recalculating the path several times. Only when passengers are rejected in the boarding procedure we need to recompute their path.

The occurrence of a disruption is incorporated in the simulation algorithm by changing the passenger graph at the appropriate point in the simulation. All containers in the set S then have their preferred path recomputed.

As the simulation algorithm is stated, it returns a function $f : A \times \mathcal{P} \to \mathbb{R}$ denoting the number of passengers from each group traveling by each arc. However, it is necessary for the feedback mechanism described in Section 5.5 to have the path decomposition of the flows rather than the flows represented by f. One can derive a greedy path decomposition from f in time $O(|\mathcal{P}| \cdot |V| \cdot |A|)$, but we need the path decomposition that represents the actual paths traveled by passengers. We can create it on the fly by storing the path history with each container in the algorithm instead.

We note for the computational complexity of the simulation algorithm, that up to two containers are added to S whenever one container is removed. This implies that in the worst case each passenger group is split a number of times that is exponential in the number of trips. It is trivial to construct instances that demonstrate asymptotic worst case behavior, but as we observe later, the computational behavior of the simulation on realistic instances is not a bottleneck in the process.

We note that we could potentially obtain a speedup of the algorithm by merging containers that represent the same passenger group if they meet at a common node. This would even prevent the worst case exponential complexity. However, we would then loose the history of paths traveled thus far in the simulation and we would loose information on the traveling patterns of groups that were split.

5.5 Feedback

The purpose of the feedback mechanism is to interpret the passenger flow returned by the simulation algorithm, and provide an optimization direction that is likely to improve the solution in the next iteration. This is performed by estimating the effect of the train capacities on the total passenger inconvenience.

More formally, the feedback mechanism is a function $F : \mathcal{T} \times \mathbb{Z}_+ \to \mathbb{R}$ that maps the trips $\mathcal{T}$ in the timetable and the non-negative integers to values in $\mathbb{R}$. For trip $t \in \mathcal{T}$ and $c \in \mathbb{Z}_+$ the value $F(t, c)$ denotes the penalty to the potential decision of assigning rolling stock with capacity c to trip t.

Consider a passenger flow $f : A \times \mathcal{P} \to \mathbb{R}$ returned by the simulation algorithm in Section 5.4. Then consider a trip $t \in \mathcal{T}$ and let $a \in A$ be the trip arc representing trip t. We achieve this by considering how many passengers wanted to travel with t in the simulation and estimating the delay of any passengers who were unable to board the train at its departure.

From the simulation algorithm we know which passenger groups attempted to travel with arc a. Let $p_1, \dots, p_k$ be those groups. Let $u_i = f(a, p_i)$ be the number of passengers from group p_i that traveled with arc a and let r_i be the number of passengers from group p_i that attempted to board arc a but were rejected.

Consider a passenger group p_i. The passengers of this group follow paths from the origin node o_{p_i} according to the path decomposition from the simulation algorithm. Suppose s_i paths contain the arc a, and t_i paths result from the rejection at the boarding procedure at arc a. Then let $Q_1, \dots, Q_{s_i}$ be the paths that contain the arc a and let $W_1, \dots, W_{t_i}$ be the paths that result from the rejection at a. For a path P let $n(P)$ be the number of passengers that traveled along P and let $v(P)$ be the inconvenience per passenger traveling along P for the group p_i. Then the term

$$\texttt{TRAVEL}(p_i, a) = \frac{\sum_{j=1}^{s_i} (n(Q_j) \cdot v(Q_j))}{\sum_{j=1}^{s_i} n(Q_j)}$$

denotes the average inconvenience per passenger from group p_i traveling on a. Similarly, the term

$$\texttt{REJECT}(p_i, a) = \frac{\sum_{j=1}^{t_i} (n(W_j) \cdot v(W_j))}{\sum_{j=1}^{t_i} n(W_j)}$$

denotes the average inconvenience per passenger from group p_i rejected upon boarding a. Using the terms $\texttt{TRAVEL}(p_i, a)$ and $\texttt{REJECT}(p_i, a)$ we estimate the contribution

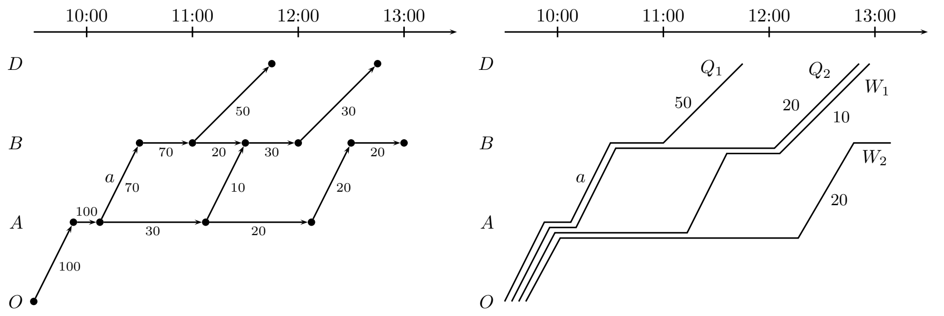


Figure 5.3: *Left*: The passenger flow on arcs for passenger group p_i. *Right*: The path decomposition.

per passenger to the total inconvenience by

$$\texttt{AVG\_COST}(a) = \frac{\sum_{i=1}^{k} r_i \cdot (\texttt{REJECT}(p_i, a) - \texttt{TRAVEL}(p_i, a))}{\sum_{i=1}^{k} r_i}$$

where $\texttt{AVG\_COST}(a)$ is the marginal contribution to the cost of the passenger flow by decreasing the capacity of arc a by a small value ϵ. We set the function in the feedback mechanism to the following value

$$F(t, c) = \max \left\{ 0 \, , \, \sum_{i=1}^{k} (u_i + r_i) - c \right\} \cdot \texttt{AVG\_COST}(a)$$

which denotes the expected cost of assigning capacity c to trip t (where a is the trip arc corresponding to t). Note that if the capacity is sufficient to hold all passengers (i.e. $c \geq \sum_{i=1}^{k} (u_i + r_i)$) then it incurs no cost to assign the capacity.

Example

Consider the example passenger flow of a group p in the left diagram of Figure 5.3. The group consists of 100 passengers traveling from origin O to destination D, starting at 9:00, and with a deadline of 13:15. Upon boarding arc $a = (A, 10{:}10)(B, 10{:}30)$ the passenger group is split where 70 passengers are able to board and the remaining 30 passengers are rejected. The group is further split at some other arcs resulting in the paths shown in the right diagram in Figure 5.3. Paths Q_1 and Q_2 follow from the boarding of a whereas paths W_1 and W_2 follow from the rejection at a.

Assume the passengers in group p had expected arrival time 11:45. Then we derive the following path sizes and values from the figure: $n(Q_1) = 50$, $v(Q_1) = 0$,

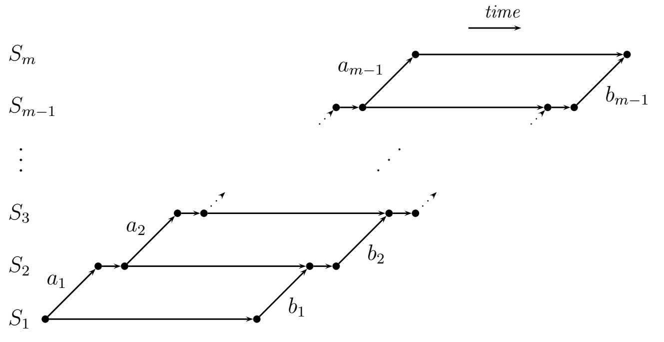


Figure 5.4: Example with passenger graph that needs $m-1$ iterations to avoid delay.

$n(Q_2) = 20$, $v(Q_2) = 60$, $n(W_1) = 10$, $v(W_1) = 60$, $n(W_2) = 20$ and $v(W_2) = 90$. The average contribution per passenger to the total inconvenience for passengers from group p respectively traveling with or rejected at a is calculated as follows.

$$\texttt{TRAVEL}(p, a) = \frac{\sum_{j=1}^{2} (n(Q_j) \cdot v(Q_j))}{\sum_{j=1}^{2} n(Q_j)} = \frac{50 \cdot 0 + 20 \cdot 60}{70} = 17.1$$

$$\texttt{REJECT}(p, a) = \frac{\sum_{j=1}^{2} (n(W_j) \cdot v(W_j))}{\sum_{j=1}^{2} n(W_j)} = \frac{10 \cdot 60 + 20 \cdot 90}{30} = 80$$

The estimated contribution to the inconvenience per passenger rejected at arc a is thus 62.9.

Observations and remarks

We note that the feedback mechanism may encourage the insertion of more capacity on an arc a resulting in more passengers traveling on that arc. However, these passengers may turn out to be delayed by the lack of capacity on other arcs instead. It may thus take several iterations before reaching an assignment of rolling stock that avoids some of the delays. Consider for example the passenger graph in Figure 5.4 with m stations $S_1, \ldots, S_m$. Two trip arcs a_i and b_i connect station S_i and S_{i+1} where b_i is later than a_i. Suppose a passenger group has origin S_1 and destination S_m and has some passengers rejected upon boarding at arc a_1. Those passengers will then travel with trip arcs $b_1, \ldots, b_{m-1}$.

The feedback mechanism will then penalize capacity shortage on arc a_1. Assigning more capacity to a_1 will then cause the passengers to be rejected at arc a_2 instead, leading to the same amount of delay. The feedback mechanism will then additionally penalize capacity shortage on arc a_2 in the next iteration and so on. In this example it will take $m-1$ iterations before the delay is avoided.

We also note that the feedback mechanism described in this section is a natural candidate for providing a new optimization direction based on the current solution. But there exist many other schemes for this task that could make sense.

5.6 Optimization

For rescheduling the rolling stock we use the Composition Model described in Section 3.4. Recall that the model contains variables $X_{t,p} \in \{0,1\}$ that denote whether composition $p \in \eta(t)$ is used for trip $t \in \mathcal{T}$ where $\eta(t)$ is the set of allowed compositions for t. The capacity of the rolling stock in composition p is $\texttt{cap}(p)$.

The objective of the model consists of two parts; the system related costs and the service related costs. At first we could define the service related costs in the ith iteration by the function $g_i(X)$ using the feedback mechanism $F : \mathcal{T} \times \mathbb{Z}_+ \to \mathbb{R}$.

$$g_i(X) = \sum_{t\in\mathcal{T}} \sum_{p\in\eta(t)} F(t, \texttt{cap}(p)) \cdot X_{t,p}$$

Then the objective function $h_i(X, Z, I)$ in the ith iteration can be defined as follows.

$$h_i(X, Z, I) = c(X, Z, I) + g_i(X)$$

where $c(X, Z, I)$ are the system related costs and we assume that the feedback is initially $g_0(X) = 0$.

However, we experienced that this approach could lead to cyclical behavior as the feedback from earlier iterations is ignored. We therefore modify the part of the objective derived from the feedback to

$$g_i(X) = (1-\alpha)g_{i-1}(X) + \alpha \sum_{t\in\mathcal{T}} \sum_{p\in\eta(t)} F(t, \texttt{cap}(p)) \cdot X_{t,p}$$

where $0 \le \alpha \le 1$ is a parameter that weighs the latest feedback against feedback from earlier iterations. This way the feedback from a certain iteration gradually becomes less relevant. Note that $\alpha = 1$ describes the special case where feedback from earlier iterations is ignored. The similar introduction of such a parameter α in the approach by Dumas and Soumis (2008) was shown to be crucial for the performance of their solution procedure.

We mention that neither the solution nor the g_i function necessarily converges with the applied feedback mechanism and the definition of α.

5.7 Lower bounds

To assess the quality of our solutions we investigate methods for constructing lower bounds on the solution value. This allows us to analyze how much of the delays is caused by the changes in the disrupted timetable and how much is caused by the shortage of capacity.

In this section we describe two methods based on relaxing some of the assumptions on the problem. The first approach is rather simple, we relax the constraints that cause the passengers to interact – namely the limitation on the capacity of rolling stock. In the second approach we relax some assumptions on the passengers' traveling strategies; we assume that capacity is not utilized by passengers in a greedy manner, but is rather allocated by the operator such that total inconvenience in the system is minimized.

5.7.1 Lower bound based on unlimited capacity

An intuitive way to relax the underlying assumptions of the system on passenger interaction is to assume that all trains have unlimited capacity. This way no passenger is ever rejected in the boarding procedure and every passenger can follow the best path in the passenger graph according to their traveling strategy. This implies that all delays experienced by passengers are caused by changes to the timetable such as canceled trips rather than by limited capacity.

This relaxation corresponds to the situation where unlimited amounts of rolling stock are available, an arbitrary amount of rolling stock may be assigned to each train, and there are no limitations on shunting possibilities at any station. It results in a lower bound on the service related cost i.e. the total amount of delay experienced by passengers in the system. We call this lower bound the *infinite capacity lower bound* (IC-LB).

Calculating the infinite capacity lower bound simply amounts to simulating the passenger groups $\mathcal{P}$ in a passenger graph $G = (V, A)$ where all arcs $a \in A$ have capacity $\texttt{cap}(a) = \infty$.

5.7.2 Lower bound based on centralized passenger flow

As described earlier, the passenger flow constitutes a multi commodity flow in the passenger graph. The quality of the flow is measured by the total inconvenience experienced by the passengers, i.e. the amount of delay and the number of passengers who do not reach their destination before their deadline. Since the passengers themselves choose their routes in the network the total flow may be suboptimal with respect to this quality measure. Suppose we relax the assumptions on the route choice to let the operator decide which passengers travel with which trains.

For passenger group $p \in \mathcal{P}$ we denote by π_p the set of paths the passengers in the group can travel with. The paths originate from the node (o_p, τ_p) in the passenger graph $G = (V, A)$ and can lead to any station in the network – not only to the destination station. This includes an empty path that represents staying at the origin station.

Define for each passenger group $p \in \mathcal{P}$ and possible traveling path $q \in \pi_p$ a variable $Y_{p,q} \in \mathbb{R}_+$. The variable states the number of passengers from group p that travel with path q. Let $c_{p,q}$ denote the inconvenience by one passenger from group p traveling with path q. Then the passenger flow with operator control may be expressed as the following linear program.

$$\min \sum_{p \in \mathcal{P}} \sum_{\substack{q \in \pi_p: \\ a \in q}} c_{p,q} Y_{p,q} \tag{5.1}$$

subject to

$$\sum_{q \in \pi_p} Y_{p,q} = n_p \qquad \forall p \in \mathcal{P} \tag{5.2}$$

$$\sum_{p \in \mathcal{P}} \sum_{\substack{q \in \pi_p: \\ a \in q}} Y_{p,q} \leq \mathtt{cap}(a) \qquad \forall a \in A \tag{5.3}$$

$$Y_{p,q} \in \mathbb{R}_+ \qquad \forall p \in \mathcal{P}, q \in \pi_p \tag{5.4}$$

The objective function (5.1) consists of the sum over the paths of the number of passengers traveling with a possible path times the contribution per passenger of the path to the quality of the flow.

Constraints (5.2) state that all passengers in a group must travel by one of the possible paths. The capacity constraints (5.3) denote that the number of passengers traveling on an arc is limited by the arc capacity. Finally constraints (5.4) state the domains of the variables.

The program (5.1) – (5.4) models the operator controlled passenger flow problem as a continuous minimum cost multi commodity flow problem. However, the model assumes that capacities on arcs are given as input (right hand side of constraints (5.3)). The arc capacities are decided through the rolling stock rescheduling process and are themselves subject to constraints on rolling stock availability and shunting possibilities. In order to incorporate the aspect of rolling stock rescheduling, we extend the model to (5.5) – (5.9) below. Here, the variables $X_{t,p} \in \{0,1\}$ model the assignment of composition p to trip t as in the Composition Model in Chapter 3.

$$\min \sum_{p \in \mathcal{P}} \sum_{\substack{q \in \pi_p: \\ a \in q}} c_{p,q} Y_{p,q} \tag{5.5}$$

subject to

$$\sum_{q \in \pi_p} Y_{p,q} = n_p \qquad \forall p \in \mathcal{P} \tag{5.6}$$

$$\sum_{p \in \mathcal{P}} \sum_{\substack{q \in \pi_p: \\ a \in q}} Y_{p,q} \leq \sum_{p \in \eta(t)} \texttt{cap}(p)\, X_{t,p} \qquad \forall a \in A, t \text{ is trip of } a \tag{5.7}$$

$$X \in \mathcal{X} \tag{5.8}$$

$$Y_{p,q} \in \mathbb{R}_+ \qquad \forall p \in \mathcal{P}, q \in \pi_p \tag{5.9}$$

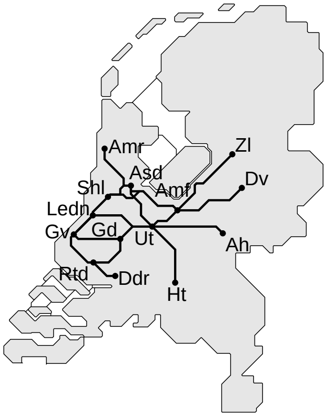
In constraints (5.7) the right hand side describes that the capacity of an arc is determined by the rolling stock composition assigned to the trip represented by the arc. The constraint (5.8) states that the vector X of composition variables must belong to the set $\mathcal{X}$ of feasible assignments of rolling stock to trips. In the railway system considered in this thesis, it is equivalent to saying that it must be a feasible solution to the Composition Model. We call this the *strong operator controlled lower bound* (SOC-LB). However, we did not solve this model. But in the discussion of future research in Section 5.9 we propose a method for solving SOC-LB using a branch-and-cut approach.

A somewhat weaker lower bound that does not require solving the complex model (5.5) – (5.9) can be constructed by relaxing constraints (5.8). Rather than requiring that all variables $X_{t,p}$ correspond to a feasible solution to the rolling stock rescheduling problem, we require for each trip $t \in \mathcal{T}$ to be assigned a composition $p \in \eta(t)$ with the largest possible capacity that is feasible in some solution $X \in \mathcal{X}$.

This enables us to decide the capacity in a preprocessing step and fix the variables $X_{t,p}$. With the capacities fixed the model is equivalent to the multi commodity flow model (5.1) – (5.4). We call this the *weak operator controlled lower bound* (WOC-LB).

We solve WOC-LB in two steps. In the first step, the Composition Model is solved once for each trip $t \in \mathcal{T}$ with the simple objective

$$\texttt{cap}_{\texttt{max}}(t) = \max \sum_{p \in \eta(t)} \texttt{cap}(p) X_{t,p}$$



Line	Stations	Frequency
500	Gv Gd Ut Amf Zl	hourly
700	Shl Amf Zl	hourly
800	Amr Zd Asd Ut Ht	half hourly
1500	Asd Amf Dv	half hourly
1600	Shl Amf Dv	hourly
1700	Gv Gd Ut Amf Dv	hourly
1900	Gv Rtd Ddr	half hourly
2000	Gv Gd Ut Ah	half hourly
2100	Asd Shl Ledn Gv Rtd Ddr	half hourly
2600	Asd Shl Ledn Gv	half hourly
2800	Rtd Gd Ut Amf	half hourly
3000	Amr Asd Ut Ah	half hourly
3500	Shl Ut Ht	half hourly
8800	Ledn Ut	half hourly
20500	Rtd Gd Ut	hourly
21700	Rtd Gd Ut	hourly

Figure 5.5: The network considered in the test instances.

thus computing the maximum possible capacity $\texttt{cap}_{\texttt{max}}(t)$ on each trip t in any feasible assignment of the rolling stock. In the second step, we solve the model (5.1) – (5.4) with the arc capacities in constraint (5.3) set to $\texttt{cap}(a) = \texttt{cap}_{\texttt{max}}(t)$ where a is the corresponding trip arc of t in the passenger graph. The multi commodity flow problem is solved using a text-book column generation procedure (see Ahuja et al. (1993)).

5.8 Computational tests

In this section we perform computational tests based on our approach. In Section 5.8.1 we describe how we generate test instances with different characteristics, and in Section 5.8.2 we report and discuss the results.

5.8.1 Instances

For testing our approach we constructed a number of instances from realistic data based on the Intercity network of NS. The instances involve the heavily utilized core part of the network connecting the 14 stations shown in Figure 5.5. This part of the network is serviced by the 16 Intercity lines listed in the figure. The lines call at the given stations and operate with the specified frequencies. On most routes there are at least four trains per hour between neighboring stations.

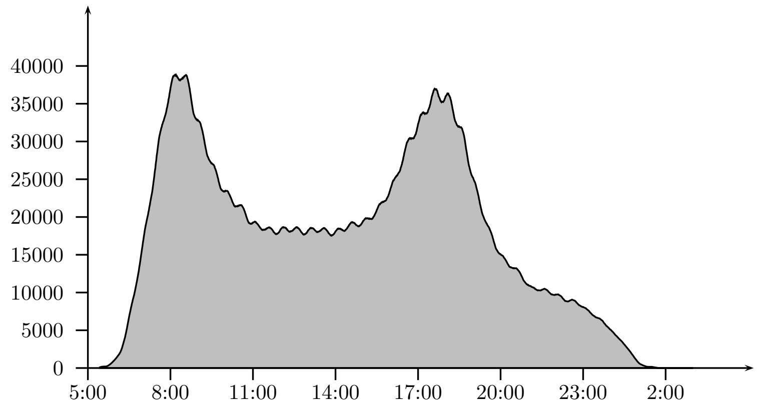


Figure 5.6: Number of passengers in the system during the day in the undisrupted situation.

We note that some of the involved lines in reality continue beyond the terminal stations shown in Figure 5.5 i.e. to the south-west of Dordrecht (Ddr), north of Alkmaar (Amr) and Zwolle (Zl), east of Deventer (Dv), south-east of 's-Hertogenbosch (Ht), and south of Arnhem (Ah). However, we perform a spatial aggregation of the network in the instances by assuming that passengers do not travel further than those six terminal stations. This is not a restriction in the verification of the approach as there are only limited rerouting possibilities in the peripheral parts of the network anyway.

The timetable is from a weekday and contains 2324 trips. All trips are assumed to be served with rolling stock of the type VIRM which is available in two variants with 4 and 6 carriages, respectively. The two variants have technical maximum capacities of 572 and 847 passengers per unit respectively. The maximum train length is assumed to be 14 carriages on all trips and it is possible for all lines to perform shunting operations at the terminal stations except Utrecht (Ut).

We created the passenger groups for the instances by matching passenger counts on each trip with OD data which resulted in 11415 passenger groups for the full day instance. To construct the deadlines of the passenger groups we assume that passengers are willing to accept at most an increase in traveling time of 50% plus 90 minutes.

Figure 5.6 shows the number of people in the system during the day in the instances. We observe that the peak hours around 8:00 and 17:00 are the busiest and therefore the most likely periods to experience capacity problems.

The disruptions considered in the computational tests all concern situations where a certain part of the network is unavailable for several hours. The timetable is updated according to current practice by canceling affected trips and turning trains on either side of the disruption.

The instances are described in Table 5.1. Each instance is named by a string like "D2T1O1R2P1" characterizing the instance. The name consists of five parts where the first part "Di" describes the place of the disruption; "D1" is between Rotterdam (Rtd) and The Hague (Gv), "D2" is between Gouda (Gd) and Utrecht (Ut), "D3" is between Utrecht (Ut) and Amersfoort (Amf), "D4" is between The Hague (Gv) and Leiden (Ledn), and "D5" is between Amsterdam (Asd) and Utrecht (Ut).

The second part "Ti" describes the time of the disruption; for instances with "T1" the blockage lasts from 16:00 to 19:00, and for instances with "T2" the blockage lasts from 11:00 to 15:00.

The third part "Oi" describes the original rolling stock schedules used for the undisrupted situation. Instances with "O1" are based on a circulation with 114 rolling stock units of the variant with 4 carriages and 44 rolling stock units of the variant with 6 carriages. The circulation is planned such that enough capacity is assigned to all trips to accommodate all passengers in the undisrupted situation, furthermore every train has significant slack capacity compared to the "full" capacity. Instances with "O2" use only 106 and 41 units of the two types respectively.

The fourth part "Ri" refers to the number of available reserve units. Instances with "R1" do not have reserve units. Instances with "R2" have three reserve units allocated in the network as one unit with 4 carriages at Amsterdam (Asd), one unit with 6 carriages at each of the stations of The Hague (Gv) and Rotterdam (Rtd). Instances with "R3" have six reserve rolling stock units distributed as two units with 4 carriages in Amsterdam (Asd), one unit with 4 carriages in Amersfoort (Amf), one unit with 6 carriages in The Hague (Gv), and one of each type of unit in Rotterdam (Rtd).

The fifth part "Pi" specifies the number of passengers in the system. Instances with "P1" have 422022 passengers distributed over the day as shown in Figure 5.6. Instances with "P2" have 15% more passengers in all passenger groups (total 485375 passengers) relatively distributed in the same way as in the "P1" instances.

Objectives

The objective in the instances consists of the system related goals and the service objectives. As described earlier the service related objectives are measured by the

Instance	Disruption	Time	Assigned units	Reserve units	Passengers	Passenger groups
D1T1O1R1P1	Rtd-Gv	16:00 - 19:00	114, 44	0	422022	11415
D1T1O1R2P1	Rtd-Gv	16:00 - 19:00	114, 44	1, 2	422022	11415
D1T1O1R3P1	Rtd-Gv	16:00 - 19:00	114, 44	4, 2	422022	11415
D1T1O2R1P1	Rtd-Gv	16:00 - 19:00	106, 41	0	422022	11415
D1T1O1R2P2	Rtd-Gv	16:00 - 19:00	114, 44	1, 2	485375	11415
D1T2O1R1P1	Rtd-Gv	11:00 - 15:00	114, 44	0	422022	11415
D2T1O1R1P1	Gd-Ut	16:00 - 19:00	114, 44	0	422022	11415
D2T1O1R2P1	Gd-Ut	16:00 - 19:00	114, 44	1, 2	422022	11415
D2T1O1R3P1	Gd-Ut	16:00 - 19:00	114, 44	4, 2	422022	11415
D2T1O2R1P1	Gd-Ut	16:00 - 19:00	106, 41	0	422022	11415
D2T1O1R2P2	Gd-Ut	16:00 - 19:00	114, 44	1, 2	485375	11415
D2T2O1R1P1	Gd-Ut	11:00 - 15:00	114, 44	0	422022	11415
D3T1O1R1P1	Ut-Amf	16:00 - 19:00	114, 44	0	422022	11415
D3T1O1R2P1	Ut-Amf	16:00 - 19:00	114, 44	1, 2	422022	11415
D3T1O1R3P1	Ut-Amf	16:00 - 19:00	114, 44	4, 2	422022	11415
D3T1O2R1P1	Ut-Amf	16:00 - 19:00	106, 41	0	422022	11415
D3T1O1R2P2	Ut-Amf	16:00 - 19:00	114, 44	1, 2	485375	11415
D3T2O1R1P1	Ut-Amf	11:00 - 15:00	114, 44	0	422022	11415
D4T1O1R1P1	Gv -Ledn	16:00 - 19:00	114, 44	0	422022	11415
D4T1O1R2P1	Gv -Ledn	16:00 - 19:00	114, 44	1, 2	422022	11415
D4T1O1R3P1	Gv -Ledn	16:00 - 19:00	114, 44	4, 2	422022	11415
D4T1O2R1P1	Gv -Ledn	16:00 - 19:00	106, 41	0	422022	11415
D4T1O1R2P2	Gv -Ledn	16:00 - 19:00	114, 44	1, 2	485375	11415
D4T2O1R1P1	Gv -Ledn	11:00 - 15:00	114, 44	0	422022	11415
D5T1O1R1P1	Asd-Ut	16:00 - 19:00	114, 44	0	422022	11415
D5T1O1R2P1	Asd-Ut	16:00 - 19:00	114, 44	1, 2	422022	11415
D5T1O1R3P1	Asd-Ut	16:00 - 19:00	114, 44	4, 2	422022	11415
D5T1O2R1P1	Asd-Ut	16:00 - 19:00	106, 41	0	422022	11415
D5T1O1R2P2	Asd-Ut	16:00 - 19:00	114, 44	1, 2	485375	11415
D5T2O1R1P1	Asd-Ut	11:00 - 15:00	114, 44	0	422022	11415

Table 5.1: Instances for computational results.

total inconvenience experienced by the passengers. More specifically, we minimize the sum of the delay minutes and the penalty for passengers who leave the system.

For the system related objectives we minimize with highest priority the number of canceled trips and with secondary priority the number of changes to the shunting process and the number of off-balances (see Section 2.6). More specifically we use a cost of 500 for introducing a new shunting operation or changing the type of operation performed. Shunting a different number of units or canceling a shunting operation is penalized by 100, while off-balances cost 400. Finally, we use a penalty of 0.0001 for carriage kilometers to ensure that for two solutions with the same value for all other objective terms, the one with lower operating cost is used. Note that all other objective parameters outweigh the total contribution from carriage kilometers.

We note that for a concrete application of the approach it is up to the decision maker to decide on the trade-off between service and system objectives. However, we limit this study to a trade-off that favors the service oriented part of the rolling stock rescheduling process and therefore we apply relatively low costs on changes to the system.

5.8.2 Results

We first investigate the performance of the approach for different values of the parameter α on a subset of the instances. Recall from Section 5.6 that α is a parameter that weighs the feedback from the current iteration against feedback from earlier iterations. We then run our approach on the remaining instances with a fixed value of α and discuss the results.

Parameter α

To analyze the effect of α we use the subset of instances named "DiT1O1R2P1" where $i = 1, \ldots, 5$. In Table 5.2 we report the performance of the approach on the instances for distinct values of α in the interval $0 \leq \alpha \leq 1$.

In each run we performed 30 iterations. The table shows the service and system cost of the best solution found as well as in which iteration the best solution was found. Finally, the table shows in the last column how many unique solutions were found in the 30 iterations.

For the instance "D1T1O1R2P1" we plotted the traversal of the algorithm in the two-dimensional objective space in separate diagrams for each α value in Figure 5.7. Similarly Figure 5.8 contains diagrams for the tests on instance "D5T1O1R2P1".

Instance	α	Service	System	Iteration	Unique solutions
D1T1O1R2P1	0.20	514983	3949	22	27
	0.35	514983	3949	20	30
	0.50	514869	4649	24	29
	0.65	518450	2749	24	30
	0.80	517241	4550	17	23
	1.00	536076	4449	3	5
D2T1O1R2P1	0.20	1241867	16252	14	30
	0.35	1244684	14551	16	30
	0.50	1245766	13851	22	30
	0.65	1244286	14451	25	30
	0.80	1244286	14651	15	30
	1.00	1250848	14951	3	7
D3T1O1R2P1	0.20	429898	3350	16	17
	0.35	429898	3350	9	19
	0.50	429898	3350	6	20
	0.65	429898	3350	5	20
	0.80	429898	3350	8	17
	1.00	429898	4150	1	3
D4T1O1R2P1	0.20	689086	7951	21	28
	0.35	689086	7851	17	29
	0.50	689086	7851	25	30
	0.65	689086	9952	2	29
	0.80	692419	8051	17	23
	1.00	717539	7652	7	8
D5T1O1R2P1	0.20	607200	9851	18	30
	0.35	607117	9151	13	30
	0.50	606274	11451	6	30
	0.65	611579	10351	21	30
	0.80	619260	8651	9	30
	1.00	644427	11051	21	30

Table 5.2: Results for instances "DiT1O1R2P1" with $i = 1, \ldots, 5$ for different values of α. Each row contains for a certain instance and α value the service and system objective of the best solution found and the iteration in which it was found. The last column shows the number of unique solutions found in the entire procedure.

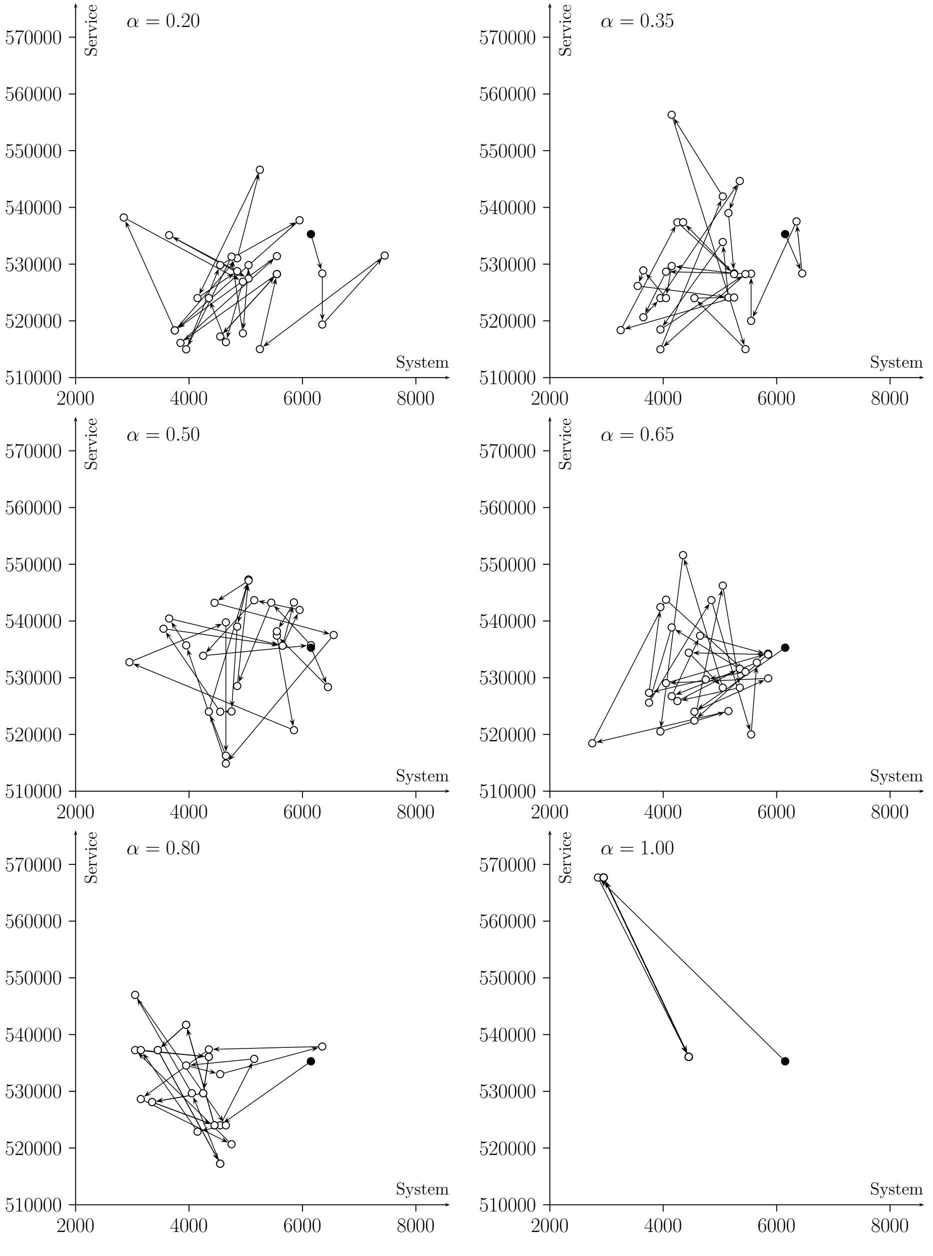


Figure 5.7: Traversal of the objective space for different settings of parameter α for instance "D1T1O1R2P1". The first iteration is marked by a solid dot.

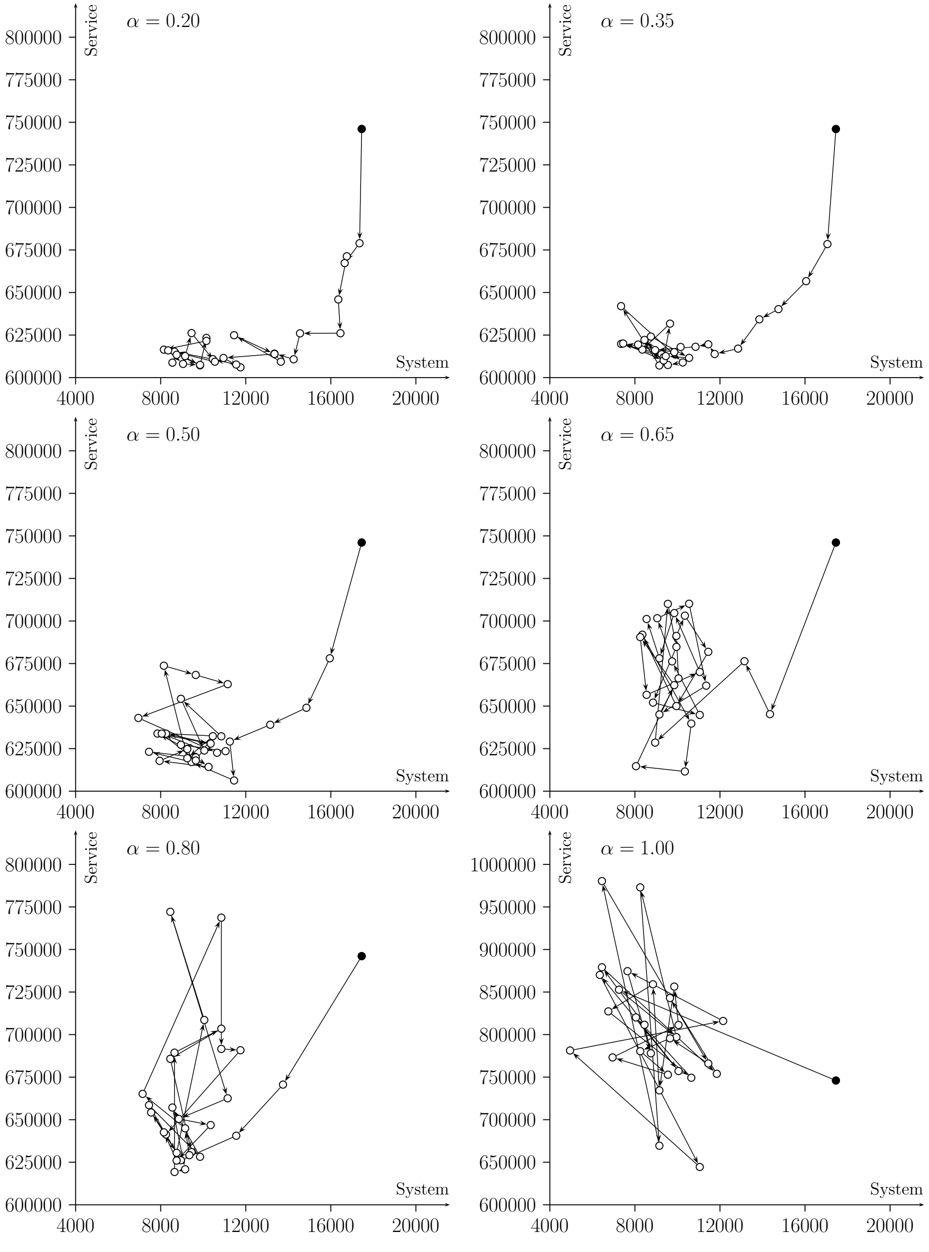


Figure 5.8: Traversal of the objective space for different settings of parameter α for instance "D5T1O1R2P1". The first iteration is marked by a solid dot. Note that the scale of the y-axis differs for $\alpha = 1.00$.

We observe that the tests with an α value of 1.00 are consistently worse than all other values for all instances. This is not surprising since only the feedback from the last iteration is taken into account in each step, and we therefore do not use the information obtained in the earlier iterations. In fact, on “D1”-“D4” we observe cyclical behavior of the algorithm. After a few iterations the algorithm alternates between the same two iterations. This behavior can be observed in Figure 5.7.

For the other values of α we observe that with the parameter $\alpha = 0.20$ the algorithm often takes longer to reach the best solution in the procedure. In Figure 5.8 it is particularly observable that the procedure takes smaller steps in the objective space with this α value. Setting the parameter to $\alpha = 0.80$ also seems inferior to the remaining values. For the tests with $\alpha = 0.35$, 0.50, and 0.65 we observe only limited difference in the performance of the algorithm. We therefore use the parameter value $\alpha = 0.35$ for the remaining computational tests.

All instances

We tested the approach on all instances using the fixed parameter value $\alpha = 0.35$. The results are shown in Table 5.3. The first column contains the instance names, and the next three columns show the minimum, average and maximum experienced running times for the optimization module over the 30 iterations of the algorithm. The next column shows the average running time for the simulation module of the algorithm. The columns named “Service” and “System” under “Best for system” denote the service and system objectives of the solutions that minimizes the system related objective. The columns named “Service” and “System” under “Best solution” denote the service and system objectives of solutions that minimize the sum of service and system objectives. Columns “IC-LB” and “WOC-LB” indicate the two types of lower bounds on the service objective.

For all instances with disruptions “D1”, “D3” and “D4” we found solutions without canceled trips. However, for all instances with disruptions “D2” and “D5” it was necessary to cancel one trip. The solutions that only concern the system objective (in columns “Best for system”) all have low system costs and relatively high service costs.

We can make a number of observations concerning the results. First, we observe that the approach is able to improve the service quality significantly in all instances at the cost of a number of changes to the system. The improvements are between 4% and 47%. For disruptions “D1” and “D3” we generally add 3 – 6 shunting operations and cancel or change 15 – 30 others. The solutions have 1 or 2 off-balances.

	OPT time			SIM	Best for system		Best solution			
Instance	min	avg	max	time	Service	System	Service	System	IC-LB	WOC-LB
D1T1O1R1P1	5.6	9.3	12.2	2.8	628211	1449	600017	4649	421007	544674
D1T1O1R2P1	5.6	8.6	14.6	2.6	578550	1249	514983	3949	421007	430681
D1T1O1R3P1	4.5	8.4	21.1	2.6	579315	1249	474465	4449	421007	424816
D1T1O2R1P1	4.4	7.6	12.2	2.8	660150	2049	622684	4750	421007	556666
D1T1O1R2P2	3.7	7.0	14.7	2.8	818667	1249	736754	6450	484126	532481
D1T2O1R1P1	10.2	17.8	41.9	2.8	303037	1149	287612	3049	287612	287612
D2T1O1R1P1	5.7	10.4	13.6	3.1	1818607	4248	1249484	15351	996484	1039719
D2T1O1R2P1	5.7	10.3	14.5	3.1	1818607	4248	1244684	14551	996484	1039719
D2T1O1R3P1	4.9	8.0	17.7	3.1	1818607	4248	1240150	12951	996484	1039719
D2T1O2R1P1	7.1	11.0	16.6	3.1	1805419	3849	1298913	16751	996484	1068677
D2T1O1R2P2	4.8	10.8	14.6	3.3	2567406	4248	1618741	17952	1145355	1263904
D2T2O1R1P1	22.4	203.3	4701.2	3.1	1115628	3248	743964	10051	721884	727088
D3T1O1R1P1	4.6	8.6	21.5	2.7	505971	2349	429898	3750	416595	426838
D3T1O1R2P1	4.5	7.4	21.1	2.6	505971	2049	429898	3350	416595	426838
D3T1O1R3P1	4.6	7.8	27.8	2.7	505971	2049	429898	3350	416595	426838
D3T1O2R1P1	4.7	10.0	21.7	2.6	505971	3049	429898	5750	416595	426838
D3T1O1R2P2	4.5	10.1	21.6	2.7	765566	2049	515438	6651	479347	500599
D3T2O1R1P1	15.9	32.8	70.3	3.0	625948	949	454710	4450	444589	450480
D4T1O1R1P1	3.2	5.3	12.9	2.6	839882	1349	698616	7451	666626	674203
D4T1O1R2P1	3.3	5.1	14.0	2.4	839882	1349	689086	7851	666626	670784
D4T1O1R3P1	3.7	5.4	12.8	2.6	839882	1349	686646	8551	666626	670784
D4T1O2R1P1	2.9	4.7	12.3	2.6	842661	1249	704199	8751	666626	675355
D4T1O1R2P2	3.4	7.5	14.0	2.7	1191093	1349	846999	10952	766343	785204
D4T2O1R1P1	6.8	12.7	21.1	2.8	494192	1048	460461	5650	459536	459572
D5T1O1R1P1	14.3	85.8	404.0	2.8	1116346	1449	610295	9551	520867	536397
D5T1O1R2P1	20.5	192.7	544.6	2.6	1116650	1449	607117	9151	520867	536397
D5T1O1R3P1	22.9	268.3	954.3	2.7	1116650	1449	596495	10151	520867	536397
D5T1O2R1P1	15.5	57.8	788.4	2.8	1122055	1449	665010	14251	520867	536397
D5T1O1R2P2	13.6	79.6	1612.8	3.0	1722815	1449	915054	18453	599038	635239
D5T2O1R1P1	6.6	148.8	3082.6	2.9	765585	549	417038	7451	388216	388216

Table 5.3: Results for all instances. The table contains computation times for the optimization and simulation module. It contains the best solutions found for the system and the best overall solutions. The last two columns contain the lower bounds based on infinite capacity and operator control respectively.

For disruptions "D2", "D4" and "D5" we generally experience more changes to the system; 10 − 17 new shunting operations are added while 20 − 35 are changed or canceled. The solutions for these disruptions have up to 7 off-balances. These changes are relatively few from a practical point of view considering the fact that the system contains 14 stations.

Second, we notice that we are generally able to alleviate more passenger inconvenience in the instances with more available reserve units. This concerns instances "R1" with no reserve units, "R2" with three reserve units, and "R3" with six reserve units. The results for these instances are the first, second and third rows respectively in each block in Table 5.3. Especially for disruption "D1" (between The Hague (Gv) and Rotterdam (Rtd)) the extra reserve units seem to be highly beneficial for the service objective. This indicates that the reserve units at those stations can be allocated where the capacity is needed.

Third, we notice that there is a significant difference between disruptions in the peak hours and in the off-peak hours. This is apparent when comparing the disruptions at time 16:00-19:00 (named "T1") to the disruptions at time 11:00-15:00 (named "T2") – first and sixth row in each block in Table 5.3 respectively. Generally, less delay minutes occur in a four hour off-peak disruption compared to a three hour disruption during the peak. Also, it is possible for all disruptions "D1",. . .,"D5" to bring the service objective close to the IC-LB in the off-peak instances. This indicates that capacity is not a bottleneck in those situations.

Fourth, we observe that the instances with less slack capacity in the original plan generally lead to rescheduled solutions with worse service objectives. This concerns "O2" (fourth row in each block in the table) which is planned using much less rolling stock than "O1" (first row), and "P2" (fifth row) which has 15% more passengers than "P1" (second row) but uses the same amount of rolling stock. The results imply that less slack capacity potentially leads to more inconvenience during disruptions.

Fifth, for disruptions "D3" and "D4" we see that the service objective of the best solutions found is quite close to the lower bound given by IC-LB. This means that most of the experienced inconvenience is due to the changed timetable rather than due to lack of capacity. For all other instances we experience that the changes to the timetable are still by far the major contributor to passenger delays, but lack of capacity often contributes around 10%-30% of the delays.

Sixth, in the lower bounds given by WOC-LB, the passenger traveling paths are decided so that the sum of the delays is minimized. In contrast to the IC-LB the arcs have limited capacity and the WOC-LB thus explains some of the delays by the

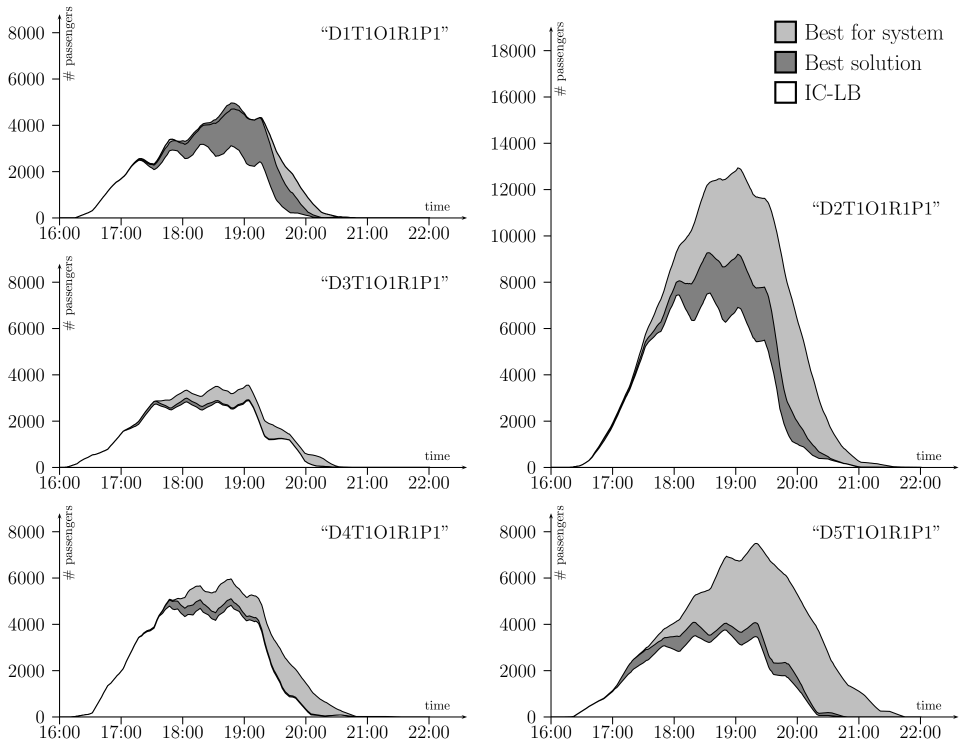


Figure 5.9: Number of additional passengers in the system during and after the disruption for instances "D1T1O1R1P1", . . ., "D5T1O1R1P1". The area under a curve is the cumulative delay.

lack of capacity. For several instances this lower bound closes a significant part of the gap between the service objective of the best found solution and the lower bound on the service objective.

In a separate set of computational tests we solved the instances with greatly different weights on the system objectives, i.e. divided or multiplied by a factor ten. But the best solutions found seemed to be comparable to the solutions reported for the current weights. This is probably due to the system costs being relatively small compared to the service costs.

The numbers of passengers that are delayed at a given time are shown in the diagrams of Figure 5.9. More specifically, the numbers of passengers in the system in addition to the passengers in the corresponding undisrupted situation are shown. The diagrams concern the solutions to the first instance in each block in Table 5.3. The number of passengers in the system is always at least as large as the number of passengers in the system in the undisrupted situation. This is because people depart at the same time and they arrive no earlier than they are able to in the undisrupted situation. Each diagram contains the IC lower bound which shows how many passengers are delayed because of the changed timetable. Furthermore, each diagram contains the curve for the best found solution (dark gray) and the solution that is best for the system (light gray). We notice that for the complicated disruptions "D2" and "D5" there is a significant gap between the best found solution and the solution that is best for the system. For disruptions "D3" and "D4" the best solution is almost identical to the lower bound. Finally, we observe that for disruption "D1" the best solution contains more delayed passengers than the best system solution at some time instants. However, after the disruption the best solution brings the delayed passengers out of the system faster than the best system solution.

Computation times

The computation times of the optimization module and the simulation module for each instance are shown in Table 5.3. For the optimization module the minimum, the average, and the maximum computation time over all iterations is shown in seconds. For the simulation time only the average time is shown since the computation time for that module is very consistent. Running times for the feedback mechanism are not shown as they are neglectable in comparison with the computation times of the other modules.

The tests were performed on an Intel Core 2 duo 3.33 GHz desktop computer with 3 GB of RAM. For the optimization we used CPLEX 11.0 on a single processor while the code for the simulation and the feedback mechanism was written in Java.

The computation time of the optimization module seems to be instance dependent. For all instances involving the disruptions "D1", "D3" and "D4" we experienced computation times of up to 70 seconds for an iteration, and on average the optimization performed each iteration in around 5 – 10 seconds. Instances involving the early disruption "T2" take longer, as a larger part of the day is rescheduled. The instances involving disruption "D2" need longer computation times, especially for the instance "D2T2O1R1P1" the computation time is on average around three minutes per iteration.

The instances involving disruption "D5" generally require longer computation times for the optimization. In fact, the iterations for instance "D5T1O1R3P1" all took at least 22 seconds and on average more than 4 minutes. We attempted to reduce the running time by altering some of the basic CPLEX parameters and adding some valid cuts. This worked on some instances but increased the running time on others. We were thus unable to find settings that consistently improved running times compared to the default settings. However, tuning CPLEX for this set of instances is out of the scope of this study. We did, nevertheless, try another MIP solver for the optimization module. We used Gurobi 3.0.0 which was faster on especially the "D5" instances but somewhat slower on a number of other instances. A comparison between the running times of the two tested MIP solvers would be unfair since Gurobi 3.0.0 utilizes both cores in the processor and is also a much newer release.

5.9 Conclusions and future research

In this chapter we described a heuristic approach for improving the service aspect of the rolling stock schedule during disruptions. The improvements in service quality come at the cost of changes to the system. For all instances in the computational tests we were able to improve the service objective and in some cases even to reach the lower bound.

In our approach we apply a number of assumptions on the behavior and interaction of passengers. Passengers are assumed to want to arrive at their destinations as quickly as possible and are assumed to leave the system if their delays exceed certain thresholds. Also, passengers are assumed to compete for the scarce capacity in the sense that capacity is assigned to groups of passengers based on their size. We claim that these assumptions reasonably reflect the real situation. But the approach is modular and can be adapted to a system with significantly different assumptions on passenger behavior by changing single components in the iterative approach.

The lower bounds provided in this chapter are admittedly rather weak for some instances. For future research we suggest investigating solution approaches for the lower bound SOC-LB. It is appealing to develop a cut-and-price approach for solving model (5.5) – (5.9). Such an approach would combine column generation for the paths of passengers and row generation based on Benders decomposition to add valid cuts for the assignment of capacity.

The computation time on most instances is appealing for real-time use, although the running time seems to depend on the structure of the solution. The approach provides a feasible solution in every iteration and may therefore be terminated when a satisfactory solution is reached or when the available computation time is up.

Another means to reducing computation time would be to utilize the rolling horizon framework introduced in Chapter 4. This integration would also allow us to account for the uncertainty of the system although it requires a more realistic notion of informedness of passengers. We have, however, chosen not to utilize the rolling horizon framework for this study of rolling stock rescheduling with dynamic passengers since it would obscure the contributions of the iterative framework.

Chapter 6

Case Study: Rolling Stock Rescheduling in Short-term Planning

In the course of developing the framework and models for disruption management of rolling stock, we realized that the models are also applicable in other stages of the planning process of the railway operations. In particular, rolling stock rescheduling problems occur in the short-term planning stage of the planning process as well. Recall from Section 2.2 that short-term planning is concerned with adapting the existing generic schedules from the tactical phase to specific days. This leads to rolling stock rescheduling problems when the timetable is modified or different passenger demands are projected. Still we intend to minimize the changes to the schedules like in real-time planning. The rescheduling problems in short-term planning are to a large extent similar to those encountered in disruption management. In fact, the major differences from disruption management are the absence of the uncertainty of disruptions and the available time to come up with solutions.

In the spring of 2009 the short-term planning department of NS was lacking the planning capacity to efficiently deal with the challenges of manually constructing rolling stock schedules for the coming summer. The need for planning capacity presented an exciting opportunity to test the rescheduling models on real-life cases and hopefully achieve implementable results.

In this chapter we discuss the experiences from two real-life case studies from the short-term planning department of NS. The chapter is based on Nielsen (2010). We describe the particularities of short-term rolling stock planning and discuss how the rescheduling models developed in this thesis can be utilized in this planning step. In addition, we elaborate on the experiences made from working with practitioners, and the importance of presenting potential solutions to the decision makers in a convenient and comprehensible manner. We describe the two cases and how our resulting solutions were implemented in practice.

The prototype models developed for short-term rolling stock rescheduling are known as TAM (*Tool voor de Aanpassing van de Materieeldiensten*[1]). The tool is based on the Composition Model combined with the Duty Path Model introduced in Chapter 3. TAM is currently in use for the short-term planning department at NS where it is being extended to incorporate more details of the planning process. Also, the interfaces with existing systems are being streamlined to speed up the whole short-term planning process even further.

6.1 Short-term planning

In Section 2.2 we introduced the short-term planning process as a series of planning tasks with a time horizon of a few days up to two months. The overall purpose of short-term planning is to adapt the generic week plan to specific calendar days. The adaptation concerns the timetable and the resource schedules. The timetable is modified by removing, adding or rerouting train services according to the available infrastructure and the passenger demand of specific days. The resource schedules are then rescheduled to serve the modified timetable.

The major difference between short-term planning and the earlier planning phases is the existence of the plans from the tactical planning steps. The scheduling conducted in the tactical phase is performed from scratch whereas short-term planning concerns adapting the existing generic schedules. We emphasize that short-term planning concerns *rescheduling* rather than scheduling from scratch.

Rolling stock rescheduling in the short-term planning phase generally has the following goals: (*i*) Minimize operational costs while providing sufficient seat capacity. (*ii*) Minimize the implied changes to the local shunting operations. (*iii*) Ensure that the modified schedules fit into the context in which they exist, i.e. ensure that modified rolling stock duties start and end at the appropriate stations. The exact weights

[1] Tool for the adaptation of rolling stock duties.

of the mentioned goals depend on the nature of the short-term rescheduling instance in question.

The short-term rolling stock rescheduling at the short-term planning department at NS is generally conducted manually and in two steps; first, analyzing how well the generic schedules fit with the specific calendar days, and second, adapting the assignment of rolling stock to each train iteratively. In the first step the generic schedules are analyzed to determine whether each train meets expected seat demand for specific calendar days. This analysis results in a list of suggestions for which train services have too little or too much capacity and therefore may have their short-term capacity adapted. In the second step the planners attempt to meet the suggestions by iteratively moving units from one train service to another.

The manual planning is a local search approach which may work well for small instances, but it may be difficult to meet all proposed changes from the analysis. Especially the very complex Noord-Oost case is a challenging instance for manual short-term planning. The application of TAM requires a somewhat different approach to short-term planning since the idea is to build a new solution for a part of the system while minimizing the implied changes from the generic schedule.

6.2 Application cases

The first major real-life application of our rescheduling model was the *Zomerplan 2009* (in English *Summer plan*). This case concerns a six week period in the summer of 2009 where many people are on holiday. In this period NS operates a reduced timetable in the sense that some trains are canceled. In addition, lower passenger demands are expected on the remaining trains in the timetable. The task was to reschedule the rolling stock on a number of intercity and regional lines to reduce operational costs while meeting the reduced demand and limiting the changes to the local operations.

We refined the methods to handle a broader set of situations like infrastructure maintenance projects. We then applied the methods to the so-called *Week 37 case.* This case involves a number of interconnected intercity lines on Saturday, September 12, to Sunday, September 13, 2009. For those lines the timetable was temporarily changed over the weekend due to major track maintenance projects between some major cities. The goal was to reschedule the rolling stock to the adapted timetable in such a way that the assignment of capacity to trips is similar to that of the generic plan. At the same time we wanted to minimize the changes to the local operations. We here describe the two cases in more detail.

6.2.1 The Zomerplan 2009

Every summer during the vacation period, NS operates a reduced timetable. In 2009 this concerned a six week period in July and August where fewer passengers travel by train. Especially the number of passengers in the peak hours of the weekdays is reduced in this period since many regular commuters are on vacation. In response, NS cancels a number of lines and operates others at reduced frequency.

Canceling train services already leads to considerable savings in operational costs, but there is a further potential for savings by operating shorter trains on the remaining lines. This was the motivation for investigating the applicability of TAM for the Zomerplan.

It was thus essential to have a reliable forecast of the expected number of passengers on each train service in the reduced timetable. To achieve this the decision makers decided to use the average of the passenger counts collected in the same period in 2008 plus some percentages to account for the variance and a further addition for the expected yearly growth in demand. After this process, the forecasted passenger demands were given as input in the form of the expected number of passengers on each trip for each weekday. The estimation of the passenger demands is not part of the TAM package, but rather part of the necessary input.

Initially, the case consisted of two connected intercity lines named the 2100 and 2600 lines. The 2100 line runs from Amsterdam (Asd) via Roosendaal (Rsd) to Vlissingen (Vs) while the 2600 line runs from Amsterdam (Asd) to The Hague (Gvc). After constructing a number of high quality solutions for these lines and further discussions with planners it was decided to include the two lines 2000 and 8800 that run from The Hague (Gvc) to Utrecht (Ut) and from Leiden (Ledn) to Utrecht (Ut), respectively. These sets of lines can thereby share rolling stock by coupling and uncoupling units in The Hague. The intercity lines are shown as solid lines in Figure 6.1.

Furthermore, seven regional lines in the south-eastern part of the country were added to the case. Lines 6800 and 6900 form a closed circulation between stations Maastricht (Mt), Roermond (Rm) and Heerlen (Hrl). Lines 4400, 5200, 6400 and 9600 form a closed circulation in the region between stations Tilburg West (Tbw), Eindhoven (Ehv), Deurne (Dn), Weert (Wt), 's-Hertogenbosch (Ht) and Nijmegen

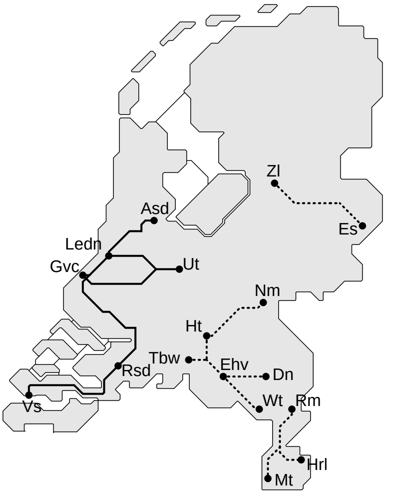


Figure 6.1: Lines included in the Zomerplan 2009 case.

(Nm). Line 7900 runs between Zwolle (Zl) and Enschede (Es). The regional train lines are shown as dashed lines in Figure 6.1.

The intercity lines in the case are all served by electrically powered double-decked units of the type VIRM which are available in two types with 4 and 6 carriages respectively. The maximum train length is 12 carriages for the 2100/2600 lines and 8 carriages for the 2000/8800 lines which results in 7 and 3 possible compositions to serve the trips on the lines, respectively. The key numbers for a single day of the instance are summarized in Table 6.1. The table includes information on the number of trips in each set of lines and the number of connections where it is possible to change the composition of a train. Also, the possible repositioning trips are included in the instance. These trips denote nightly repositioning trips between stations and non-service trains from stations to nearby depots.

The 6800/6900 lines and the 4400/5200/6400/9600 lines are served by single-decked electrical units of the type Mat '64 which is available in lengths of 2 and 4 carriages. The maximum allowed train length is 6 carriages which means 6 different compositions are possible to assign to the trains. The 7900 line is served by diesel powered units of the type DM '90 which consists of 2 carriages. A maximum train length of 8 carriages results in 4 possible compositions to be assigned.

Lines	# trips	# repositioning trips	# compositions	# shunting possibilities
2100, 2600	190	31	7	171
2000, 8800	138	24	3	108
6800, 6900	135	19	6	84
4400, 5200, 6400, 9600	352	51	6	262
7900	78	10	4	89

Table 6.1: Number of trips and repositioning trips of all parts of the instance along with the number of allowed compositions on trips and the number of connections with shunting possibilities. The numbers are for a single day.

6.2.2 The week 37 case

During planned infrastructure maintenance projects, NS often has to temporarily modify the timetable. Such projects take place almost every weekend and the resulting changes to the assignment of resources are handled in the short-term planning. Often these situations can be handled by relatively simple modifications to the resource assignment, but for some cases the changes are more involved.

In the weekend of week 37 in 2009, two major infrastructure projects were planned on the Dutch railway network. The tracks between Zwolle (Zl) and Groningen (Gn) would be out of service the entire weekend and the direct connection between The Hague (Gvc) and Utrecht (Ut) would be closed until Sunday 13:00. This has consequences for a number of lines that utilize these parts of the network in the generic schedules. A number of regional train lines were canceled or rerouted which led to relatively small rolling stock rescheduling problems that were solved manually. But the maintenance projects also had consequences for the complex set of lines known as the Noord-Oost lines which was introduced in Section 4.7.1.

In the generic plan the trains on the Noord-Oost lines undergo complex shunting operations in Zwolle (Zl) where the north-bound trains on the 500 and 700 lines are split and continue in the direction of Leeuwarden (Lw) and Groningen (Gn). Correspondingly, the south-bound trains from the two northern terminal stations are combined in Zwolle. Also, one or two units are often uncoupled from the north-bound trains just like one or two units are coupled to the south-bound trains making the station of Zwolle an important hub for the Noord-Oost lines. In a similar fashion, the trains on the 500 and 1700 lines are combined and split in Utrecht (Ut) when

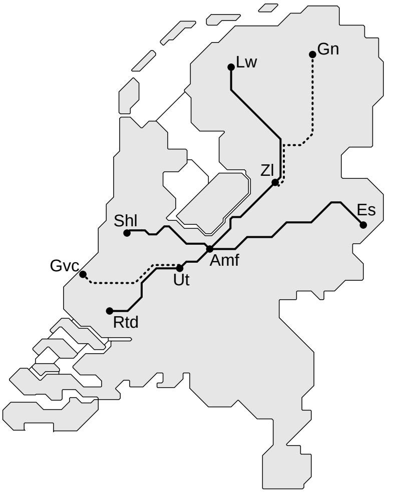


Figure 6.2: Lines included in the week 37 case. Solid lines are preserved and dashed lines are canceled.

going to or coming from Rotterdam (Rtd) and The Hague (Gvc). However, usually no units are coupled or uncoupled in Utrecht.

The timetable for the Noord-Oost lines was adapted for the weekend by canceling the trips from Zwolle (Zl) to Groningen (Gn) on the 500 and 700 lines which means the trains would no longer be split or combined in Zwolle. Similarly, the splitting and combining in Utrecht is given up until Sunday 13:00 as the 500 and 1700 lines just continue to Rotterdam (Rtd). The map in Figure 6.2 shows the Noord-Oost lines, the parts of the lines that are not operated in this particular case are shown as dashed.

The Noord-Oost lines are operated by the so called Koploper units which are single-decked electrically powered intercity units that are available in lengths of 3 and 4 carriages. Most of the trips can accommodate trains of either up to 12 carriages or up to 15 carriages which leads to 16 and 31 different compositions respectively. The key numbers for the instance are summarized in Table 6.2.

The number of passengers in the case was not expected to deviate significantly from the generic plan. So the objective in the case was to reschedule the rolling stock to assign similar capacity as in the generic plan to the remaining train services. At the same time the changes to the local plans should be minimized.

Lines	# trips	# repositioning trips	# compositions	# shunting possibilities
Noord-Oost	1270	197	16/31	768

Table 6.2: Number of trips and repositioning trips of all parts of the week 37 instance along with the number of allowed compositions on trips and the number of connections with shunting possibilities. The numbers are for the entire weekend.

6.3 Solving the cases

To solve the cases we needed to transform the cases to fit into the solution methodology – this involved a number of techniques and assumptions. We decided to break down the cases into smaller instances and to apply Tam to those. In this section we describe how this decomposition was performed and how the models were tailored for the resulting instances. Furthermore, we discuss how we cooperated with the planners of NS in the process.

6.3.1 Decomposition

The rolling stock schedules for the weekends of the Zomerplan 2009 were not to be changed which means that the rolling stock balance between Sunday evening and Monday morning was given by the rolling stock schedules for the weekend. The same holds for the rolling stock balance between Friday evening and Saturday morning. The passenger forecasts were identical for all weekdays.

The given passenger forecasts combined with the fixed Monday morning and Friday evening balances immediately implied that the construction of a rolling stock schedule for all six weeks boils down to the construction of a one week rolling stock schedule that is repeated six times.

Furthermore, the timetables of each day of the week were identical with the exception of a few optional repositioning trips on Monday and Friday. This implied that we could use more or less the same assignment of rolling stock on each single day. We therefore decided to make three schedules. The first schedule was for Monday and used the fixed rolling stock balances from Sunday evening for the morning position of all available units. The second schedule would be used for Tuesday, Wednesday and Thursday, and would have a *cyclical* inventory which means that the morning balance should be equal to the evening balance. The third schedule was for Friday and it used the fixed rolling stock balances from Saturday morning for the evening position of all available units.

As mentioned earlier, the lines included in the Zomerplan case formed closed circulations for the intercity lines 2100, 2600, 2000, 8800 served with the type VIRM, the regional lines 6800, 6900 served with Mat ’64, the regional lines 4400, 5200, 6400, 9600 also served with Mat ’64 and the regional line 7900 served with DM ’90. The decision to create separate solutions for Monday, for Tuesday/Wednesday/Thursday, and for Friday thus resulted in 12 instances in total.

For the week 37 case we decided to divide the case into two instances by creating schedules for Saturday and Sunday separately. The balance on Saturday morning was given by the position of the trains on Friday evening in the generic rolling stock schedule and in the same way the balance on Sunday evening was given by the positions of trains on Monday morning in the generic rolling stock schedule.

6.3.2 Preparing the data

TAM is a tool that solves an instance of RSRP and thus needs its input structured as described in Section 3.1.1 and formalized in Section 3.2.4.

The data describing the generic schedule was given as a database file containing the generic rolling stock duties. Furthermore, the data on changes to the timetable for specific weekdays were given as a simple spreadsheet. The modified passenger demand was given as the expected number of passengers per train service between each pair of major stations along the involved lines. The shunting possibilities were simply given as a set of rules describing where shunting is possible in addition to the shunting operations in the generic schedules.

We implemented in TAM an automated tool that combines the database of generic rolling stock duties with the information on the changes to the timetable to construct the set of trips $\mathcal{T}$ and the set of connections $\mathcal{C}$ for each instance. The automated tool defines a trip as a part of a train service between two stations where it is not possible to change the assigned rolling stock underway. Furthermore, the automated tool combines the information on the generic rolling stock duties with the given shunting rules to create the set of allowed composition changes $\varrho(c)$ as well as coupling and uncoupling times $\tau^+(c)$ and $\tau^-(c)$ for each connection $c \in \mathcal{C}$. The automated tool ensures that all shunting operations that are performed in the generic rolling stock plan are also allowed in the instances. This means that if there are no changes to the timetable and the rolling stock availability, the generic plan is a feasible solution to the instances constructed by the automated tool.

6.3.3 Model

To create the circulation for each instance TAM uses the Composition Model (3.1) – (3.14) with a number of the objective terms introduced in Section 3.4. We discuss the objective function below.

In accordance with the rules given by the planners we allowed some deviations in the balances at some stations located close to each other since the off-balances could easily be resolved by repositioning trips. We modeled this by adding constraints (3.20) to the model for the evening balances. We here repeat the constraint for completeness.

$$I^{\infty}_{s,m} = i^{\infty}_{s,m} + D^{\infty,+}_{s,m} - D^{\infty,-}_{s,m} \qquad \forall s \in \mathcal{S}, m \in \mathcal{M}$$

Recall that these constraints account for the deviations in the evening balances. Similar constraints are added to the model for the morning off-balances based on corresponding deviation variables $D^{0,+}_{s,m}$ and $D^{0,-}_{s,m}$. These variables denote the surplus or shortage, respectively, of units of type m at station s in the morning inventories.

For the schedules for Tuesday, Wednesday and Thursday of the Zomerplan we added constraint (6.1) below to ensure the cyclicity of the inventory in the schedules.

$$I^{0}_{s,m} = I^{\infty}_{s,m} \qquad \forall s \in \mathcal{S}, m \in \mathcal{M} \tag{6.1}$$

Specifically, the constraints state that the number of units of type $m \in \mathcal{M}$ available at station $s \in \mathcal{S}$ in the morning is equal to the number of units of the same type that ends its duty at the station in the evening. A cyclic inventory ensures that the same schedule is repeatable on consecutive days which is exactly what we want for those three weekdays.

For the objective function we used the linear function (6.2). The terms in the objective function were introduced and discussed in detail in Section 3.4.

$$\begin{aligned}
&\sum_{t\in\mathcal{T}}\sum_{p\in\eta(t)} v\cdot\texttt{carr}(p)\cdot\texttt{km}(t)\cdot X_{t,p} + \\
&\sum_{t\in\mathcal{T}}\sum_{p\in\eta(t)} w\cdot\max\{d_t-\texttt{cap}(p),0\}\cdot\texttt{km}(t)\cdot X_{t,p} + \\
&\sum_{c\in\mathcal{C}}\sum_{q\in\varrho(c)} \gamma_{c,q}Z_{c,q} + \\
&\sum_{s\in\mathcal{S}}\sum_{m\in\mathcal{M}} \left(d^{\infty,+}_{s,m}D^{\infty,+}_{s,m} + d^{\infty,-}_{s,m}D^{\infty,-}_{s,m}\right) + \\
&\sum_{s\in\mathcal{S}}\sum_{m\in\mathcal{M}} \left(d^{0,+}_{s,m}D^{0,+}_{s,m} + d^{0,-}_{s,m}D^{0,-}_{s,m}\right)
\end{aligned} \tag{6.2}$$

Here the first term is (3.15) which accounts for operational costs through penalizing carriage kilometers. Recall that carriage kilometers are the key measure of the operational cost of a rolling stock schedule. The second term is (3.17) and relates to the service level requirements by penalizing the assignment of too little capacity. The third objective term is (3.19) and is added to account for the changes to the shunting patterns. The fourth objective term is (3.21) and penalizes off-balances at the end of the day. Similarly, the fifth objective term penalizes off-balances at the beginning of the day.

The model needs a number of parameters for the conflicting objectives.

- Parameter v denotes the cost of one carriage kilometer.
- Parameter w denotes the cost of one seat shortage kilometer.
- Parameters $\gamma_{c,q}$ denote the cost of performing composition change q at connection c.
- Parameters $d^{\infty,+}_{s,m}$ and $d^{\infty,-}_{s,m}$ denote the cost of a surplus or shortage of one unit of type m at station s at the end of the day. Similarly, parameters $d^{0,+}_{s,m}$ and $d^{0,-}_{s,m}$ denote the cost of a surplus or shortage of a unit at the beginning of the day.

The cost $\gamma_{c,q}$ of performing composition change q at connection c depends on the shunting operation taking place at connection c in the generic schedule. The possible changes are: Introducing a new shunting operation, canceling a shunting operation, changing the type of operation, and changing the type or number of units shunted (see Section 2.6.1).

For the Zomerplan 2009 case it is straightforward to compare the shunting operations in a potential solution to those of the generic plan since the short-term timetable for the involved lines is identical, and each new connection is identical to an old one. For the week 37 case, however, there are changes to the timetable in the form of canceled trips. This means that some connections are not identical to old ones. More specifically, the splitting and combining of trains in Zwolle and Utrecht are changed to through going trains. Since shunting is not allowed in Utrecht the changed connections there do not pose a problem. But in Zwolle it is possible to couple and uncouple up to two units to or from a train so the question arose on how to rate potential shunting operations at this station. After discussions with planners we decided to treat all introduced shunting operations in Zwolle as minor changes to the generic plans on par with changing the number or type of units shunted.

Duty generation

The Composition Model results in an assignment of compositions to the trips in the modified set of trips $\mathcal{T}$, and an assignment of composition changes to connections. TAM utilizes the Duty Path Model (3.70) – (3.73) for converting the composition assignment to a set of duties. The Duty Path Model allows us to create modified duties that resemble the original duties as much as possible.

6.3.4 Cooperation with practitioners

Throughout the process of creating solutions for the cases we cooperated with the practitioners by meeting frequently to discuss a number of issues: Realistic values of parameters settings; details on how to match the new timetable with the old one; the quality of solutions and how they could be improved in their opinion.

In this cooperative process, two instruments proved particularly useful. One was a small visualization tool we designed for the cases, and the other was the ability to construct multiple solutions.

Visualization

We constructed a small visualization tool to assist the cooperation with the planners. This tool shows a graphical representation of all duties and accompanying shunting operations in a solution as well as off-balances, train compositions, capacities and seat shortages. Furthermore, the tool gives an easy overview of all key numbers such as expected seat shortages, carriage kilometers, and number of changes to the shunting operations. Figure 6.3 shows screenshots from the visualization tool.

In addition to showing a solution, the tool also allows a user to directly see the changes from the generic plan. In particular, we presented the changes to the shunting operations by a screen intuitively showing all shunting changes of a station. This proved very useful when the planners assessed the likelihood of a potential solution being accepted by local planners.

Another useful feature of the visualization module is that one can easily evaluate the consequences of the rescheduling for the crew planning. In particular, the minimum number of conductors needed on a train depends on the types and amount of rolling stock assigned to the train.

With this tool the decision makers could thus easily inspect the fine details of the solutions, as well as evaluate potential savings, service in the form of whether

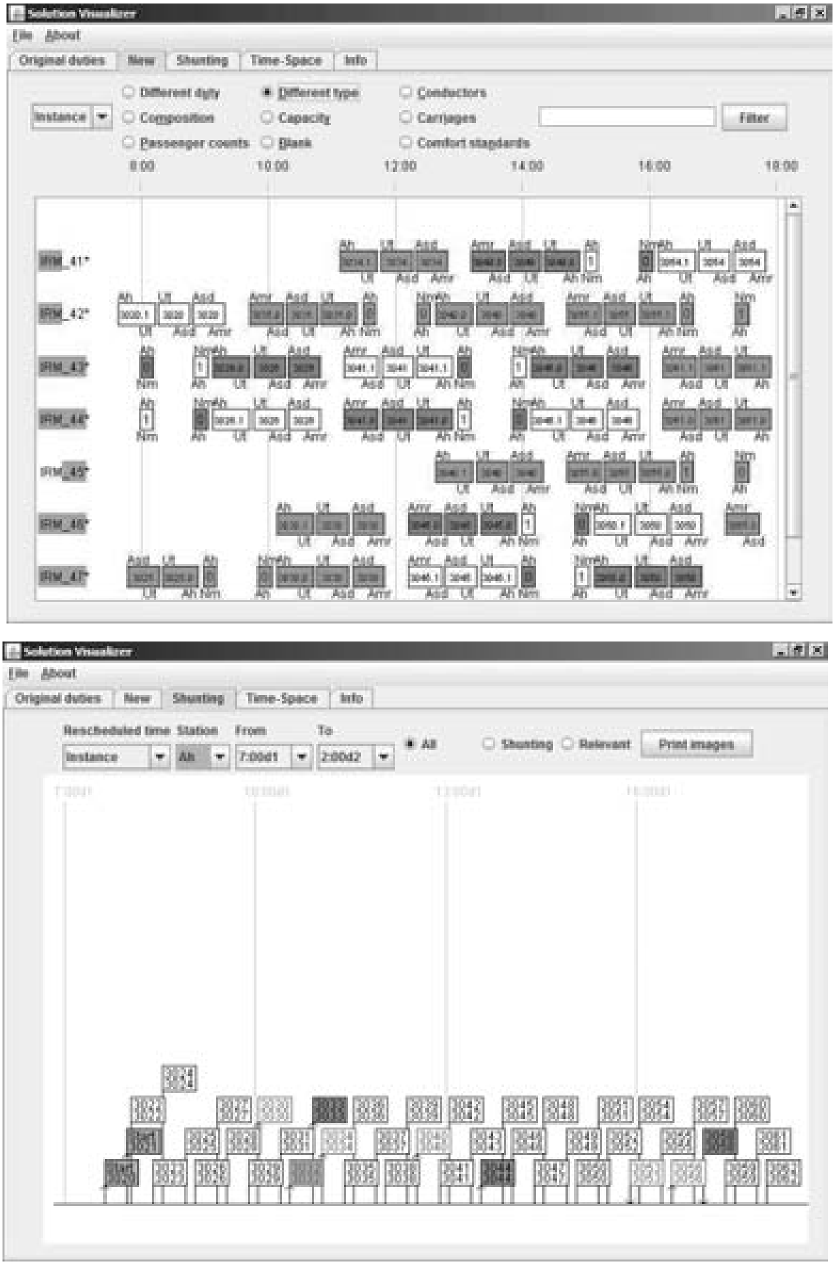


Figure 6.3: Screenshots from the visualization tool of TAM. *Top*: A set of rolling stock duties where certain colors indicate changes to the original duties. *Bottom*: Time line showing the changes to the local shunting operations.

capacity meets demand, changes to local operations, consequences for crew planning, and the need for repositioning trips to deal with off-balances.

Multiple solutions

Some of the objectives in the process are difficult to quantify such as the cost of introducing a new shunting operation at a specific connection. In particular, it was difficult for the decision makers to weigh the conflicting objectives against each other; for example when using a given set of parameters one new shunting operation can be saved by running a number of extra carriage kilometers.

We therefore decided to create multiple solutions by varying the parameters in the objective function, and to visualize the solutions simultaneously. This allowed the decision makers to explore the trade-off between the involved objectives effectively. Especially it made the cost of new shunting operations apparent in terms of carriage kilometers, seat shortages and other objectives.

6.4 Computational results

In both cases we found an appropriate set of parameters that resulted in acceptable schedules. We here describe the parameters used and discuss the resulting schedules.

6.4.1 Zomerplan 2009

For the Zomerplan we tested several sets of parameters. The decision makers finally chose to implement the schedules resulting from the parameters in Table 6.3. In this set of parameters off-balances are penalized severely as well as carriage kilometers. These penalties represent the desire to create a set of schedules that prioritize low operational costs.

The instances of the Zomerplan 2009 case were solved with the commercial MIP solver CPLEX 11.1 on a standard desktop PC. The running time of the solver was a few seconds on each instance which allowed us to quickly compute alternative schedules when undesirable features occurred.

The schedules for the Zomerplan were finally accepted by the decision makers. After thorough analysis they decided to make two manual changes to the plans before implementing them; the capacity on a single train in the morning peak hours on the 2600 line between The Hague (Gvc) and Amsterdam (Asd) was increased to ensure some slack capacity on that particular train. Furthermore, the uncoupling and later

Parameter	Description	Value
v	carriage kilometer	0.6
w	seat shortage kilometer	0.12
$\gamma_{c,q}$	new shunting	20
	canceled shunting	5
	changed type of operation	10
	changed type/number of units	5
$d^{0,+}_{s,m}$, $d^{0,-}_{s,m}$	beginning-of-day off-balance	200
$d^{\infty,+}_{s,m}$, $d^{\infty,-}_{s,m}$	end-of-day off-balance	200

Table 6.3: Parameter values used for the Zomerplan 2009.

coupling of a single unit in 's-Hertogenbosch (Ht) from a train on the 4400 line was canceled due to lacking shunting capacity in the station.

The final rolling stock schedules for the Zomerplan contained a significant reduction of carriage kilometer over the six week period compared to the generic schedules. These savings are in addition to the savings achieved by canceling a number of lines and reducing the frequency on some others. The reduction in carriage kilometers is equivalent to 14% (1.6 million) of the total number of carriage kilometers in the generic schedules for the involved lines.

6.4.2 The week 37 case

For the week 37 case the operational cost was much less pressing than in the Zomerplan. We therefore decided in agreement with the planners to set the cost of carriage kilometers $v = 0$ and thereby focus on the remaining objectives. The case was split into two instances – one for each day. We first created the schedule for Saturday with only a light penalty for evening off-balances and afterwards used the resulting evening balance as the target for Sunday morning with a heavy penalty on deviations from this target. Table 6.4 contains the final parameters used for the case.

We note that major changes to the shunting plans are also quite heavily penalized to reflect the objective of changing the local plans as little as possible. Off-balances and changes to the shunting plans are weighted against the third objective; meeting capacity demand. A rather light penalty was chosen for seat shortage kilometers to reflect their relative importance in this case.

Parameter	Description	Saturday	Sunday
v	carriage kilometer	0	0
w	seat shortage kilometer	0.025	0.025
$\gamma_{c,q}$	new shunting	100	100
	canceled shunting	15	5
	changed type of operation	100	100
	changed type/number of units	15	5
$d^{0,+}_{s,m}$, $d^{0,-}_{s,m}$	beginning-of-day off-balance	45	500
$d^{\infty,+}_{s,m}$, $d^{\infty,-}_{s,m}$	end-of-day off-balance	5	200

Table 6.4: Parameter values used for the week 37 case.

The final rolling stock schedules for the Week 37 case contained a number of carriage kilometers that was comparable to the original schedules. This was expected since saving carriage kilometers was not a priority in the computations. The solutions met the service requirements, which is considered very challenging in manual planning for the Noord-Oost lines. The shunting pattern at all involved stations was largely intact except at station Zwolle (Zl). The relatively large number of changed shunting operations at Zwolle is due to the fact that the pattern of connections was changed at that station.

6.4.3 Further results

In addition to the reductions in operational costs, the application of TAM has achieved a number of other advantages in the short-term planning process. Most importantly, the time it takes from the assessment of the rescheduling needs for a specific weekday until the time a modified rolling stock schedule is available is considerably reduced through the application of TAM. This is due to the short computation time of the underlying models and the fact that solutions can easily be modified or recomputed once the data is available. The reduction in production time thereby leads to an increase in planning capacity as manual short-term planning may take several days or even weeks to come up with a single feasible solution to a complex case.

An important remark for the application of TAM in the short-term planning department is that the tool quickly gained the acceptance of both the planners and the managers. In fact, during the development of TAM there evolved quite an aura of enthusiasm for the whole project among the practitioners. This enthusiasm was important when we needed input from the practitioners on various conceptual ideas for the models and the visualization.

Furthermore, the introduction of Tam in the short-term planning process has delivered useful insights into the nature of short-term planning problems. The insights involve (*i*) the trade-off between the different perspectives of the objective, (*ii*) the structure of "good" solutions, and (*iii*) potential ways to improve the working methods of the short-term planning department.

For item (*i*) the ability of Tam to deliver several alternative solutions has improved the understanding of the trade-off between the conflicting terms in the objective. When planners evaluated the proposed solutions to the instances, they realized to what extend a reduction in operational costs decreases service quality or implies changes to local operations.

For item (*ii*) the solutions computed by Tam often contained certain patterns of shunting operations. Some of these patterns were easily accepted by local planners whereas others were more problematic. Especially the introduction of shunting operations in the peak hours led to some discussions with the practitioners. Also, the assignment of long compositions to late departing trains was frowned upon by crew planners since long trains require more conductors. The application of Tam thus led to a better understanding of patterns and structures in solutions that were likely to be accepted by crew planners and local planners.

For item (*iii*) the introduction of Tam may change the way short-term planning is conducted. In the usual manual planning, the generic rolling stock schedule is usually adapted in a stepwise fashion by solving conflicts one by one. Tam on the other hand delivers a new solution that contains no conflicts and resembles the generic solution as much as possible. So rather than attempting to solve the puzzle of adapting the generic schedules, the planners now create a new solution and ensure that the new solution minimizes the implied changes to other planning processes.

6.5 Practical challenges

The collection and formatting of input data still constitutes a bottleneck in the application of Tam in short-term planning. This is because Tam relies on a fairly large amount of detailed data. However, at the time of writing, the further development of Tam has been taken over by the innovation department of NS and is being extended to take more details of short-term planning into account. This includes streamlining the process of collecting and formatting the significant amount of data. Also, attempts are made to combine other planning tools with TAM. These include in particular decision support tools for local planning since the verification of the modified rolling stock schedules at local planning poses a bottleneck in the current planning process.

Another practical aspect of the application of TAM is the fact that TAM, as it is currently implemented, requires a specialist to operate it. More specifically, the user is likely to need some knowledge of the underlying low-level implementation since all non-trivial extensions to the underlying models must be programmed as Java classes following a prescribed application programming interface. It will be a managerial question to decide whether TAM should be operated by specialists or by personnel with less technical insight. The last option requires the development of an appropriate graphical user interface that allows the user to utilize the full functionality of TAM without having to manipulate the implementation. If TAM eventually can be operated by non-specialists, it has the advantage that the experienced rolling stock planners can utilize TAM directly.

Currently, TAM is a stand-alone tool that is able to read and interpret data exported from the main information system of NS. Once TAM has computed an acceptable solution it can be written back into the information system. It is, however, the goal that TAM eventually becomes an integrated part of the main decision support and information system of NS. At the time of writing, a new information system for rolling stock scheduling called DONNA is being developed and adopted at NS. DONNA includes a graphical user interface and TAM is likely to be integrated into this system.

6.6 Discussion and conclusions

In this chapter we described the concrete application of the Rolling Stock Rescheduling Problem from Chapter 3 to a number of cases from short-term planning at NS. The models have been implemented in the planning tool TAM which also contains methods to visualize the solutions and key statistics as well as the changes from the generic schedules and any potential conflicts.

TAM has been successfully applied to two complex short-term planning cases; the Zomerplan 2009 and the Week 37 case. For the Zomerplan the solutions were implemented with only minor manual modifications and led to a decrease in carriage kilometers by 14% for the involved lines. For the Week 37 case the application of TAM resulted in a solution that was accepted by both the short-term planning department and local planners.

The experiences with the application of TAM in short-term planning raises a central question about the way rolling stock planning is conducted: *Are the generic schedules necessary?* It seems redundant to first create a detailed generic plan and then completely overhaul the plan to fit to the requirements of specific days. However, most weeks fit well to the specifications of the generic week plan and for those weeks it may be a good idea to have one highly optimized and well-tested plan. The generic plan serves this purpose well.

The use of a generic schedule is also pragmatic; manual short-term planning has a relatively long throughput time and the generic plan is traditionally used as basis for the modified situation by stepwise resolving any conflicts. But TAM changes both of these aspects – it speeds up the process and it creates a new plan rather than updating the existing one. It is therefore the opinion of the author that the overall rolling stock planning process could become more flexible if the generic plan is less binding.

Another issue is that in manual short-term planning the planners generally operate with just one rolling stock schedule. It starts out as the generic plan and changes to the plan are performed iteratively, but at any point in time the plan with the changes is considered the current plan. TAM can potentially change this manner of working by offering several alternative solutions to the planning problem.

Lastly, we note that TAM has been applied to the Zomerplan 2010 for the 12 week period from June 14 to September 5. This year it was decided to apply TAM to the rolling stock assignment of all intercity lines except those that use locomotive-hauled carriages. Furthermore, TAM was combined with a prototype tool for predicting capacity demand based on historical data (see Hoogenraad et al. (2010)). The application of TAM for the Zomerplan 2010 resulted in total savings of 4.1 million carriage kilometers compared to the generic plan. In addition, NS could operate their intercity services with 8 rolling stock units less during the summer period.

Chapter 7

Conclusions

In this thesis we investigated Operations Research models for rolling stock rescheduling in passenger railways. The models have been applied to a number of problems in disruption management and short-term planning at the major Dutch passenger railway operator NS. In Chapter 2 we gave an overview of the planning process at a passenger railway operator along with a review of relevant literature. In Chapter 3 we introduced the Rolling Stock Rescheduling Problem and developed models for the problem. In Chapters 4 to 6 we applied the rolling stock rescheduling models to problems in disruption management and in short-term planning.

7.1 Research questions

The main research question of this thesis reads as follows:

> *How can rolling stock rescheduling be modeled and how can the models be applied in the decision making process in a passenger railway system?*

We dissected the main question into six sub-questions. The first question concerns the characterization of rescheduling problems and the identification of rescheduling problems in the planning and operational process at NS.

1. *Which specific problems in the planning and operations of a passenger railway system can be characterized as rolling stock rescheduling problems?*

In Chapter 1 we characterized a rescheduling problem as the problem of adapting an existing solution to a modified situation. In the context of rolling stock in passenger railways it is the problem of adapting the rolling stock schedules to a modified timetable due to blocked infrastructure, unavailability of rolling stock units, or a significant change in demand.

Two specific application areas of rolling stock rescheduling were identified at NS: The first is in disruption management while the second is in the planning of rolling stock in the short-term planning phase.

The real time scheduling of rolling stock in disruption management is indeed an application of rolling stock rescheduling because the problems considered exist in a time-critical environment and usually concern only a subset of the system.

The rolling stock scheduling performed in the short-term planning phase was also characterized as an application of rolling stock rescheduling. Here the existing generic rolling stock schedule is adapted to specific calendar days with modified demand, particular rolling stock availability, or unavailable infrastructure due to maintenance projects.

The second question concerns the managerial goals associated with rolling stock rescheduling.

2. *Which aspects characterize the problems in rolling stock rescheduling in terms of managerial goals?*

In Chapter 2 we identified a number of aspects of the overall managerial goal and classified the aspects into four categories: Service, efficiency, system and process. The service category concerns passenger related aspects such as reliability and punctuality of the train services, but also providing sufficient seat capacity and providing adequate information to the passengers in case of disruptions. The efficiency category covers operational costs such as fuel, salaries and maintenance. The system category relates to the robustness of the system, i.e. the ability to absorb small disturbances, and the ability to recover from disruptions. The process category involves the late planning stages and the actual operations where the stability of the system is a major concern.

The identification of the rescheduling problems led to the third question.

3. *How can the specific problems that arise in the rolling stock rescheduling context be modeled and solved?*

In Chapter 3 we investigated several models for rolling stock rescheduling. We divided the problem into two sub-problems – circulation generation and duty generation – and provided models for both steps. The models are all MIP models that are solved using commercial-strength MIP software. Extensive computational tests revealed that the Composition Model combined with the Duty Path Model generally provides the best results in the least running time.

The fourth question concerns the modeling of the managerial goals.

4. *How can the trade-off between the different aspects of the managerial goals be modeled?*

In Chapter 3 we showed how several service and efficiency related aspects can be taken into account directly in the rescheduling models. Furthermore, we argued that the process perspectives could be handled by penalizing certain features in the models. These features primarily concern implied changes to the shunting operations as such changes may destabilize the system.

In Chapter 4 we studied a rolling horizon approach for rolling stock rescheduling in disruption management. The study included an investigation of the trade-off between efficiency, system performance, and computation time.

We further extended the focus on the service perspective in Chapter 5 by accounting for the dynamics of passenger flows during disruptions. In this study we include the inconvenience caused by insufficient capacities on trains, and again we trade off several conflicting objectives.

In Chapter 6 we considered the trade-off between several conflicting goals in practice when contructing rolling stock schedules for two complex short-term planning cases. We constructed a set of solutions by varying the parameters in the models and let experienced planners evaluate which solutions to use.

The fifth question relates to how uncertainty can be accounted for in rolling stock rescheduling in disruption management.

5. *How can the inherent uncertainty of disruptions be incorporated in the rescheduling process?*

We suggested three ways to take the uncertainty into account: (*i*) the rolling horizon approach studied in Chapter 4, (*ii*) the use of deviation costs applied in Chapters 3 to 5, and (*iii*) the ability to produce several alternative solutions through short computation times.

For (*i*) we specifically showed that the uncertainty of real time operations can be modeled as an online combinatorial optimization problem. The state of the system is gradually updated a number of times during the development of a disruption, and the decision maker does not have any knowledge of later updates. We used a rescheduling framework based on a rolling horizon to solve the online problem, and explored the trade-off between the involved objectives and computation time.

Item (*ii*) was in particular applied to potential changes to the shunting plans. In the real time context the exact local consequences of changing a planned shunting movement are often unclear, therefore we apply deviation costs in the models in Chapter 3. The deviation costs prevent the obtained solutions from incurring significant changes to the uncertain parallel processes.

With respect to (*iii*) we note that the short computation times experienced when employing the rolling horizon approach allow the user to generate and inspect several solutions. This opens the possibility for performing a what-if analysis, which again allows the decision maker to forecome some of the uncertainty of the situation.

The sixth question concerns the passengers.

6. *How can we improve the service quality for the passengers in the rescheduling process?*

We considered the interdependency of the rolling stock rescheduling and the flow of passengers. The interdependency is highly relevant since the passenger flow sets the demand for capacity on the trips. In Chapter 5 we incorporated the dynamics of passenger flows during disruptions by applying a simulation algorithm to estimate the inconvenience for the passengers given a certain rolling stock assignment. In particular, we studied the delays caused by insufficient capacities on trains and the trade-off between service and system objectives.

We found that explicitly accounting for passenger behavior may have added value for the service aspect of the rolling stock rescheduling – in particular in cases where passengers have several possible routes in the network to their destinations.

7.2 Contributions

For the planning process, the models developed in this thesis have already had a significant impact. In Chapter 6 we presented two case studies of the application of rolling stock rescheduling in the short-term planning phase. The rolling stock rescheduling models were implemented as the planning tool TAM. The planning tool successfully provided solutions to the two cases that were implemented in practice. The perspective of TAM is that it is likely to further lead to a revision of the entire rolling stock planning procedure; in the future there is going to be less early generic planning and more late adaptation to specific calender days. Planning the resources closer to the actual operations is only possible if a flexible planning tool like TAM is available.

Rolling stock rescheduling during disruptions is currently handled manually with limited computerized support. The computerized support generally consists of conflict detection and feasibility checks of the manually created solutions. The integration of a decision support tool into the disruption management systems would allow the user to respond more flexibly to disruptions. In particular, such an integration allows the decision maker not only to respond quickly, but also to estimate the consequences of several potential solutions before implementing one of them. The rolling stock rescheduling models suggested in this thesis can account for several aspects of the managerial goals. This allows the decision maker to explore the trade-off between the aspects before deciding which solutions to implement.

We believe there are promising perspectives in developing a decision support system with a built-in optimization module based on the rescheduling models presented in this thesis. In particular, the results for the rolling horizon framework in Chapter 4 indicate that the rescheduling methods can deal with the inherent uncertainty of disruptions.

For the theoretical implications, we note that we contribute to the understanding of uncertainty in real time optimization by developing a solution framework for disruption management of passenger railway rolling stock in Chapter 4. The framework is based on rescheduling the rolling stock in a rolling horizon manner and thereby gradually recovering the schedules. We have shown that the rolling horizon allows a trade-off between solution quality and computational time. Generally, using more computational time leads to solutions of higher quality, but there are no guarantees for that due to the uncertainty.

7.3 Future work

The results in this thesis suggest a number of applications of rolling stock rescheduling models in passenger railway operations. Still, there are several open questions with regards to both the practical and the theoretical implications of the models.

7.3.1 Practical challenges

The models developed in Chapter 3 rely on a fairly large number of parameters. These parameters set the exact weights of the involved objectives and thus control the trade-off between conflicting objectives. The exact setting of the parameters that reflects the wishes of the decision makers is still to be determined. Such fine-tuning requires testing the models with different parameters on a large number of realistic instances, and discussing the outcomes with both managers and planners.

Furthermore, it is a practical challenge to integrate the models into the existing information systems of the operator. Such an integration particularly concerns streamlining the data collection process.

For the disruption management framework based on rolling horizon rescheduling introduced in Chapter 4, we still need to verify the applicability of the models in real-time. For the rescheduling models embedded in the decision support tool Tam (see Chapter 6) we were able to confirm their practicality in a number of cases. But in the disruption management context a true verification of the rescheduling approach requires a better integration with the existing data management systems to access the necessary data in a timely manner. Developing the interfaces needed for such an integration is beyond the scope of an Operations Research based study, but we suggest first testing the methods in a simulated environment.

A significant challenge for the models with dynamic passenger flows presented in Chapter 5 is the validation of the methods in practice. Especially, the passenger behavior predicted by the models has to be compared to real life observations, thereafter the parameters of the models have to be calibrated accordingly. It is also a practical challenge to get the proper data for the models since they rely on quite detailed OD-information. For the data collection task it may be useful to study the data provided by the recently introduced electronic tickets in the Netherlands.

The rescheduling methods implemented in the decision support tool Tam have already proven their practical relevance for the short-term rescheduling process. However, there remain several practical challenges with respect to the application of the methods. The models rely on a large number of parameters which have to be fine-tuned, the tool has to be better integrated into the existing systems, the user interface could benefit from improvements, and it should be decided whether the tool should be operated by planners or Operations Research specialists. These questions should be addressed by the management of NS.

7.3.2 Areas for future research

In this thesis we have considered rolling stock rescheduling as an isolated resource allocation problem. However, as discussed in the introduction, rolling stock is only one of the three main resources to be rescheduled – the two others are the infrastructure and the crew. It could be interesting to investigate the possibilities of integrating the rescheduling of several of these resources in the process. Even if such an integration is not computationally tractable the study may still deliver insight into the restrictions of separate rescheduling processes, and into the interdependencies between the resources. First of all, it is an open question how to model the integration, and second, it is a further challenge how to deal with the integration in online rescheduling.

For the framework for rolling stock rescheduling in disruption management in Chapter 4 it could be interesting to study how several relevant aspects could be included in the framework. These aspects include maintenance of rolling stock units, rollng stock units with limited availability, and the stochasticity of disruptions. For the latter we note that it may be possible to exploit the knowledge of how disruptions usually evolve to develop stochastic models for rolling stock rescheduling.

The study of rolling stock rescheduling with passenger dynamics presented in Chapter 5 applies a specific set of assumptions on the passenger behavior. We claim that the framework could be utilized for studying the effects of providing better real-time information to passengers, or the effects of introducing seat reservation systems. Such studies would require changing the set of assumptions on passenger behavior. Furthermore, it could be interesting to include other realistic real time decisions in addition to capacity allocation in the model, like delay management and the possibility of adapting the stopping patterns of train services.

Further, it would be interesting to investigate better lower bounds on the service objective as described in Section 5.7.2. Better lower bounds are likely to be obtained by a cut-and-price approach combining column generation for the paths of passengers and row generation based on Benders decomposition to add valid cuts for the assignment of capacity.

Appendix A

Glossary

Carriage: An indivisible piece of rolling stock that can carry passengers.

Rolling stock unit: A piece of rolling stock consisting of several *carriages*. A unit has its own engines and can be operated individually.

Rolling stock type: A set of *rolling stock units* with the same technical specifications. During planning, units of the same type are usually considered interchangeable while during operations they are distinguished by their maintenance requirements.

Composition: An ordered combination of *rolling stock units* attached to each other.

Crew: The personnel that operates the trains. We distinguish two types of crew: drivers and conductors.

Train service: An announced movement of a train from one terminal station to another terminal station with a number of calls underway. A train service has a train number and fixed departure/arrival times at all involved stations.

Timetable: The set of all *train services*.

Trip: A part of a *train service* from a departure station to an arrival station where it is not possible to change the *composition* in between.

Connection: The planned continuation of zero, one or two incoming *trips* on zero, one or two outgoing *trips*.

Splitting: A *connection* where the train assigned to an incoming *trip* is split into two trains assigned to two outgoing *trips*.

Combining: A *connection* where the trains assigned to two incoming *trips* are combined into a single train assigned to an outgoing *trip*.

Shunting: The movements of *rolling stock units* inside railway nodes. Shunting includes the *coupling* and *uncoupling* of *rolling stock units* to and from trains, as well as the storage of idle units in the shunt yard.

Shunting operation: The action of retrieving a *rolling stock unit* in the shunt yard and *coupling* it to a train, or *uncoupling* a unit from a train and sending it to the shunt yard.

Coupling: The act of attaching one or more *rolling stock units* to a train.

Uncoupling: The act of detaching one or more *rolling stock units* from a train.

Composition change: *Coupling* or *uncoupling* one or more *rolling stock units* during a connection.

Task: The job for a *rolling stock unit* of being assigned to a certain position in a *composition* on a *trip*.

Duty: The sequence of *tasks* to be performed by a single *rolling stock unit* for a certain planning period, usually a day.

Rolling stock circulation: The global assignment of *compositions* to *trips*.

Inventory: An abstraction of the storage of *rolling stock units* at the stations. The inventory keeps track of the number of *rolling stock units* available for coupling at any time during the planning period. When a unit is coupled to a train it is deducted from the inventory, and when a unit is uncoupled from a train it is added to the inventory.

Train line: A set of *train services* that call at a given set of stations.

Turning pattern: The matching of incoming *trips* with outgoing *trips* at a terminal station. During planning the turning pattern is often fixed whereas during operations the turning pattern may be more flexible.

Off-balance: A surplus of one *rolling stock unit* in the inventory of a station at the beginning or the end of the planning horizon. The surplus implies a deficit in the inventory of another station.

Repositioning trip: A train ride without passengers used for bringing rolling stock from one station to another. Repositioning trips are generally used to resolve off-balances.

Disturbance: A minor incident, such as a delay, that affects the operations of the railway system. Handling a disturbance usually requires little or no action from the operator.

Disruption: A major incident that affects the operations of the of the railway system. Disruptions require substantial deviations from the planned operations.

Appendix B

Notations

B.1 Notations in Chapter 3

B.1.1 Common parameters

$\mathcal{T}$	the set of trips in the modified timetable.
$\mathcal{C}$	the set of connections in the modified timetable.
$\pi(t)$	the predecessor connection of trip t.
$\sigma(t)$	the successor connection of trip t.
IN_c	the set of incoming trips in connection c.
OUT_c	the set of outgoing trips in connection c.
$\mathcal{M}$	the set of rolling stock types.
$\mathcal{S}$	the set of stations.
$\mathcal{P}$	the set of compositions.
$\nu_m(p)$	the number of rolling stock units of type m in composition p.
$\eta(t)$	the set of allowed compositions for trip t.
$\varrho(c)$	the set of allowed composition changes at connection c.
$p_{q,t}$	the composition assigned to incoming trip t in composition change q.
$p'_{q,t}$	the composition assigned to outgoing trip t in composition change q.
$\tau_a(c)$	the arrival time of the first incoming trip at connection c.
$\tau_d(c)$	the departure time of the last outgoing trip at connection c.
$\tau^+(c)$	the time at which rolling stock units are retrieved from inventory to be coupled at connection c.

$\tau^-(c)$ — the time at which rolling stock units uncoupled at connection c are available in the inventory.

$i^0_{s,m}$ — the number of rolling stock units of type m in the inventory at station s at the beginning of the planning horizon.

$i^\infty_{s,m}$ — the number of rolling stock units of type m in the inventory at station s at the end of the planning horizon.

$\texttt{km}(t)$ — the number of kilometers in trip t.

$\texttt{carr}(p)$ — the number of carriages in composition p.

$\texttt{cap}(p)$ — the seat capacity of composition p.

$\texttt{con}(p)$ — the number of conductors needed to serve composition p.

B.1.2 Notations for the Composition Model

Parameters

$\alpha_{q,m}$ — the number of rolling stock units of type m uncoupled in composition change q.

$\beta_{q,m}$ — the number of rolling stock units of type m coupled in composition change q.

Variables

$X_{t,p}$ — whether composition $p \in \eta(t)$ is used for trip $t \in \mathcal{T}$.

$Z_{c,q}$ — whether composition change $q \in \varrho(c)$ is used for connection $c \in \mathcal{C}$.

$I_{c,m}$ — number of units of type $m \in \mathcal{M}$ in the inventory at station $s(c)$ at time $\tau^+(c)$.

$I^0_{s,m}$ — number of units of type $m \in \mathcal{M}$ at station $s \in \mathcal{S}$ at the beginning of the planning period.

$I^\infty_{s,m}$ — number of units of type $m \in \mathcal{M}$ at station $s \in \mathcal{S}$ at the end of the planning period.

$C_{c,m}$ — number of units of type $m \in \mathcal{M}$ that are coupled during connection c.

$U_{c,m}$ — number of units of type $m \in \mathcal{M}$ that are uncoupled during connection c.

$D^{\infty,+}_{s,m}$ — number of units of type $m \in \mathcal{M}$ in excess at station $s \in \mathcal{S}$ at the end of the planning period.

$D^{\infty,-}_{s,m}$ — number of units of type $m \in \mathcal{M}$ in shortage at station $s \in \mathcal{S}$ at the end of the planning period.

$A^{\infty,+}_{s,M}$ number of units of type m in the set of units with similar characteristics M in excess at station $s \in \mathcal{S}$ at the end of the planning period.

$A^{\infty,-}_{s,M}$ number of units of type m in the set of units with similar characteristics M in shortage at station $s \in \mathcal{S}$ at the end of the planning period.

Parameters for the flexible turning patterns

$\tilde{\mathcal{C}}$ the set of connections in the modified timetable at the station at which flexible turning patterns are applied.

$\mathcal{T}_1$ the set of trips in the modified timetable arriving at the station at which flexible turning patterns are applied.

$\mathcal{T}_2$ the set of trips in the modified timetable departing from the station at which flexible turning patterns are applied.

$\mathrm{UN}_{p,t}$ the set of allowed compositions that may result from uncoupling zero or more units from the composition p at the arrival of trip t.

$\mathrm{CO}_{p,t}$ the set of compositions from which one can construct p by coupling zero or more units at the departure of trip t.

$\delta_m(p_1, p_2)$ the difference in the number of units of type m between composition p_1 and p_2.

$\mathtt{plat}(s)$ the number of available platforms at station s.

Variables for the flexible turning patterns

$Y_{t,p,p'}$ whether the train assigned to trip t arrives with composition p and leaves composition p' at the platform.

$W_{t,p,p'}$ whether the train assigned to trip t departs with composition p and takes composition p' from the platform.

$P_{c,p}$ number of trains with composition p that are parked at the platforms of station $s(c)$ at time $\tau_d(c)$ for connection c.

Parameters for keeping the current duty structure

$\mathcal{N}$ the set of uncoupling/coupling relations.

$r(n)$ the number of units involved in relation n.

$m(n)$ the type of unit involved in relation n.

$c_1(n)$ the connection at which the units of relation n are uncoupled.

$c_2(n)$ the connection at which the units of relation n are coupled.

Variables for keeping the current duty structure

S_n	number of units participating in uncoupling/coupling relation $n \in \mathcal{N}$.

Parameters for continuity

Strings	the set of strings of 0s and 1s up to the appropriate length.
Start	the set of strings that only contains 1s.
End	the set of strings with at least the required number of 1s.
Allow(q)	the set of pairs of strings (r, r') where string r is changed to r' through the composition change q.
len(p)	the number of units in composition p.
len(r)	the number of digits in the string r.

Variables for continuity

$H_{t_i,r}$	whether the string $r \in$ Strings is assigned to trip t_i.
$G_{c_i,r,r'}$	whether the string r is assigned to trip t_i and r' is assigned to trip t_{i+1}.

B.1.3 Notations for the Task Model

Parameters

tasks(t)	the set of possible tasks in the train that serves trip t.
pos(k)	the position of task k.
trip(k)	the trip of task k.
$\mathcal{R}(s)$	the set of relevant times at station s.
d_0	time instant that precedes all other relevant times.
d_∞	time instant that is after all other relevant times.
prev(d)	the previous relevant time at station s for $d \in \mathcal{R}(s)$.
next(d)	the next relevant time at station s for $d \in \mathcal{R}(s)$.
$H(c)$	the set of inequalities that describe the convex hull of the allowed settings of the variables involved in connection c.
$\alpha^s_{m,k,k'}$	the coefficient for variable $R_{m,k,k'}$ for inequality s.
$\beta^s_{m,k}$	the coefficient for variable $U_{m,k}$ for inequality s.
$\gamma^s_{m,k}$	the coefficient for variable $C_{m,k}$ for inequality s.

ζ_s	the right hand side of inequality s.

Variables

$T_{m,k}$	whether a rolling stock unit of type $m \in \mathcal{M}$ is assigned to task k for $\texttt{trip}(k) \in \mathcal{T}$.
$R_{m,k,k'}$	whether a rolling stock unit of type $m \in \mathcal{M}$ assigned to task k is transferred to task k' in a connection.
$U_{m,k}$	whether a unit of type $m \in \mathcal{M}$ is uncoupled after doing task k for $\texttt{trip}(k) \in \mathcal{T}$.
$C_{m,k}$	whether a unit of type $m \in \mathcal{M}$ is coupled to do task k for $\texttt{trip}(k) \in \mathcal{T}$.
$I_{s,m,d}$	number of units type $m \in \mathcal{M}$ are in the inventory at station $s \in \mathcal{S}$ from relevant time $d \in \mathcal{R}(s)\backslash\{d_\infty\}$ to the next relevant time $\texttt{next}(d) \in \mathcal{R}(s)$.
$X_{t,p}$	whether composition $p \in \eta(t)$ is used for trip $t \in \mathcal{T}$.
$Z_{c,q}$	whether composition change $q \in \varrho(c)$ is used for connection $c \in \mathcal{C}$.

Parameters for the reformulation of the Task Model

$a_{p,m,k}$	whether a unit of type m is assigned to position $\texttt{pos}(k)$ in composition p.
$b_{q,m,k,k'}$	whether a unit of type m is transferred from task k of incoming trip $\texttt{trip}(k)$ to task k' in outgoing trip $\texttt{trip}(k')$ in the composition change q.
$u_{q,m,k}$	whether a unit of type m is uncoupled from task k in incoming trip $\texttt{trip}(k)$ in the composition change q.
$g_{q,m,k}$	whether a unit of type m is coupled to task k in outgoing trip $\texttt{trip}(k)$ in the composition change q.

Variables for the reformulation of the Task Model

$X_{t,p}$	whether composition $p \in \eta(t)$ is used for trip $t \in \mathcal{T}$.
$Z_{c,q}$	whether composition change $q \in \varrho(c)$ is used for connection $c \in \mathcal{C}$.

B.1.4 Notations for duty generation

Parameters

$\mathcal{G}$	the duty graph.
$\mathcal{V}$	the vertex set of the duty graph. Vertices represent fundamental duties.
$\mathcal{A}$	the arc set of the duty graph. Arcs represent pairs of fundamental duties that can be performed after eachother.
$s(v)$	the station where the fundamental duty v starts.
$e(v)$	the station where the fundamental duty v ends.
$\delta(u, v)$	the time needed to uncouple the unit from the train at $e(u)$, send it to the inventory, retrieve it from the inventory and couple it to the train of v.
$\delta^{\mathtt{in}}(v)$	the set of incoming arcs at vertex v.
$\delta^{\mathtt{out}}(v)$	the set of outgoing arcs at vertex v.
$\mathcal{U}$	the set of original rolling stock duties.
$h_{v,p}$	denotes whether vertex v is contained in path p.
$s_{u,p}$	the similarity between duties u and p.

Variables for the Duty Flow Model

K_a	whether vertices u and v are performed in succession by the same duty for arc $a = (u, v) \in \mathcal{A}$.
B_v	whether vertex $v \in \mathcal{V}$ is the first vertex in a duty.
E_v	whether vertex $v \in \mathcal{V}$ is the last vertex in a duty.

Variables for the Duty Path Model

$Q_{u,p}$	whether original duty $u \in \mathcal{U}$ is assigned to path $p \in \mathcal{P}$ in the modified situation.

Variables for the Duty-Task Model

$J_{m,k,k'}$	whether a unit of type $m \in \mathcal{M}$ performs task k, is uncoupled and then later coupled to task k'.

$J^0_{m,k}$	whether a unit of type m is coupled to task k without being assigned to any trip earlier.
$J^\infty_{m,k}$	whether a unit of type m is uncoupled from task k and is not used in any later tasks.
$J^{\texttt{idle}}_{s,m}$	the number of units of type m that are idle during the entire planning period at station s.

B.2 Notations in Chapter 4

Notations for the rolling horizon framework

$\langle t_i, \mathcal{S}_i \rangle$	the scenario $\mathcal{S}_i$ is available at time t_i.
h	horizon parameter denoting how far ahead to take the current information into account.
p	update parameter denoting how often the circulation should be updated.
a	the time at which intermediate inventories start to be taken into account.
b	the time at which intermediate inventories are taken into account with full weight.
$\varrho(t)$	parameter in $[0, 1]$ denoting the relative weight of off-balances at time t.

B.3 Notations in Chapter 5

Notations for rolling stock rescheduling with dynamic passenger flows

G	the passenger graph.
V	the node set of the passenger graph where $(s, \tau) \in V$ denotes the departure/arrival of a train at time τ at station s.
A	the arc set of the passenger graph.
$\mathcal{P}$	the set of passenger groups.
n_p	the size of passenger group p.
τ_p	the time at which the passenger group p enters the system.
o_p	the origin station of passenger group p.
d_p	the destination station of passenger group p.

$\tilde{t}_p$	the deadline of passenger group p.
S_p	the traveling strategy of passenger group p.
$\texttt{Paths}(G)$	the set of paths in the passenger graph $\mathcal{G}$.
$f(a,p)$	the flow of passenger group p on arc a.
$F(t,c)$	the feedback mechanism penalty for assigning capacity c to trip t.
$\texttt{TRAVEL}(p,a)$	average inconvenience per passenger from group p traveling on a.
$\texttt{REJECT}(p,a)$	average inconvenience per passenger from group p rejected upon boarding a.
$\texttt{AVG\_COST}(a)$	estimated marginal contribution to the cost of the passenger flow by decreasing the capacity of arc a.
α	parameter denoting how much we rely on the latest feedback information.

Parameters for the lower bounds

π_p	the set of all possible traveling paths for passenger group p.
c_p, q	the inconvenience for one passenger from passenger gorup p when traveling by path q.
$\texttt{cap}_{\texttt{max}}(t)$	the maximum capacity on trip t in any rolling stock schedule.

Variables for the lower bounds

$Y_{p,q}$	number of passengers from group p traveling along path q.

Appendix C

Computational Results for Rolling Horizon Approach

This appendix contains the computational results for the other instances that were mentioned in Section 4.7.1. Also, it contains a short study of the trade-off between service quality and efficiency in the horizon approach.

C.1 Horizon parameters

In addition to the instances used in Section 4.7 we tested the rolling horizon approach described in Chapter 4 on three other sets of instances. The first set involves the 3000 line (see Figure 4.1) with a disruption between Schagen and Alkmaar at time 12:00. The disruption lasts 4 hours. The set has 20 instances. The second set involves the 3000 line with a disruption between Utrecht and Arnhem at time 19:00. The disruption lasts 3 hours. The set has 16 instances. The third set involves the Noord-Oost lines with a disruption between Utrecht and Amersfoort at time 19:30. The disruption lasts 2 hours. The set has 10 instances.

Figures C.1, C.2 and C.3 show the objective values of the instances and the trade-off between off-balances and introducing new shunting operations. We observe that setting a to around 18:00 to 19:00 yields the best trade-off for the 3000 line. For the disruptions that take place in the evening we notice that there is no point in starting the balancing process earlier since there are no off-balances in the intermediate inventories before the disruption.

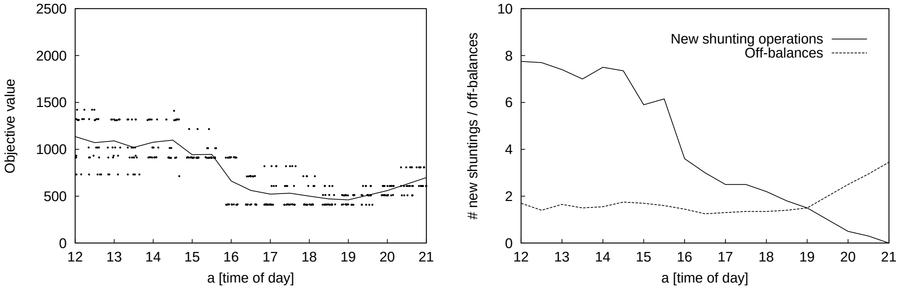


Figure C.1: *Left*: Objective function of each instance for the disruption at 12:00 on the 3000 line as a function of parameter a. The line represents the average. *Right*: Average number of new shunting operations and off-balances.

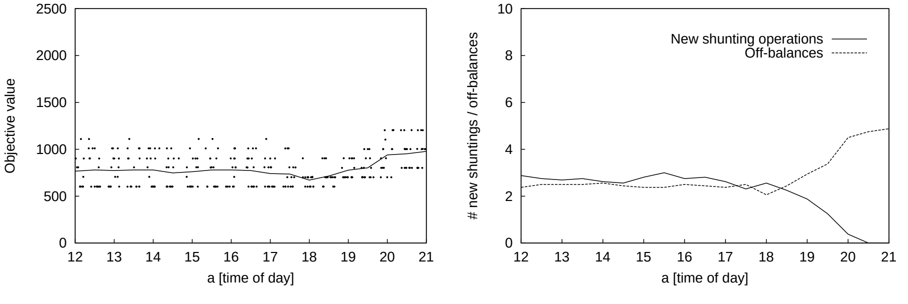


Figure C.2: *Left*: Objective function of each instance for the disruption at 19:00 on the 3000 line as a function of parameter a. The line represents the average. *Right*: Average number of new shunting operations and off-balances.

Objective	Cost
Cancel trip	10,000
Off-balance	20
New shunting operation	10
Minor shunting changes	1, 2 or 5
Seat shortage kilometer	0.0004–0.0025
Carriage kilometer	0.01

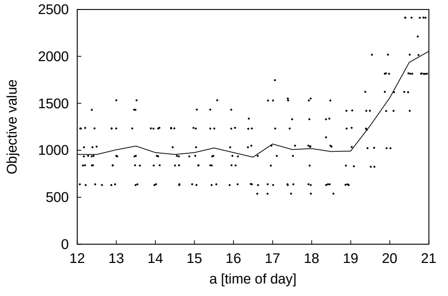
Table C.1: Objective parameters with service and efficiency aspects.

Figure C.4 shows the relationship between the objective value and the horizon length h. Generally, longer horizons result in better solutions, although for horizons longer than 3:30 hours the added value seems to diminish.

C.2 Impact on service quality and efficiency

To further explore the effects of rescheduling on service quality and efficiency, we used the objective weights in Table C.1: Introducing a new shunting operation costs 10, minor changes to shunting operations cost 1, 2 or 5 depending on their nature. Off-balances cost 20. A carriage kilometer costs 0.01, while we use varying seat shortage kilometers costs. The expected passenger counts are based on the undisrupted situation which means that the impact of the disruption on the passenger flows is not taken into account. Canceling a trip costs 10,000. This set of weights balances the practical aspects of the process with the service and efficiency perspectives since assigning too little capacity to trains is penalized as well as assigning long trains to low demand trips.

We conducted experiments with a horizon of $h = 3$ hours with rescheduling every $p = 1$ hours. Seat shortage costs between 0.0004 and 0.0025 are used while all other cost coefficients are held constant. The top diagram in Figure C.5 shows the relationship between the resulting number of seat shortage kilometers and carriage kilometers averaged over all instances. We note a clear trade-off between carriage kilometers and seat shortage kilometers. The originally planned rolling stock circulation used 369,751 carriage kilometers of which 4,255 are canceled directly due to the disruption. At the same time the original plan had 52,553 seat shortage kilometers. Depending on the parameter settings the rescheduled plan achieves a somewhat



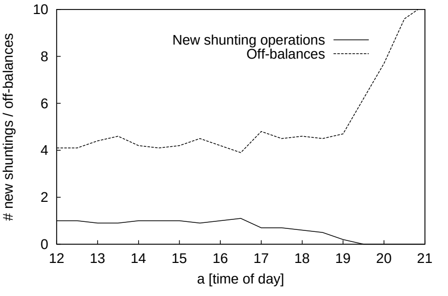


Figure C.3: *Left*: Objective function of each instance for the disruption at 19:30 on the Noord-Oost lines as a function of parameter a. The line represents the average. *Right*: Average number of new shunting operations and off-balances.

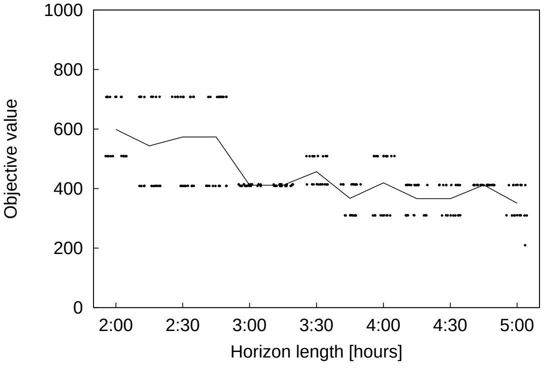


Figure C.4: Objective function of each instance for the disruption at 12:00 on the 3000 line as a function of parameter h. The line represents the average.

lower number of carriage kilometers with at least the double number of seat shortage kilometers.

The bottom diagram in Figure C.5 shows the trade-off between seat shortage kilometers and another characteristic of the solution: the changes to the shunting plans. Again we note that providing better service, in the sense of limiting seat shortages, comes at a cost of some other undesired features like more changes to the shunting plans.

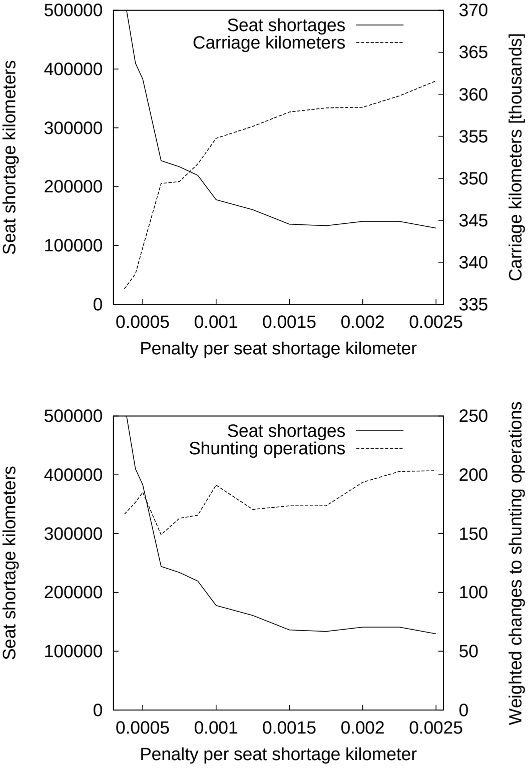


Figure C.5: *Top*: Seat shortage kilometers and carriage kilometers for different values of the seat shortage penalty. *Bottom*: Seat shortage kilometers and the weighted number of changes to the shunting operations.

Bibliography

ABARA, J. (1989). *Applying Integer Linear Programming to the Fleet Assignment Problem*. Interfaces, 19(4):20–28.

ABBINK, E.J.W., FISCHETTI, M., KROON, L.G., TIMMER, G., VROMANS, M.J.C.M. (2005). *Reinventing Crew Scheduling at Netherlands Railways*. Interfaces, 35(5):393–401.

ABBINK, E.J.W., MOBACH, D.G.A., FIOOLE, P.J., KROON, L.G., VAN DER HEIJDEN, E.H.T., WIJNGAARDS, N.J.E. (2009). *Actor-Agent Application for Train Driver Rescheduling*. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 513–520. International Foundation for Autonomous Agents and Multiagent Systems.

ABBINK, E.J.W, VAN DEN BERG, B.W.V., KROON, L.G., SALOMON, M. (2004). *Allocation of Railway Rolling Stock for Passenger Trains*. Transportation Science, 38(1):33–42.

AHUJA, R., MAGNANTI, T., ORLIN, J. (1993). *Network Flows: Theory, Algorithms and Applications*. Prentice-Hall, Englewood Cliffs, NJ.

AHUJA, R.K., LIU, J., ORLIN, J.B., SHARMA, D., SHUGHART, L.A. (2005). *Solving Real-life Locomotive Scheduling problems*. Transportation Science, 39(4):503–517.

ALFIERI, A., GROOT, R., KROON, L.G., SCHRIJVER, A. (2006). *Efficient Circulation of Railway Rolling Stock*. Transportation Science, 40(3):378–391.

ASSAD, A.A. (1980). *Models for Rail Transportation*. Transportation Research A, 14(3):205–220.

Ball, M., Barnhart, C., Nemhauser, G., Odoni, A. (2007). *Air Transportation: Irregular Operations and Control*, volume 14 of *Handbooks in Operations Research and Management Science*, chapter 1. North-Holland.

Bard, J.F., Yu, G., Argüello, M.F. (2001). *Optimizing Aircraft Routings in response to Groundings and Delays*. IIE Transactions, 33(10):931–947.

Ben-Khedher, N., Kintanar, J., Queille, C., Stripling, W. (1998). *Schedule Optimization at SNCF: From Conception to Day of Departure*. Interfaces, 28(1):6–23.

Ben-Tal, A., El Ghaoui, L., Nemirovski, A. (2009). *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press.

Billionnet, A. (2003). *Using Integer Programming to Solve the Train-Platforming Problem*. Transportation Science, 37(2):213–222.

Bodin, L., Golden, B. (1981). *Classification in Vehicle Routing and Scheduling*. Networks, 11(2):97–108.

Borndörfer, R., Grotschel, M., Pfetsch, M.E. (2008). *Models for Line Planning in Public Transport*. In *Computer Aided Scheduling of Public Transport*, Lecture Notes in Economics and Mathematical Systems, pages 363–378. Springer.

Borodin, A., El-Yaniv, R. (1998). *Online Computation and Competitive Analysis*. Cambridge University Press.

Brännlund, U., Lindberg, P.O., Nõu, A., Nilsson, J.E. (1998). *Railway Timetabling Using Lagrangian Relaxation*. Transportation Science, 32(4):358–369.

Bratu, S., Barnhart, C. (2006). *Flight Operations Recovery: New Approaches Considering Passenger Recovery*. Journal of Scheduling, 9(3):279–298.

Brucker, P., Hurink, J., Rolfes, T. (2003). *Routing of Railway Carriages*. Journal of Global Optimization, 27(2):313–332.

Budai, G., Maróti, G., Dekker, R., Huisman, D., Kroon, L.G. (2009). *Rescheduling in Passenger Railways: the Rolling Stock Rebalancing Problem*. Journal of Scheduling, 13(3):1–17.

Bunte, S., Kliewer, N. (2009). *An Overview on Vehicle Scheduling Models*. Public Transport, 1(4):299–317.

Bussieck, M.R., Lindner, T., Lübbecke, M.E. (2004). *A Fast Algorithm for Near Cost Optimal Line Plans*. Mathematical Methods of Operations Research, 59(2):205–220.

Cacchiani, V., Caprara, A., Galli, L., Kroon, L.G., Maróti, G., Toth, P. (2008a). *Price of Recoverability for Railway Rolling Stock Planning*. Technical Report ARRIVAL-TR-0121, ARRIVAL project.

Cacchiani, V., Caprara, A., Kroon, L.G., Maróti, G., Toth, P. (2008b). *Move-up Train Units*. Technical Report ARRIVAL-TR-0122, ARRIVAL project.

Cacchiani, V., Caprara, A., Toth, P. (2010a). *Scheduling Extra Freight Trains on Railway Networks*. Transportation Research Part B: Methodological, 44(2):215–231.

——— (2010b). *Solving a Real-World Train-Unit Assignment Problem*. Mathematical Programming, 124(1):207–231.

Caimi, G., Fuchsberger, M., Laumanns, M., Schüpbach, K. (2007). *Periodic Railway Timetabling with Event Flexibility*. In Liebchen, C., Ahuja, R.K., Mesa, J.A. (editors), *ATMOS 2007*.

Caprara, A., Fischetti, M., Guida, P.L., Toth, P., Vigo, D. (1999a). *Solution of Large Scale Railway Crew Planning Problems: the Italian Experience*. In Wilson, N.H.M. (editor), *Computer Aided Transit Scheduling*. Springer-Verlag.

Caprara, A., Fischetti, M., Toth, P. (1999b). *A Heuristic Method for the Set Covering Problem*. Operations Research, 47(5):730–743.

——— (2002). *Modeling and Solving the Train Timetabling Problem*. Operations Research, 50(5):851–861.

Caprara, A., Galli, L., Toth, P. (2007a). *Solution of the Train Platforming Problem*. In Liebchen, C., Ahuja, R.K., Mesa, J.A. (editors), *ATMOS 2007*.

Caprara, A., Kroon, L.G., Monaci, M., Peeters, M., Toth, P. (2007b). *Passenger Railway Optimization*. In Barnhart, C., Laporte, G. (editors), *Handbook in OR & MS vol. 14*, pages 129–187. Elsevier.

Carey, M., Carville, S. (2003). *Scheduling and Platforming Trains at Busy Complex Stations*. Transportation Research Part A: Policy and Practice, 37(3):195–224.

Christof, T., Löbel, A. (2009). *PORTA software*. `http://www.iwr.uni-heidelberg.de/groups/comopt/software/PORTA`.

Clausen, J., Larsen, A., Larsen, J., Rezanova, N.J. (2010). *Disruption Management in the Airline Industry – Concepts, Models and Methods*. Computers & Operations Research, 37(5):809–821.

Cordeau, J.F., Desaulniers, G., Lingaya, N., Soumis, F., Desrosiers, J. (2001a). *Simultaneous Locomotive and Car Assignment at VIA Rail Canada*. Transportation Research Part B: Methodological, 35(8):767–787.

Cordeau, J.F., Soumis, F., Desrosiers, J. (2000). *A Benders Decomposition Approach for the Locomotive and Car Assignment Problem*. Transportation Science, 34(2):133–149.

——— (2001b). *Simultaneous Assignment of Locomotives and Cars to Passenger Trains*. Operations Research, 49(4):531–548.

Cordeau, J.F., Toth, P., Vigo, D. (1998). *A Survey of Optimization Models for Train Routing and Scheduling*. Transportation Science, 32(4):380–404.

Corman, F., D'Ariano, A., Pacciarelli, D., Pranzo, M. (2010). *A Tabu Search Algorithm for Rerouting Trains during Rail Operations*. Transportation Research Part B: Methodological, 44(1):175–192.

Cornelsen, S., Di Stefano, G. (2007). *Track Assignment*. Journal of Discrete Algorithms, 5(2):250–261.

D'Ariano, A. (2008). *Improving Real-Time Train Dispatching: Models, Algorithms and Applications*. Ph.D. thesis, Delft University of Technology, The Netherlands.

De Almeida, D., Marcos, N., Vianey, S., Weber, C. (2008). *Robustness of Rolling Stock Schedules*. Technical Report ARRIVAL-TR-0129, ARRIVAL project.

Di Stefano, G., Koči, M.L. (2004). *A Graph Theoretical Approach to the Shunting Problem*. Electronic Notes in Theoretical Computer Science, 92:16–33.

Dumas, J., Aithnard, F., Soumis, F. (2009). *Improving the Objective Function of the Fleet Assignment Problem*. Transportation Research Part B: Methodological, 43(4):466–475.

Dumas, J., Soumis, F. (2008). *Passenger Flow Model for Airline Networks*. Transportation Science, 42(2):197–207.

Filar, J.A., Manyem, P., White, K. (2001). *How Airlines and Airports Recover from Schedule Perturbations: A Survey*. Annals of Operations Research, 108(1):315–333.

Fioole, P.J., Kroon, L.G., Maróti, G., Schrijver, A. (2006). *A Rolling Stock Circulation Model for Combining and Splitting of Passenger Trains*. European Journal of Operational Research, 174(2):1281–1297.

Flier, A.G.H., Gaurav, A., Nunkesser, M. (2008). *Combinatorial Aspects of Move-up Crews*. In *Operations Research Proceedings 2008*. Springer.

Folkmann, M., Jespersen, J., Nielsen, M.N. (2007). *Estimates on Rolling Stock and Crew in DSB S-tog based on Timetables*. Lecture Notes in Computer Science, 4359:91–107.

Freling, R., Lentink, R.M., Kroon, L.G., Huisman, D. (2005). *Shunting of Passenger Train Units in a Railway Station*. Transportation Science, 39(2):261–272.

Fuchsberger, M. (2007). *Solving the Train Scheduling Problem in a Main Station Area via a Resource Constrained Space-time Integer Multi-commodity Flow*. Master's thesis, Institute for Operations Research, ETH Zürich, Switzerland.

Ghiani, G., Guerriero, F., Laporte, G., Musmanno, R. (2003). *Real-Time Vehicle Routing: Solution Concepts, Algorithms and Parallel Computing Strategies*. European Journal of Operational Research, 151(1):1–11.

GINKEL, A., SCHÖBEL, A. (2007). *To Wait or Not to Wait? The Bicriteria Delay Management Problem in Public Transportation*. Transportation Science, 41(4):527–538.

GIOVANNI, L.D., HEILPORN, G., LABBÉ, M. (2007). *Optimization Models for the Delay Management Problem in Public Transportation*. European Journal of Operational Research, 189(3):762–774.

GOOSSENS, J.W. (2004). *Models and Algorithms for Railway Line Planning Problems*. Ph.D. thesis, University of Maastricht, The Netherlands.

GOOSSENS, J.W., VAN HOESEL, S., KROON, L.G. (2004). *A Branch-and-Cut Approach for Solving Railway Line-planning Problems*. Transportation Science, 38(3):379–393.

GRÖTSCHEL, M., KRUMKE, S.O., RAMBAU, J., WINTER, T., ZIMMERMANN, U.T. (2001). *Combinatorial Online Optimization in Real Time*. In GRÖTSCHEL, M., KRUMKE, S.O., RAMBAU (editors), *Online Optimization of Large Scale Systems*, pages 679–704. Springer.

HANE, C.A., BARNHART, C., JOHNSON, E.L., MARSTEN, R.E., NEMHAUSER, G.L., SIGISMONDI, G. (1995). *The Fleet Assignment Problem: Solving a Large-Scale Integer Program*. Mathematical Programming, 70(1):211–232.

HOOGENRAAD, J., LENTINK, R., DE VRIES, B. (2010). *Hoeveel Reizigers Zitten er in de Trein?*. STAtOR, 11(1):8–10. In Dutch.

HUISMAN, D. (2007). *A Column Generation Approach for the Rail Crew Rescheduling Problem*. European Journal of Operational Research, 180(1):163–173.

HUISMAN, D., KROON, L.G., LENTINK, R.M., VROMANS, M. (2005). *Operations Research in Passenger Railway Transportation*. Statistica Neerlandica, 59(4):467–497.

JESPERSEN-GROTH, J., CLAUSEN, J. (2006). *Optimal Reinsertion of Cancelled Train Lines*. Technical report, Informatics and Mathematical Modelling, Technical University of Denmark.

Jespersen-Groth, J., Potthoff, D., Clausen, J., Huisman, D., Kroon, L.G., Maróti, G., Nielsen, M.N. (2010). *Disruption Management in Passenger Railway Transportation*. In Ahuja, R., Möhring, R., Zaroliagis, C. (editors), *Special Volume on Robust and Online Large-scale Optimization*. Springer.

Karlin, A.R., Manasse, M.S., Rudolph, L., Sleator, D.D. (1988). *Competitive Snoopy Caching*. Algorithmica, 3(1):79–119.

Kohl, N., Larsen, A., Larsen, J., Ross, A., Tiourine, S. (2007). *Airline disruption management – Perspectives, experiences and outlook*. Journal of Air Transport Management, 13(3):149–162.

Kouvelis, P., Yu, G. (1997). *Robust Discrete Optimization and Its Applications*. Kluwer Academic Publishers, Boston, MA.

Kroon, L.G., Fischetti, M. (2001). *Crew Scheduling for Netherlands Railways "Destination: Customer"*. In Voss, S., Daduna, J.R. (editors), *Computer-Aided Scheduling of Public Transport*, pages 181–201. Springer, Berlin, Germany.

Kroon, L.G., Huisman, D., Abbink, E., Fioole, P.J., Fischetti, M., Maróti, G., Schrijver, A., Steenbeek, A., Ybema, R. (2009). *The New Dutch Timetable: The OR Revolution*. Interfaces, 39(1):6–17.

Kroon, L.G., Lentink, R.M., Schrijver, A. (2008a). *Shunting of Passenger Train Units: An Integrated Approach*. Transportation Science, 42(4):436–449.

Kroon, L.G., Maróti, G. (2008). *Robust Train Routing*. Technical Report ARRIVAL-TR-0123, ARRIVAL project.

Kroon, L.G., Maróti, G., Helmrich, M.R., Vromans, M., Dekker, R. (2008b). *Stochastic Improvement of Cyclic Railway Timetables*. Transportation Research Part B: Methodological, 42(6):553–570.

Kroon, L.G., Peeters, L.W.P. (2003). *A Variable Trip Time Model for Cyclic Railway Timetabling*. Transportation Science, 37(2):198–212.

Kroon, L.G., Romeijn, H.E., Zwaneveld, P.J. (1997). *Routing Trains Through Railway Stations: Complexity Issues*. European Journal of Operational Research, 98(3):485–498.

LENTINK, R.M. (2006). *Algorithmic Decision Support for Shunt Planning*. Ph.D. thesis, Erasmus University, Rotterdam, The Netherlands.

LI, J.Q., MIRCHANDANI, P.B., BORENSTEIN, D. (2007). *The Vehicle Rescheduling Problem: Model and Algorithms*. Networks, 50(3):211–229.

LIEBCHEN, C. (2006). *Periodic Timetable Optimization in Public Transport*. Ph.D. thesis, Technische Universität Berlin, Germany.

LINGAYA, N., CORDEAU, J.F., DESAULNIERS, G., DESROSIERS, J., SOUMIS, F. (2002). *Operational Car Assignment at VIA Rail Canada*. Transportation Research Part B: Methodological, 36(9):755–778.

MARÓTI, G. (2006). *Operations Research Models for Railway Rolling Stock Planning*. Ph.D. thesis, Eindhoven University of Technology, Eindhoven, The Netherlands.

MARÓTI, G., KROON, L.G. (2005). *Maintenance Routing for Train Units: The Transition Model*. Transportation Science, 39(4):518–525.

——— (2007). *Maintenance Routing for Train Units: the Interchange Model*. Computers & Operations Research, 34(4):1121–1140.

MELLOULI, T., SUHL, L. (2007). *Rotation Planning of Locomotive and Carriage Groups with Shared Capacities*. In GERAETS, F., KROON, L.G., SCHÖBEL, A., WAGNER, D., ZAROLIAGIS, C. (editors), *Algorithmic Methods for Railway Optimization*, volume 4359 of *Lecture Notes in Computer Science*, pages 276–294. Springer.

MORGADO, E.M., MARTINS, J.P. (1998). *CREWS_NS: Scheduling Train Crews in The Netherlands*. AI Magazine, 19(1):25–38.

NIELSEN, L.K. (2010). *Short-term Rolling Stock Scheduling at a Major Railway Operator*. ORbit, 16:32–35.

NIELSEN, L.K., KROON, L.G., MARÓTI, G. (2010). *A Rolling Horizon Approach for Disruption Management of Railway Rolling Stock*. Submitted to *European Journal of Operational Research.*

NIELSEN, L.K., MARÓTI, G. (2009). *Disruption Management of Rolling Stock in Passenger Railway Transportation*. Technical Report ERS-2009-046-LIS, Erasmus Research Institute of Management (ERIM). Awarded the 2009 EURO Management Science Strategic Innovation Prize.

OPPENHEIM, N. (1995). *Urban Travel Demand Modeling: From Individual Choices to General Equilibrium*. Wiley-Interscience.

ORLIN, J. (1988). *A Faster Strongly Polynomial Minimum Cost Flow Algorithm*. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 377–387. ACM.

PEETERS, L.W.P., KROON, L.G. (2001). *A Cycle Based Optimization Model for the Cyclic Railway Timetabling Problem*. In *Computer Aided Scheduling of Public Transport*, pages 275–296. Springer.

PEETERS, M., KROON, L.G. (2008). *Circulation of Railway Rolling Stock: a Branch-and-Price Approach*. Computers & Operations Research, 35(2):538–556.

POTTHOFF, D., HUISMAN, D., DESAULNIERS, G. (2009). *Column Generation with Dynamic Duty Selection for Railway Crew Rescheduling*. Technical Report ARRIVAL-TR-0209, ARRIVAL Project. To appear in *Transportation Science*.

POTTER, S. (2003). *Transport Energy and Emissions: Urban Public Transport*, chapter 13, pages 247–262. Handbook of Transport and the Environment. Emerald Group Pub Ltd.

POTTHOFF, D. (2010). *Railway Crew Rescheduling: Novel Approaches and Extensions*. Ph.D. thesis, Erasmus University, Rotterdam, The Netherlands.

REZANOVA, N.J., RYAN, D.M. (2010). *The Train Driver Recovery Problem – A Set Partitioning Based Model and Solution Method*. Computers & Operations Research, 37(5):845–856.

ROSENBERGER, J.M., JOHNSON, E.L., NEMHAUSER, G.L. (2003). *Rerouting Aircraft for Airline Recovery*. Transportation Science, 37(4):408–421.

ROUILLON, S., DESAULNIERS, G., SOUMIS, F. (2006). *An Extended Branch-and-Bound Method for Locomotive Assignment*. Transportation Research Part B: Methodological, 40(5):404–423.

RUSZCZYŃSKI, A., SHAPIRO, A. (editors) (2003). *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*. Elsevier Science Ltd.

SCHACHTEBECK, M. (2010). *Delay Management in Public Transportation: Capacities, Robustness, and Integration*. Ph.D. thesis, Universität Göttingen, Germany.

SCHÖBEL, A., SCHOLL, S. (2006). *Line Planning with Minimal Transfers*. In *5th workshop on algorithmic methods and models for optimization of railways*, Dagstuhl seminar proceedings, vol. 06901.

SCHOLL, S. (2005). *Customer-Oriented Line Planning*. Ph.D. thesis, Kaiserslautern University of Technology, Germany.

SCHRIJVER, A. (1998). *Routing and Timetabling by Topological Search*. International Congress of Mathematicians, 3:687–695.

SCHRIJVER, A., STEENBEEK, A. (1994). *Dienstregelingontwikkeling voor RailNed (Timetable Construction for RailNed)*. Technical report, Center for Mathematics and Computer Science, Amsterdam, The Netherlands. In Dutch.

SÉGUIN, R., POTVIN, J.Y., GENDREAU, M., CRAINIC, T.G., MARCOTTE, P. (1997). *Real-Time Decision Problems: An Operational Research Perspective*. Journal of the Operational Research Society, 48(2):162–174.

SERAFINI, P., UKOVICH, W. (1989). *A Mathematical Model for Periodic Scheduling Problems*. SIAM Journal on Discrete Mathematics, 2(4):550–581.

SHAPIRO, R.J., HASSETT, K.A., ARNOLD, F.S. (2002). *Conserving Energy and Preserving the Environment: The Role of Public Transportation*. `http://www.apta.com/gap/policyresearch/Documents/shapiro.pdf`.

SHEBALOV, S., KLABJAN, D. (2006). *Robust Airline Crew Pairing: Move-up Crews*. Transportation Science, 40(3):300–312.

SISCOG (2010). *CREWS software by Siscog*. `http://www.siscog.eu`.

TEODOROVIC, D., STOJKOVIC, G. (1990). *Model for Operational Daily Airline Scheduling*. Transportation Planning and Technology, 14(4):273–285.

THENGVALL, B.G., BARD, J.F., YU, G. (2000). *Balancing User Preferences for Aircraft Schedule Recovery During Irregular Operations*. IIE Transactions, 32(3):181–193.

Thengvall, B.G., Yu, G., Bard, J.F. (2001). *Multiple Fleet Aircraft Schedule Recovery Following Hub Closures*. Transportation Research Part A: Policy and Practice, 35(4):289–308.

Törnquist, J. (2005). *Computer-Based Decision Support for Railway Traffic Scheduling and Dispatching: A Review of Models and Algorithms*. In Kroon, L.G., Möhring, R.H. (editors), *ATMOS 2005 (Algorithmic MeThods and Models for Optimization of RailwayS), Palma de Mallorca, Spain, October*.

——— (2007). *Railway Traffic Disturbance Management – An Experimental Analysis of Disturbance Complexity, Management Objectives and Limitations in Planning Horizon*. Transportation Research Part A: Policy and Practice, 41(3):249–266.

Vaidyanathan, B., Ahuja, R.K., Liu, J., Shughart, L.A. (2008). *Real-life Locomotive Planning: New Formulations and Computational Results*. Transportation Research Part B: Methodological, 42(2):147–168.

Veelenturf, L.P., Potthoff, D., Huisman, D., Kroon, L.G. (2010). *Railway Crew Rescheduling with Retiming*. Econometric Institute Report EI 2009-24, Erasmus University, Rotterdam, Econometric Institute.

Vromans, M. (2005). *Reliability of Railway Systems*. Ph.D. thesis, Erasmus University, Rotterdam, The Netherlands.

Walker, C.G., Snowdon, J.N., Ryan, D.M. (2005). *Simultaneous Disruption Recovery of a Train Timetable and Crew Roster in Real Time*. Computers & Operations Research, 32(8):2077–2094.

Wets, R.J.-B. (2002). *Stochastic Programming Models: Wait-and-See Versus Here-and-Now*. In Greengard, C., Ruszczyński, A. (editors), *Decision Making Under Uncertainty*. Springer.

Yan, S., Lin, C.G. (1997). *Airline Scheduling for the Temporary Closure of Airports*. Transportation Science, 31(1):72–82.

Yan, S., Tang, C., Shieh, C.-L. (2005). *A Simulation Framework for Evaluating Airline Temporary Schedule Adjustments Following Incidents*. Transportation Planning and Technology, 28(3):189–211.

Yu, G., Qi, X. (2004). *Disruption Management: Framework, Models and Applications*. World Scientific.

Zwaneveld, P.J., Kroon, L.G., Romeijn, H.E., Salomon, M., Dauzère-Pérès, S., van Hoesel, S.P.M., Ambergen, H.W. (1996). *Routing Trains Through Railway Stations: Model Formulation and Algorithms*. Transportation Science, 30(3):181–194.

Zwaneveld, P.J., Kroon, L.G., van Hoesel, S.P.M. (2001). *Routing Trains Through a Railway Station Based on a Node Packing Model*. European Journal of Operational Research, 128(1):14–33.

About the Author

Lars Kjær Nielsen was born on September 30th 1979 in Herning, Denmark. He obtained his pre-university education at the Nordfyns Gymnasium, Søndersø, in 1999 before serving 8 months military service from 1999 to 2000. From 2000 Lars studied computer science and mathematics at the University of Southern Denmark, Odense, and obtained his Master's degree in early 2006. The Master's project was conducted in cooperation with the railway operator DSB S-tog A/S. This cooperation led to a four month position as an analyst at DSB S-tog A/S in Copenhagen where he applied the methods developed in the Master's thesis to real life data.

In September 2006 Lars moved to Rotterdam to become a PhD candidate at the Department of Decision and Information Sciences of the Rotterdam School of Management, Erasmus University. During this period he attended conferences such as INFORMS Annual Meeting in Washington D.C. 2008, 23rd European Conference on Operational Research in Bonn 2009, and TRISTAN VII in Tromsø 2010.

Lars was awarded the first prize in the 2008 INFORMS Railway Applications Section Student Paper Award and together with Gábor Maróti he was awarded the prestigious EURO Management Science Strategic Innovation Prize 2009. At the time of writing, a paper based on the results of this thesis is in review at the *European Journal of Operational Research*.

Summary

Rescheduling is the action of adapting a schedule to a modified situation. In this thesis we study the rescheduling of passenger railway rolling stock in disruption management and in the short-term planning phase. All examples and test instances are based on the major Dutch railway operator Nederlandse Spoorwegen.

The overall goal of the thesis is to identify processes in the operational procedures that can be improved by computerized algorithmic support, and develop Operations Research models for these tasks.

In Chapter 1 we introduce the topic of rolling stock rescheduling in passenger railways and set up the research questions for the thesis. We elaborate on the contributions and the methodology which is based on Operations Research techniques.

In Chapter 2 we describe the planning and operations of a passenger railway system. The planning of the operations is divided into five phases: Strategic, tactical, short-term, daily, and real-time planning. Each phase has a shorter time horizon than the previous and concerns decisions on finer details of the eventual operations. In short-term planning the generic rolling stock schedules from the tactical planning phase are adapted to specific calendar days. This is an application of rolling stock rescheduling since the problem is to adapt an existing plan rather than construct a new one from scratch. Similarly, rolling stock rescheduling occurs during disruption management where the rolling stock schedules are modified to a disruption. A further challenge in the disruption management process is that the problem exists in a time-critical and uncertain environment.

The operational process consists of monitoring the timetable and the resources, detecting upcoming conflicts, and reacting to the conflicts. We describe the actors, and their responsibilities and options in the process. For the rolling stock in particular, the dispatchers have several options for rescheduling the duties. The dispatcher may exchange duties between rolling stock units, or uncouple or couple units from

the planned trains in order to assign the capacity to where it is needed. Also, the dispatcher may exchange entire train compositions at terminal stations and reposition rolling stock by inserting empty trains between stations. These options all affect the locally planned shunting operations.

In Chapter 3 we formalize the Rolling Stock Rescheduling Problem as the problem of adapting a set of rolling stock duties to a modified situation. The modified situation concerns an updated timetable, a change in rolling stock availability, or a change in demand.

We suggest a solution method based on a two-step approach; first we generate a new circulation of the rolling stock, i.e. an assignment of rolling stock types to trips and composition changes to connections. Second, we translate the circulation to a set of rolling stock duties.

For the circulation generation phase we suggest two models, the Composition Model and the Task Model. They are MIP models based on integer multi commodity flows in two different networks. We solve the models using standard MIP solving software. We also present two different MIP models for the duty generation phase, and perform computational comparisons of the different models on several realistic instances. We find that the Composition Model together with the Duty Path Model performs best for the Rolling Stock Rescheduling Problem. The two models in tandem allow the decision maker to trade off several objectives by varying the weights of the involved parameters.

Disruptions in a railway system may be caused by several external or internal factors. The impact and duration of a disruption is generally quite uncertain when the disruption occurs, and the development of the situation is therefore monitored closely. In Chapter 4 we develop a framework for disruption management of rolling stock based on an online combinatorial optimization problem: The state of the system is described by a certain scenario, and at any time, the scenario may be updated based on the development of the situation.

We propose a solution methodology based on rescheduling with a rolling horizon. The rolling stock is rescheduled for all trips within a certain time horizon and the plans are then revised periodically as updates to the scenario become available. The solution method allows the decision maker to trade off the involved objectives and even implicitly account for decisions that are out of the scope at the time of rescheduling. Computational tests show promising results on a number of instances from NS intercity lines.

The MIP models developed so far all assume that capacity demand is given as input. However passengers react to a disruption by choosing new routes in the network in both time and space. In Chapter 5 we consider the dynamics of passenger flows and incorporate the passenger behavior by applying an iterative solution procedure. The procedure combines the optimization of the rolling stock assignment with detailed simulation of the passenger flow.

We test the iterative solution approach on large realistic instances and show how the approach is able to decrease the inconvenience for the passengers at minor operational costs for the operator. Generally, the resulting solutions assign extra capacity to the trains on the rerouting possibilities around the disrupted area which allows more passengers to get to their destination in a timely manner.

We had the unique opportunity to solve real-life rolling stock rescheduling problems in cooperation with the short-term planning department at NS. In Chapter 6 we describe how we applied the rolling stock rescheduling models to two concrete cases and implemented the results in practice. One case concerns the summer schedules of 2009 where the capacity demands are lower than in generic weeks. The other case concerns a weekend with a major infrastructure maintenance project. In both cases the rescheduling models were able to provide multiple solutions in neglectable computation time. This again allowed the decision makers to explore the trade-off between the involved objectives, and implement the solutions that were most suitable in each case. The application of our rolling stock rescheduling models led to significant savings for the operational costs and increased planning capacity of the short-term planning department.

The optimization models for short-term planning are integrated into a planning tool where it is combined with a graphical user interface.

Samenvatting (Summary in Dutch)

In dit proefschrift bestuderen we de bijsturing van het materieel voor reizigerstreinen. Bijsturen is het op korte termijn, of zelfs direct, aanpassen van een plan naar aanleiding van een gewijzigde situatie. Alle voorbeelden en tests in dit proefschrift zijn gebaseerd op de situatie bij de Nederlandse Spoorwegen (NS).

De doelstelling van dit proefschrift is: (*i*) het bestuderen in hoeverre processen en procedures in materieelbijsturing op operationeel niveau profijt hebben van geautomatiseerde algoritmische ondersteuning en (*ii*) het ontwikkelen van besliskundige modellen en oplosmethoden hiervoor.

Hoofdstuk 1 vormt de inleiding tot de materieelbijsturing van reizigerstreinen en introduceert de onderzoeksvragen. We bespreken de bijdrage van dit onderzoek en de methodologie die gebaseerd is op de toepassing van besliskundige technieken.

In Hoofdstuk 2 beschrijven we de plannings- en bijsturingsprocessen die zich voordoen in een netwerk van reizigerstreinen. De planning is op te delen in vijf fasen, respectievelijk van lange naar korte planningshorizon zijn dat: Strategisch, tactisch, korte termijn, dagelijks, en real-time. Met de afname van de planningshorizon stijgt het detailniveau in de planning. In de korte termijn planning wordt de algemene materieelplanning aangepast aan de bijzondere omstandigheden van kalender dagen. Dit is een vorm van herplanning omdat het op basis van het algemene materieelplan gebeurt en er geen compleet nieuw plan gemaakt wordt. Ook tijdens een real-time verstoring is bijsturing van het materieel noodzakelijk. Naast het vinden van een aangepast materieelplan is een tweede uitdaging dit zo snel mogelijk te kunnen doen. Om de oplossing van praktische waarde te laten zijn, moet hierbij rekening gehouden worden met de onzekerheid in het proces.

Het bijsturingsproces bestaat uit het monitoren van de dienstregeling, het materieel en het personeel, het voorspellen en detecteren van conflicten, en het reageren op conflicten. We beschrijven de actoren en hun verantwoordelijkheden in het proces. Voor het herplannen van het materieel hebben de bijstuurders verschillende opties. Zo kan een bijstuurder taken tussen treinstellen uitwisselen of treinstellen van de geplande treinen aftrappen/bijplaatsen om meer of minder capaciteit toe te wijzen op de plekken waar dat nodig is. Een bijstuurder kan ook hele treinsamenstellingen uitwisselen bij het eindstation, en het materieel herpositioneren door nieuwe leeg materieel treinen in te plannen tussen stations. Deze bijsturingsmogelijkheden hebben alle gevolgen voor de lokaal geplande rangeerbewegingen.

In Hoofdstuk 3 formaliseren we het Materieel Bijsturings Probleem. Dit probleem bestaat uit het aanpassen van een set materieeldiensten aan een gewijzigde situatie. Dit kan gaan om een verandering in de dienstregeling, in de beschikbaarheid van materieel, of in de vraag naar capaciteit.

Wij stellen een oplossingsmethode voor op basis van een twee-staps benadering. In de eerste stap wordt de materieelomloop herpland, ofwel er wordt een koppeling gemaakt tussen materieeltypen en ritten, en de benodigde compositieveranderingen op de stations worden gegenereerd. In de tweede stap vertalen we de materieelomloop naar een set van materieeldiensten.

Voor het generen van de materieelomloop presenteren we twee Mixed Integer Programming (MIP) modellen: het Composition Model en het Task Model. Beide zijn gebasseerd op multi-commodity flows met verschillende onderliggende netwerken. We lossen de modellen op met behulp van standaard MIP software. Daarnaast definiëren we ook twee verschillende MIP modellen voor het genereren van de materieeldiensten. We voeren verschillende tests uit op basis van voorbeelden uit de praktijk om de modellen met elkaar te vergelijken. Uit de tests concluderen we dat het Composition Model samen met het Duty Path Model het best presteren voor het oplossen van het Materieel Bijsturings Probleem. De combinatie van de twee modellen geeft de gebruiker de mogelijkheid om de verschillende doelstellingen tegen elkaar af te wegen.

Verstoringen in een spoorwegsysteem kunnen worden veroorzaakt door verschillende externe en interne factoren. De impact en de duur van een verstoring is over het algemeen onzeker op het moment dat de verstoring optreedt. Daarom volgen de bijstuurders de ontwikkeling van de situatie op de voet. In Hoofdstuk 4 ontwikkelen we een kader voor de bijsturing van het materieel op basis van een online combinatorisch optimaliseringsprobleem: De toestand van het systeem wordt beschreven door een specifiek scenario en op elk tijdstip kan het scenario worden bijgewerkt gebaseerd op de ontwikkeling van de situatie.

Wij stellen een oplossingsmethode voor die gebaseerd is op bijsturing met een rollende tijdshorizon. Voor alle ritten binnen een bepaalde tijdshorizon wordt het materieel herpland en de plannen worden vervolgens periodiek herzien als scenario updates beschikbaar zijn. De oplossingsmethode geeft de bijstuurder de mogelijkheid de verschillende doelstellingen tegen elkaar af te wegen door er zelf gewichten aan toe te wijzen. Ook is het mogelijk om impliciet rekening te houden met beslissingen die buiten de huidige tijdshorizon vallen. De resultaten van de tests op reëele instanties van een aantal NS intercity lijnen zien er veelbelovend uit.

De tot nu toe beschreven MIP modellen gaan er van uit dat de vraag van de reizigers naar capaciteit gegeven is. Maar reizigers kunnen reageren op een verstoring door te kiezen voor een nieuwe route door het netwerk, in tijd en ruimte. In Hoofdstuk 5 beschouwen we de dynamiek van de reizigersstromen en nemen we het gedrag van reizigers mee door het toepassen van een iteratieve oplossingsprocedure. De procedure combineert het optimaliseren van de materieelomloop met een gedetailleerde simulatie van de reizigersstromen.

We testen de iteratieve oplossingsprocedure op een aantal realistische instanties en laten zien hoe de aanpak in staat is om het ongemak voor de reizigers te verkleinen terwijl de operationele kosten van de vervoerder laag blijven. In het algemeen is het zo dat de oplossingen van deze methode extra capaciteit toewijzen aan die treinen die betrokken zijn bij de omreismogelijkheden voor reizigers in het verstoorde gebied. Hierdoor komen meer reizigers eerder aan op hun bestemming.

Wij hadden de unieke kans om enkele echte materieel bijsturingsproblemen van de NS op te lossen in samenwerking met de korte termijn plannings afdeling van NS. In Hoofdstuk 6 beschrijven we hoe we de materieel herplanningsmodellen toegepast hebben op twee concrete gevallen en hoe we de resultaten in de praktijk hebben geïmplementeerd. De ene case betreft de zomer dienstregeling van 2009, waar de vraag naar capaciteit lager is dan in generieke weken. De andere case betreft een weekend met groot onderhoud aan het spoor. In beide gevallen waren de herplanningsmodellen in staat om in korte tijd verschillende oplossingen te berekenen. De verscheidenheid aan oplossingen zorgde ervoor dat de planners zelf in zeer korte tijd de trade-off tussen de verschillende doelstellingen konden zien en daaruit de beste oplossing konden kiezen en implementeren. De toepassing van onze materieel bijsturingsmodellen heeft geleid tot aanzienlijke besparingen in de operationele kosten, en tot een toename in de planningscapaciteit van de korte termijn planning afdeling.

De optimalisatie modellen voor de korte termijn planning worden op dit moment geïntegreerd in een planningstool waar het wordt gecombineerd met een grafische gebruikersinterface. Dit interface is belangrijk om de gebruiker te laten zien welke veranderingen in de materieelomloop zijn geïntroduceerd.

ERIM Ph.D. Series Research in Management

Erasmus Research Institute of Management (ERIM)

ERIM Ph.D. Series Research in Management

ERIM Electronic Series Portal: http://hdl.handle.net/1765/1

Acciaro, M., *Bundling Strategies in Global Supply Chains*, Promotor: Prof.dr. H.E. Haralambides, EPS-2010-197-LIS, ISBN: 978-90-5892-240-3, http://hdl.handle.net/1765/1

Agatz, N.A.H., *Demand Management in E-Fulfillment*, Promotor: Prof.dr.ir. J.A.E.E. van Nunen, EPS-2009-163-LIS, ISBN: 978-90-5892-200-7, http://hdl.handle.net/1765/15425

Alexiev, A., *Exploratory Innovation: The Role of Organizational and Top Management Team Social Capital*, Promotors: Prof.dr. F.A.J. van den Bosch & Prof.dr. H.W. Volberda, EPS-2010-208-STR, ISBN: 978-90-5892-249-6, http://hdl.handle.net/1765/1

Althuizen, N.A.P., *Analogical Reasoning as a Decision Support Principle for Weakly Structured Marketing Problems*, Promotor: Prof.dr.ir. B. Wierenga, EPS-2006-095-MKT, ISBN: 90-5892-129-8, http://hdl.handle.net/1765/8190

Alvarez, H.L., *Distributed Collaborative Learning Communities Enabled by Information Communication Technology*, Promotor: Prof.dr. K. Kumar, EPS-2006-080-LIS, ISBN: 90-5892-112-3, http://hdl.handle.net/1765/7830

Appelman, J.H., *Governance of Global Interorganizational Tourism Networks: Changing Forms of Co-ordination between the Travel Agency and Aviation Sector*, Promotors: Prof.dr. F.M. Go & Prof.dr. B. Nooteboom, EPS-2004-036-MKT, ISBN: 90-5892-060-7, http://hdl.handle.net/1765/1199

Asperen, E. van, *Essays on Port, Container, and Bulk Chemical Logistics Optimization*, Promotor: Prof.dr.ir. R. Dekker, EPS-2009-181-LIS, ISBN: 978-90-5892-222-9, http://hdl.handle.net/1765/1

Assem, M.J. van den, *Deal or No Deal? Decision Making under Risk in a Large-Stake TV Game Show and Related Experiments*, Promotor: Prof.dr. J. Spronk, EPS-2008-138-F&A, ISBN: 978-90-5892-173-4, http://hdl.handle.net/1765/13566

Baquero, G., *On Hedge Fund Performance, Capital Flows and Investor Psychology*, Promotor: Prof.dr. M.J.C.M. Verbeek, EPS-2006-094-F&A, ISBN: 90-5892-131-X, http://hdl.handle.net/1765/8192

Berens, G., *Corporate Branding: The Development of Corporate Associations and their Influence on Stakeholder Reactions*, Promotor: Prof.dr. C.B.M. van Riel, EPS-2004-039-ORG, ISBN: 90-5892-065-8, http://hdl.handle.net/1765/1273

Berghe, D.A.F. van den, *Working Across Borders: Multinational Enterprises and the Internationalization of Employment*, Promotors: Prof.dr. R.J.M. van Tulder & Prof.dr. E.J.J. Schenk, EPS-2003-029-ORG, ISBN: 90-5892-05-34, http://hdl.handle.net/1765/1041

Berghman, L.A., *Strategic Innovation Capacity: A Mixed Method Study on Deliberate Strategic Learning Mechanisms*, Promotor: Prof.dr. P. Mattyssens, EPS-2006-087-MKT, ISBN: 90-5892-120-4, http://hdl.handle.net/1765/7991

Bezemer, P.J., *Diffusion of Corporate Governance Beliefs: Board Independence and the Emergence of a Shareholder Value Orientation in the Netherlands*, Promotors: Prof.dr.ing. F.A.J. van den Bosch & Prof.dr. H.W. Volberda, EPS-2009-192-STR, ISBN: 978-90-5892-232-8, http://hdl.handle.net/1765/1

Bijman, W.J.J., *Essays on Agricultural Co-operatives: Governance Structure in Fruit and Vegetable Chains*, Promotor: Prof.dr. G.W.J. Hendrikse, EPS-2002-015-ORG, ISBN: 90-5892-024-0, http://hdl.handle.net/1765/867

Binken, J.L.G., *System Markets: Indirect Network Effects in Action, or Inaction?*, Promotor: Prof.dr. S. Stremersch, EPS-2010-213-MKT, ISBN: 978-90-5892-260-1, http://hdl.handle.net/1765/21186

Bispo, A., *Labour Market Segmentation: An investigation into the Dutch hospitality industry*, Promotors: Prof.dr. G.H.M. Evers & Prof.dr. A.R. Thurik, EPS-2007-108-ORG, ISBN: 90-5892-136-9, http://hdl.handle.net/1765/10283

Blindenbach-Driessen, F., *Innovation Management in Project-Based Firms*, Promotor: Prof.dr. S.L. van de Velde, EPS-2006-082-LIS, ISBN: 90-5892-110-7, http://hdl.handle.net/1765/7828

Boer, C.A., *Distributed Simulation in Industry*, Promotors: Prof.dr. A. de Bruin & Prof.dr.ir. A. Verbraeck, EPS-2005-065-LIS, ISBN: 90-5892-093-3, http://hdl.handle.net/1765/6925

Boer, N.I., *Knowledge Sharing within Organizations: A situated and Relational Perspective*, Promotor: Prof.dr. K. Kumar, EPS-2005-060-LIS, ISBN: 90-5892-086-0, http://hdl.handle.net/1765/6770

Boer-Sorbán, K., *Agent-Based Simulation of Financial Markets: A modular, Continuous-Time Approach*, Promotor: Prof.dr. A. de Bruin, EPS-2008-119-LIS, ISBN: 90-5892-155-0, http://hdl.handle.net/1765/10870

Boon, C.T., *HRM and Fit: Survival of the Fittest!?*, Promotors: Prof.dr. J. Paauwe & Prof.dr. D.N. den Hartog, EPS-2008-129-ORG, ISBN: 978-90-5892-162-8, http://hdl.handle.net/1765/12606

Borst, W.A.M., *Understanding Crowdsourcing: Effects of Motivation and Rewards on Participation and Performance in Voluntary Online Activities*, Promotors: Prof.dr.ir. J.C.M. van den Ende & Prof.dr.ir. H.W.G.M. van Heck, EPS-2010-221-LIS, ISBN: 978-90-5892-262-5, http://hdl.handle.net/1765/1

Braun, E., *City Marketing: Towards an Integrated Approach*, Promotor: Prof.dr. L. van den Berg, EPS-2008-142-MKT, ISBN: 978-90-5892-180-2, http://hdl.handle.net/1765/13694

Brito, M.P. de, *Managing Reverse Logistics or Reversing Logistics Management?*, Promotors: Prof.dr.ir. R. Dekker & Prof.dr. M. B. M. de Koster, EPS-2004-035-LIS, ISBN: 90-5892-058-5, http://hdl.handle.net/1765/1132

Brohm, R., *Polycentric Order in Organizations: A Dialogue between Michael Polanyi and IT-Consultants on Knowledge, Morality, and Organization*, Promotors: Prof.dr. G. W. J. Hendrikse & Prof.dr. H. K. Letiche, EPS-2005-063-ORG, ISBN: 90-5892-095-X, http://hdl.handle.net/1765/6911

Brumme, W.-H., *Manufacturing Capability Switching in the High-Tech Electronics Technology Life Cycle*, Promotors: Prof.dr.ir. J.A.E.E. van Nunen & Prof.dr.ir. L.N. Van Wassenhove, EPS-2008-126-LIS, ISBN: 978-90-5892-150-5, http://hdl.handle.net/1765/12103

Budiono, D.P., *The Analysis of Mutual Fund Performance: Evidence from U.S. Equity Mutual Funds*, Promotor: Prof.dr.ir. M.J.C.M. Verbeek, EPS-2010-185-F&A, ISBN: 978-90-5892-224-3, http://hdl.handle.net/1765/1

Burgers, J.H., *Managing Corporate Venturing: Multilevel Studies on Project Autonomy, Integration, Knowledge Relatedness, and Phases in the New Business Development Process*, Promotors: Prof.dr.ir. F.A.J. Van den Bosch & Prof.dr. H.W. Volberda, EPS-2008-136-STR, ISBN: 978-90-5892-174-1, http://hdl.handle.net/1765/13484

Campbell, R.A.J., *Rethinking Risk in International Financial Markets*, Promotor: Prof.dr. C.G. Koedijk, EPS-2001-005-F&A, ISBN: 90-5892-008-9, http://hdl.handle.net/1765/306

Carvalho de Mesquita Ferreira, L., *Attention Mosaics: Studies of Organizational Attention*, Promotors: Prof.dr. P.M.A.R. Heugens & Prof.dr. J. van Oosterhout, EPS-2010-205-ORG, ISBN: 978-90-5892-242-7, http://hdl.handle.net/1765/1

Chen, C.-M., *Evaluation and Design of Supply Chain Operations Using DEA*, Promotor: Prof.dr. J.A.E.E. van Nunen, EPS-2009-172-LIS, ISBN: 978-90-5892-209-0, http://hdl.handle.net/1765/1

Chen, H., *Individual Mobile Communication Services and Tariffs*, Promotor: Prof.dr. L.F.J.M. Pau, EPS-2008-123-LIS, ISBN: 90-5892-158-1, http://hdl.handle.net/1765/11141

Chen, Y., *Labour Flexibility in Chinas Companies: An Empirical Study*, Promotors: Prof.dr. A. Buitendam & Prof.dr. B. Krug, EPS-2001-006-ORG, ISBN: 90-5892-012-7, http://hdl.handle.net/1765/307

Damen, F.J.A., *Taking the Lead: The Role of Affect in Leadership Effectiveness*, Promotor: Prof.dr. D.L. van Knippenberg, EPS-2007-107-ORG, http://hdl.handle.net/1765/10282

Daniševská, P., *Empirical Studies on Financial Intermediation and Corporate Policies*, Promotor: Prof.dr. C.G. Koedijk, EPS-2004-044-F&A, ISBN: 90-5892-070-4, http://hdl.handle.net/1765/1518

Defilippi Angeldonis, E.F., *Access Regulation for Naturally Monopolistic Port Terminals: Lessons from Regulated Network Industries*, Promotor: Prof.dr. H.E. Haralambides, EPS-2010-204-LIS, ISBN: 978-90-5892-245-8, http://hdl.handle.net/1765/1

Delporte-Vermeiren, D.J.E., *Improving the Flexibility and Profitability of ICT-enabled Business Networks: An Assessment Method and Tool*, Promotor: Prof. mr. dr. P.H.M. Vervest & Prof.dr.ir. H.W.G.M. van Heck, EPS-2003-020-LIS, ISBN: 90-5892-040-2, http://hdl.handle.net/1765/359

Derwall, J.M.M., *The Economic Virtues of SRI and CSR*, Promotor: Prof.dr. C.G. Koedijk, EPS-2007-101-F&A, ISBN: 90-5892-132-8, http://hdl.handle.net/1765/8986

Diepen, M. van, *Dynamics and Competition in Charitable Giving*, Promotor: Prof.dr. Ph.H.B.F. Franses, EPS-2009-159-MKT, ISBN: 978-90-5892-188-8, http://hdl.handle.net/1765/14526

Dietvorst, R.C., *Neural Mechanisms Underlying Social Intelligence and Their Relationship with the Performance of Sales Managers*, Promotor: Prof.dr. W.J.M.I. Verbeke, EPS-2010-215-MKT, ISBN: 978-90-5892-257-1, http://hdl.handle.net/1765/21188

Dietz, H.M.S., *Managing (Sales)People towards Perfomance: HR Strategy, Leadership & Teamwork*, Promotor: Prof.dr. G.W.J. Hendrikse, EPS-2009-168-ORG, ISBN: 978-90-5892-210-6, http://hdl.handle.net/1765/1

Dijksterhuis, M., *Organizational Dynamics of Cognition and Action in the Changing Dutch and US Banking Industries*, Promotors: Prof.dr.ir. F.A.J. van den Bosch & Prof.dr. H.W. Volberda, EPS-2003-026-STR, ISBN: 90-5892-048-8, http://hdl.handle.net/1765/1037

Eijk, A.R. van der, *Behind Networks: Knowledge Transfer, Favor Exchange and Performance*, Promotors: Prof.dr. S.L. van de Velde & Prof.dr.drs. W.A. Dolfsma, EPS-2009-161-LIS, ISBN: 978-90-5892-190-1, http://hdl.handle.net/1765/14613

Elstak, M.N., *Flipping the Identity Coin: The Comparative Effect of Perceived, Projected and Desired Organizational Identity on Organizational Identification and Desired Behavior*, Promotor: Prof.dr. C.B.M. van Riel, EPS-2008-117-ORG, ISBN: 90-5892-148-2, http://hdl.handle.net/1765/10723

Erken, H.P.G., *Productivity, R&D and Entrepreneurship*, Promotor: Prof.dr. A.R. Thurik, EPS-2008-147-ORG, ISBN: 978-90-5892-179-6, http://hdl.handle.net/1765/14004

Fenema, P.C. van, *Coordination and Control of Globally Distributed Software Projects*, Promotor: Prof.dr. K. Kumar, EPS-2002-019-LIS, ISBN: 90-5892-030-5, `http://hdl.handle.net/1765/360`

Feng, L., *Motivation, Coordination and Cognition in Cooperatives*, Promotor: Prof.dr. G.W.J. Hendrikse, EPS-2010-220-ORG, ISBN: 90-5892-261-8, `http://hdl.handle.net/1765/21680`

Fleischmann, M., *Quantitative Models for Reverse Logistics*, Promotors: Prof.dr.ir. J.A.E.E. van Nunen & Prof.dr.ir. R. Dekker, EPS-2000-002-LIS, ISBN: 35-4041-711-7, `http://hdl.handle.net/1765/1044`

Flier, B., *Strategic Renewal of European Financial Incumbents: Coevolution of Environmental Selection, Institutional Effects, and Managerial Intentionality*, Promotors: Prof.dr.ir. F.A.J. van den Bosch & Prof.dr. H.W. Volberda, EPS-2003-033-STR, ISBN: 90-5892-055-0, `http://hdl.handle.net/1765/1071`

Fok, D., *Advanced Econometric Marketing Models*, Promotor: Prof.dr. Ph.H.B.F. Franses, EPS-2003-027-MKT, ISBN: 90-5892-049-6, `http://hdl.handle.net/1765/1035`

Ganzaroli, A., *Creating Trust between Local and Global Systems*, Promotors: Prof.dr. K. Kumar & Prof.dr. R.M. Lee, EPS-2002-018-LIS, ISBN: 90-5892-031-3, `http://hdl.handle.net/1765/361`

Gertsen, H.F.M., *Riding a Tiger without Being Eaten: How Companies and Analysts Tame Financial Restatements and Influence Corporate Reputation*, Promotor: Prof.dr. C.B.M. van Riel, EPS-2009-171-ORG, ISBN: 90-5892-214-4, `http://hdl.handle.net/1765/1`

Gilsing, V.A., *Exploration, Exploitation and Co-evolution in Innovation Networks*, Promotors: Prof.dr. B. Nooteboom & Prof.dr. J.P.M. Groenewegen, EPS-2003-032-ORG, ISBN: 90-5892-054-2, `http://hdl.handle.net/1765/1040`

Gijsbers, G.W., *Agricultural Innovation in Asia: Drivers, Paradigms and Performance*, Promotor: Prof.dr. R.J.M. van Tulder, EPS-2009-156-ORG, ISBN: 978-90-5892-191-8, `http://hdl.handle.net/1765/14524`

Gong, Y., *Stochastic Modelling and Analysis of Warehouse Operations*, Promotors: Prof.dr. M.B.M. de Koster & Prof.dr. S.L. van de Velde, EPS-2009-180-LIS, ISBN: 978-90-5892-219-9, `http://hdl.handle.net/1765/1`

Govers, R., *Virtual Tourism Destination Image: Glocal Identities Constructed, Perceived and Experienced*, Promotors: Prof.dr. F.M. Go & Prof.dr. K. Kumar, EPS-2005-069-MKT, ISBN: 90-5892-107-7, `http://hdl.handle.net/1765/6981`

Graaf, G. de, *Tractable Morality: Customer Discourses of Bankers, Veterinarians and Charity Workers*, Promotors: Prof.dr. F. Leijnse & Prof.dr. T. van Willigenburg, EPS-2003-031-ORG, ISBN: 90-5892-051-8, `http://hdl.handle.net/1765/1038`

Greeven, M.J., *Innovation in an Uncertain Institutional Environment: Private Software Entrepreneurs in Hangzhou, China*, Promotor: Prof.dr. B. Krug, EPS-2009-164-ORG, ISBN: 978-90-5892-202-1, `http://hdl.handle.net/1765/15426`

Groot, E.A. de, *Essays on Economic Cycles*, Promotors: Prof.dr. Ph.H.B.F. Franses & Prof.dr. H.R. Commandeur, EPS-2006-091-MKT, ISBN: 90-5892-123-9, http://hdl.handle.net/1765/8216

Guenster, N.K., *Investment Strategies Based on Social Responsibility and Bubbles*, Promotor: Prof.dr. C.G. Koedijk, EPS-2008-175-F&A, ISBN: 978-90-5892-206-9, http://hdl.handle.net/1765/1

Gutkowska, A.B., *Essays on the Dynamic Portfolio Choice*, Promotor: Prof.dr. A.C.F. Vorst, EPS-2006-085-F&A, ISBN: 90-5892-118-2, http://hdl.handle.net/1765/7994

Hagemeijer, R.E., *The Unmasking of the Other*, Promotors: Prof.dr. S.J. Magala & Prof.dr. H.K. Letiche, EPS-2005-068-ORG, ISBN: 90-5892-097-6, http://hdl.handle.net/1765/6963

Hakimi, N.A., *Leader Empowering Behaviour: The Leaders Perspective: Understanding the Motivation behind Leader Empowering Behaviour*, Promotor: Prof.dr. D.L. van Knippenberg, EPS-2010-184-ORG, http://hdl.handle.net/1765/1

Halderen, M.D. van, *Organizational Identity Expressiveness and Perception Management: Principles for Expressing the Organizational Identity in Order to Manage the Perceptions and Behavioral Reactions of External Stakeholders*, Promotor: Prof.dr. S.B.M. van Riel, EPS-2008-122-ORG, ISBN: 90-5892-153-6, http://hdl.handle.net/1765/10872

Hartigh, E. den, *Increasing Returns and Firm Performance: An Empirical Study*, Promotor: Prof.dr. H.R. Commandeur, EPS-2005-067-STR, ISBN: 90-5892-098-4, http://hdl.handle.net/1765/6939

Hensmans, M., *A Republican Settlement Theory of the Firm: Applied to Retail Banks in England and the Netherlands (1830-2007)*, Promotors: Prof.dr. A. Jolink& Prof.dr. S.J. Magala, EPS-2010-193-ORG, ISBN: 90-5892-235-9, http://hdl.handle.net/1765/1

Hermans, J.M., *ICT in Information Services; Use and Deployment of the Dutch Securities Trade, 1860-1970*, Promotor: Prof.dr. drs. F.H.A. Janszen, EPS-2004-046-ORG, ISBN: 90-5892-072-0, http://hdl.handle.net/1765/1793

Hernandez Mireles, C., *Marketing Modeling for New Products*, Promotor: Prof.dr. P.H. Franses, EPS-2010-202-MKT, ISBN: 90-5892-237-3, http://hdl.handle.net/1765/1

Hessels, S.J.A., *International Entrepreneurship: Value Creation Across National Borders*, Promotor: Prof.dr. A.R. Thurik, EPS-2008-144-ORG, ISBN: 978-90-5892-181-9, http://hdl.handle.net/1765/13942

Heugens, P.P.M.A.R., *Strategic Issues Management: Implications for Corporate Performance*, Promotors: Prof.dr.ir. F.A.J. van den Bosch & Prof.dr. C.B.M. van Riel, EPS-2001-007-STR, ISBN: 90-5892-009-9, http://hdl.handle.net/1765/358

Heuvel, W. van den, *The Economic Lot-Sizing Problem: New Results and Extensions*, Promotor: Prof.dr. A.P.L. Wagelmans, EPS-2006-093-LIS, ISBN: 90-5892-124-7, http://hdl.handle.net/1765/1805

Hoedemaekers, C.M.W., *Performance, Pinned down: A Lacanian Analysis of Subjectivity at Work*, Promotors: Prof.dr. S. Magala & Prof.dr. D.H. den Hartog, EPS-2008-121-ORG, ISBN: 90-5892-156-7, http://hdl.handle.net/1765/10871

Hooghiemstra, R., *The Construction of Reality: Cultural Differences in Self-serving Behaviour in Accounting Narratives*, Promotors: Prof.dr. L.G. van der Tas RA & Prof.dr. A.Th.H. Pruyn, EPS-2003-025-F&A, ISBN: 90-5892-047-X, http://hdl.handle.net/1765/871

Hu, Y., *Essays on the Governance of Agricultural Products: Cooperatives and Contract Farming*, Promotors: Prof.dr. G.W.J. Hendrkse & Prof.dr. B. Krug, EPS-2007-113-ORG, ISBN: 90-5892-145-1, http://hdl.handle.net/1765/10535

Huang, X., *An Analysis of Occupational Pension Provision: From Evaluation to Redesigh*, Promotors: Prof.dr. M.J.C.M. Verbeek & Prof.dr. R.J. Mahieu, EPS-2010-196-F&A, ISBN: 90-5892-239-7, http://hdl.handle.net/1765/1

Huij, J.J., *New Insights into Mutual Funds: Performance and Family Strategies*, Promotor: Prof.dr. M.C.J.M. Verbeek, EPS-2007-099-F&A, ISBN: 90-5892-134-4, http://hdl.handle.net/1765/9398

Huurman, C.I., *Dealing with Electricity Prices*, Promotor: Prof.dr. C.D. Koedijk, EPS-2007-098-F&A, ISBN: 90-5892-130-1, http://hdl.handle.net/1765/9399

Iastrebova, K., *Managers Information Overload: The Impact of Coping Strategies on Decision-Making Performance*, Promotor: Prof.dr. H.G. van Dissel, EPS-2006-077-LIS, ISBN: 90-5892-111-5, http://hdl.handle.net/1765/7329

Iwaarden, J.D. van, *Changing Quality Controls: The Effects of Increasing Product Variety and Shortening Product Life Cycles*, Promotors: Prof.dr. B.G. Dale & Prof.dr. A.R.T. Williams, EPS-2006-084-ORG, ISBN: 90-5892-117-4, http://hdl.handle.net/1765/7992

Jansen, J.J.P., *Ambidextrous Organizations*, Promotors: Prof.dr.ir. F.A.J. Van den Bosch & Prof.dr. H.W. Volberda, EPS-2005-055-STR, ISBN: 90-5892-081-X, http://hdl.handle.net/1765/6774

Jaspers, F.P.H., *Organizing Systemic Innovation*, Promotor: Prof.dr.ir. J.C.M. van den Ende, EPS-2009-160-ORG, ISBN: 978-90-5892-197-0, http://hdl.handle.net/1765/14974

Jennen, M.G.J., *Empirical Essays on Office Market Dynamics*, Promotors: Prof.dr. C.G. Koedijk & Prof.dr. D. Brounen, EPS-2008-140-F&A, ISBN: 978-90-5892-176-5, http://hdl.handle.net/1765/13692

Jiang, T., *Capital Structure Determinants and Governance Structure Variety in Franchising*, Promotors: Prof.dr. G. Hendrikse & Prof.dr. A. de Jong, EPS-2009-158-F&A, ISBN: 978-90-5892-199-4, http://hdl.handle.net/1765/14975

Jiao, T., *Essays in Financial Accounting*, Promotor: Prof.dr. G.M.H. Mertens, EPS-2009-176-F&A, ISBN: 978-90-5892-211-3, http://hdl.handle.net/1765/1

Jong, C. de, *Dealing with Derivatives: Studies on the Role, Informational Content and Pricing of Financial Derivatives*, Promotor: Prof.dr. C.G. Koedijk, EPS-2003-023-F&A, ISBN: 90-5892-043-7, http://hdl.handle.net/1765/1043

Kaa, G. van, *Standard Battles for Complex Systems: Empirical Research on the Home Network*, Promotors: Prof.dr.ir. J. van den Ende & Prof.dr.ir. H.W.G.M. van Heck, EPS-2009-166-ORG, ISBN: 978-90-5892-205-2, http://hdl.handle.net/1765/1

Kagie, M., *Advances in Online Shopping Interfaces: Product Catalog Maps and Recommender Systems*, Promotor: Prof.dr. P.J.F. Groenen, EPS-2010-195-MKT, ISBN: 978-90-5892-233-5, http://hdl.handle.net/1765/1

Keizer, A.B., *The Changing Logic of Japanese Employment Practices: A Firm-Level Analysis of Four Industries*, Promotors: Prof.dr. J.A. Stam & Prof.dr. J.P.M. Groenewegen, EPS-2005-057-ORG, ISBN: 90-5892-087-9, http://hdl.handle.net/1765/6667

Kijkuit, R.C., *Social Networks in the Front End: The Organizational Life of an Idea*, Promotor: Prof.dr. B. Nooteboom, EPS-2007-104-ORG, ISBN: 90-5892-137-6, http://hdl.handle.net/1765/10074

Kippers, J., *Empirical Studies on Cash Payments*, Promotor: Prof.dr. Ph.H.B.F. Franses, EPS-2004-043-F&A, ISBN: 90-5892-069-0, http://hdl.handle.net/1765/1520

Klein, M.H., *Poverty Alleviation through Sustainable Strategic Business Models: Essays on Poverty Alleviation as a Business Strategy*, Promotor: Prof.dr. H.R. Commandeur, EPS-2008-135-STR, ISBN: 978-90-5892-168-0, http://hdl.handle.net/1765/13482

Knapp, S., *The Econometrics of Maritime Safety: Recommendations to Enhance Safety at Sea*, Promotor: Prof.dr. Ph.H.B.F. Franses, EPS-2007-096-ORG, ISBN: 90-5892-127-1, http://hdl.handle.net/1765/7913

Kole, E., *On Crises, Crashes and Comovements*, Promotors: Prof.dr. C.G. Koedijk & Prof.dr. M.J.C.M. Verbeek, EPS-2006-083-F&A, ISBN: 90-5892-114-X, http://hdl.handle.net/1765/7829

Kooij-de Bode, J.M., *Distributed Information and Group Decision-Making: Effects of Diversity and Affect*, Promotor: Prof.dr. D.L. van Knippenberg, EPS-2007-115-ORG, http://hdl.handle.net/1765/10722

Koppius, O.R., *Information Architecture and Electronic Market Performance*, Promotors: Prof.dr. P.H.M. Vervest & Prof.dr.ir. H.W.G.M. van Heck, EPS-2002-013-LIS, ISBN: 90-5892-023-2, http://hdl.handle.net/1765/921

Kotlarsky, J., *Management of Globally Distributed Component-Based Software Development Projects*, Promotor: Prof.dr. K. Kumar, EPS-2005-059-LIS, ISBN: 90-5892-088-7, http://hdl.handle.net/1765/6772

Krauth, E.I., *Real-Time Planning Support: A Task-Technology Fit Perspective*, Promotors: Prof.dr. S.L. van de Velde & Prof.dr. J. van Hillegersberg, EPS-2008-155-LIS, ISBN: 978-90-5892-193-2, http://hdl.handle.net/1765/14521

Kuilman, J., *The Re-Emergence of Foreign Banks in Shanghai: An Ecological Analysis*, Promotor: Prof.dr. B. Krug, EPS-2005-066-ORG, ISBN: 90-5892-096-8, http://hdl.handle.net/1765/6926

Kwee, Z., *Investigating Three Key Principles of Sustained Strategic Renewal: A Longitudinal Study of Long-Lived Firms*, Promotors: Prof.dr.ir. F.A.J. Van den Bosch & Prof.dr. H.W. Volberda, EPS-2009-174-STR, ISBN: 90-5892-212-0, http://hdl.handle.net/1765/1

Langen, P.W. de, *The Performance of Seaport Clusters: A Framework to Analyze Cluster Performance and an Application to the Seaport Clusters of Durban, Rotterdam and the Lower Mississippi*, Promotor: Prof.dr. B. Nooteboom & Prof. drs. H.W.H. Welters, EPS-2004-034-LIS, ISBN: 90-5892-056-9, http://hdl.handle.net/1765/1133

Larco Martinelli, J.A., *Incorporating Worker-Specific Factors in Operations Management Models*, Promotors: Prof.dr.ir. J. Dul & Prof.dr. M.B.M. de Koster, EPS-2010-217-LIS, ISBN: 90-5892-263-2, http://hdl.handle.net/1765/21527

Le Anh, T., *Intelligent Control of Vehicle-Based Internal Transport Systems*, Promotors: Prof.dr. M.B.M. de Koster & Prof.dr.ir. R. Dekker, EPS-2005-051-LIS, ISBN: 90-5892-079-8, http://hdl.handle.net/1765/6554

Le-Duc, T., *Design and Control of Efficient Order Picking Processes*, Promotor: Prof.dr. M.B.M. de Koster, EPS-2005-064-LIS, ISBN: 90-5892-094-1, http://hdl.handle.net/1765/6910

Leeuwen, E.P. van, *Recovered-Resource Dependent Industries and the Strategic Renewal of Incumbent Firm: A Multi-Level Study of Recovered Resource Dependence Management and Strategic Renewal in the European Paper and Board Industry*, Promotors: Prof.dr.ir. F.A.J. Van den Bosch & Prof.dr. H.W. Volberda, EPS-2007-109-STR, ISBN: 90-5892-140-6, http://hdl.handle.net/1765/10183

Lentink, R.M., *Algorithmic Decision Support for Shunt Planning*, Promotors: Prof.dr. L.G. Kroon & Prof.dr.ir. J.A.E.E. van Nunen, EPS-2006-073-LIS, ISBN: 90-5892-104-2, http://hdl.handle.net/1765/7328

Li, T., *Informedness and Customer-Centric Revenue Management*, Promotors: Prof.dr. P.H.M. Vervest & Prof.dr.ir. H.W.G.M. van Heck, EPS-2009-146-LIS, ISBN: 978-90-5892-195-6, http://hdl.handle.net/1765/14525

Liang, G., *New Competition: Foreign Direct Investment and Industrial Development in China*, Promotor: Prof.dr. R.J.M. van Tulder, EPS-2004-047-ORG, ISBN: 90-5892-073-9, http://hdl.handle.net/1765/1795

Liere, D.W. van, *Network Horizon and the Dynamics of Network Positions: A Multi-Method Multi-Level Longitudinal Study of Interfirm Networks*, Promotor: Prof.dr. P.H.M. Vervest, EPS-2007-105-LIS, ISBN: 90-5892-139-0, http://hdl.handle.net/1765/10181

Loef, J., *Incongruity between Ads and Consumer Expectations of Advertising*, Promotors: Prof.dr. W.F. van Raaij & Prof.dr. G. Antonides, EPS-2002-017-MKT, ISBN: 90-5892-028-3, http://hdl.handle.net/1765/869

Londoño, M. del Pilar, *Institutional Arrangements that Affect Free Trade Agreements: Economic Rationality Versus Interest Groups*, Promotors: Prof.dr. H.E. Haralambides & Prof.dr. J.F. Francois, EPS-2006-078-LIS, ISBN: 90-5892-108-5, http://hdl.handle.net/1765/7578

Maas, A.A. van der,, *Strategy Implementation in a Small Island Context: An Integrative Framework*, Promotor: Prof.dr. H.G. van Dissel, EPS-2008-127-LIS, ISBN: 978-90-5892-160-4, http://hdl.handle.net/1765/12278

Maas, K.E.G., *Corporate Social Performance: From Output Measurement to Impact Measurement*, Promotor: Prof.dr. H.R. Commandeur, EPS-2009-182-STR, ISBN: 978-90-5892-225-0, http://hdl.handle.net/1765/1

Maeseneire, W. de,, *Essays on Firm Valuation and Value Appropriation*, Promotor: Prof.dr. J.T.J. Smit, EPS-2005-053-F&A, ISBN: 90-5892-082-8, http://hdl.handle.net/1765/6768

Mandele, L.M. van der,, *Leadership and the Inflection Point: A Longitudinal Perspective*, Promotors: Prof.dr. H.W. Volberda & Prof.dr. H.R. Commandeur, EPS-2004-042-STR, ISBN: 90-5892-067-4, http://hdl.handle.net/1765/1302

Meer, J.R. van der, *Operational Control of Internal Transport*, Promotors: Prof.dr. M.B.M. de Koster & Prof.dr.ir. R. Dekker, EPS-2000-001-LIS, ISBN: 90-5892-004-6, http://hdl.handle.net/1765/859

Mentink, A., *Essays on Corporate Bonds*, Promotor: Prof.dr. A.C.F. Vorst, EPS-2005-070-F&A, ISBN: 90-5892-100-X, http://hdl.handle.net/1765/7121

Meyer, R.J.H., *Mapping the Mind of the Strategist: A Quantitative Methodology for Measuring the Strategic Beliefs of Executives*, Promotor: Prof.dr. R.J.M. van Tulder, EPS-2007-106-ORG, ISBN: 978-90-5892-141-3, http://hdl.handle.net/1765/10182

Miltenburg, P.R., *Effects of Modular Sourcing on Manufacturing Flexibility in the Automotive Industry: A Study among German OEMs*, Promotors: Prof.dr. J. Paauwe & Prof.dr. H.R. Commandeur, EPS-2003-030-ORG, ISBN: 90-5892-052-6, http://hdl.handle.net/1765/1039

Moerman, G.A., *Empirical Studies on Asset Pricing and Banking in the Euro Area*, Promotor: Prof.dr. C.G. Koedijk, EPS-2005-058-F&A, ISBN: 90-5892-090-9, http://hdl.handle.net/1765/6666

Moitra, D., *Globalization of R&D: Leveraging Offshoring for Innovative Capability and Organizational Flexibility*, Promotor: Prof.dr. K. Kumar, EPS-2008-150-LIS, ISBN: 978-90-5892-184-0, http://hdl.handle.net/1765/14081

Mol, M.M., *Outsourcing, Supplier-relations and Internationalisation: Global Source Strategy as a Chinese Puzzle*, Promotor: Prof.dr. R.J.M. van Tulder, EPS-2001-010-ORG, ISBN: 90-5892-014-3, http://hdl.handle.net/1765/355

Mom, T.J.M., *Managers Exploration and Exploitation Activities: The Influence of Organizational Factors and Knowledge Inflows*, Promotors: Prof.dr.ir. F.A.J. Van den Bosch & Prof.dr. H.W. Volberda, EPS-2006-079-STR, ISBN: 90-5892-116-6, http://hdl.handle.net/1765

Moonen, J.M., *Multi-Agent Systems for Transportation Planning and Coordination*, Promotors: Prof.dr. J. van Hillegersberg & Prof.dr. S.L. van de Velde, EPS-2009-177-LIS, ISBN: 978-90-5892-216-8, http://hdl.handle.net/1

Mulder, A., *Government Dilemmas in the Private Provision of Public Goods*, Promotor: Prof.dr. R.J.M. van Tulder, EPS-2004-045-ORG, ISBN: 90-5892-071-2, http://hdl.handle.net/1765/1790

Muller, A.R., *The Rise of Regionalism: Core Company Strategies Under The Second Wave of Integration*, Promotor: Prof.dr. R.J.M. van Tulder, EPS-2004-038-ORG, ISBN: 90-5892-062-3, http://hdl.handle.net/1765/1272

Nalbantov, G.I., *Essays on Some Recent Penalization Methods with Applications in Finance and Marketing*, Promotor: Prof. dr P.J.F. Groenen, EPS-2008-132-F&A, ISBN: 978-90-5892-166-6, http://hdl.handle.net/1765/13319

Nederveen Pieterse, A., *Goal Orientation in Teams: The Role of Diversity*, Promotor: Prof.dr. D.L. van Knippenberg, EPS-2009-162-ORG, http://hdl.handle.net/1765/15240

Nguyen, T.T., *Capital Structure, Strategic Competition, and Governance*, Promotor: Prof.dr. A. de Jong, EPS-2008-148-F&A, ISBN: 90-5892-178-9, http://hdl.handle.net/1765/14005

Niesten, E.M.M.I., *Regulation, Governance and Adaptation: Governance Transformations in the Dutch and French Liberalizing Electricity Industries*, Promotors: Prof.dr. A. Jolink & Prof.dr. J.P.M. Groenewegen, EPS-2009-170-ORG, ISBN: 978-90-5892-208-3, http://hdl.handle.net/1765/1

Nieuwenboer, N.A. den, *Seeing the Shadow of the Self*, Promotor: Prof.dr. S.P. Kaptein, EPS-2008-151-ORG, ISBN: 978-90-5892-182-6, http://hdl.handle.net/1765/14223

Nijdam, M.H., *Leader Firms: The Value of Companies for the Competitiveness of the Rotterdam Seaport Cluster*, Promotor: Prof.dr. R.J.M. van Tulder, EPS-2010-216-ORG, ISBN: 978-90-5892-256-4, http://hdl.handle.net/1765/21405

Ning, H., *Hierarchical Portfolio Management: Theory and Applications*, Promotor: Prof.dr. J. Spronk, EPS-2007-118-F&A, ISBN: 90-5892-152-9, http://hdl.handle.net/1765/10868

Noeverman, J., *Management Control Systems, Evaluative Style, and Behaviour: Exploring the Concept and Behavioural Consequences of Evaluative Style*, Promotors: Prof.dr. E.G.J. Vosselman & Prof.dr. A.R.T. Williams, EPS-2007-120-F&A, ISBN: 90-5892-151-2, http://hdl.handle.net/1765/10869

Noordegraaf-Eelens, L.H.J., *Contested Communication: A Critical Analysis of Central Bank Speech*, Promotor: Prof.dr. Ph.H.B.F. Franses, EPS-2010-209-MKT, ISBN: 978-90-5892-254-0, http://hdl.handle.net/1765/21061

Nuijten, I., *Servant Leadership: Paradox or Diamond in the Rough? A Multidimensional Measure and Empirical Evidence*, Promotor: Prof.dr. D.L. van Knippenberg, EPS-2009-183-ORG, http://hdl.handle.net/1765/1

Oosterhout, J. van,, *The Quest for Legitimacy: On Authority and Responsibility in Governance*, Promotor: Prof.dr. T. van Willigenburg & Prof.mr. H.R. van Gunsteren, EPS-2002-012-ORG, ISBN: 90-5892-022-4, http://hdl.handle.net/1765/362

Oosterhout, M. van,, *Business Agility and Information Technology in Service Organizations*, Promotor: Prof,dr.ir. H.W.G.M. van Heck, EPS-2010-198-LIS, ISBN: 90-5092-236-6, http://hdl.handle.net/1765/1

Oostrum, J.M. van,, *Applying Mathematical Models to Surgical Patient Planning*, Promotor: Prof.dr. A.P.M. Wagelmans, EPS-2009-179-LIS, ISBN: 978-90-5892-217-5, http://hdl.handle.net/1765/1

Otgaar, A.H.J., *Industrial Tourism: Where the Public Meets the Private*, Promotor: Prof.dr. L. van den Berg, EPS-2010-219-ORG, ISBN: 90-5892-259-5, http://hdl.handle.net/1765/21585

Paape, L., *Corporate Governance: The Impact on the Role, Position, and Scope of Services of the Internal Audit Function*, Promotors: Prof.dr. G.J. van der Pijl & Prof.dr. H. Commandeur, EPS-2007-111-MKT, ISBN: 90-5892-143-7, http://hdl.handle.net/1765/10417

Pak, K., *Revenue Management: New Features and Models*, Promotor: Prof.dr.ir. R. Dekker, EPS-2005-061-LIS, ISBN: 90-5892-092-5, http://hdl.handle.net/1765/362/6771

Pattikawa, L.H., *Innovation in the Pharmaceutical Industry: Evidence from Drug Introduction in the U.S.*, Promotor: Prof.dr. H.R.Commandeur, EPS-2007-102-MKT, ISBN: 90-5892-135-2, http://hdl.handle.net/1765/9626

Peeters, L.W.P., *Cyclic Railway Timetable Optimization*, Promotors: Prof.dr. L.G. Kroon & Prof.dr.ir. J.A.E.E. van Nunen, EPS-2003-022-LIS, ISBN: 90-5892-042-9, http://hdl.handle.net/1765/429

Pietersz, R., *Pricing Models for Bermudan-style Interest Rate Derivatives*, Promotors: Prof.dr. A.A.J. Pelsser & Prof.dr. A.C.F. Vorst, EPS-2005-071-F&A, ISBN: 90-5892-099-2, http://hdl.handle.net/1765/7122

Pince, C., *Advances in Inventory Management: Dynamic Models*, Promotor: Prof.dr.ir. R. Dekker, EPS-2010-199-LIS, ISBN: 978-90-5892-243-4, http://hdl.handle.net/1765/1

Poel, A.M. van der, *Empirical Essays in Corporate Finance and Financial Reporting*, Promotors: Prof.dr. A. de Jong & Prof.dr. G.M.H. Mertens, EPS-2007-133-F&A, ISBN: 978-90-5892-165-9, http://hdl.handle.net/1765/13320

Popova, V., *Knowledge Discovery and Monotonicity*, Promotor: Prof.dr. A. de Bruin, EPS-2004-037-LIS, ISBN: 90-5892-061-5, http://hdl.handle.net/1765/1201

Potthoff, D., *Railway Crew Rescheduling: Novel Approaches and Extensions*, Promotors: Prof.dr. A.P.M. Wagelmans & Prof.dr. L.G. Kroon, EPS-2010-210-LIS, ISBN: 90-5892-250-2, http://hdl.handle.net/1765/21084

Pouchkarev, I., *Performance Evaluation of Constrained Portfolios*, Promotor: Prof.dr. J. Spronk & Dr. W.G.P.M. Hallerbach, EPS-2005-052-F&A, ISBN: 90-5892-083-6, http://hdl.handle.net/1765/6731

Prins, R., *Modeling Consumer Adoption and Usage of Value-Added Mobile Services*, Promotors: Prof.dr. Ph.H.B.F. Franses & Prof.dr. P.C. Verhoef, EPS-2008-128-MKT, ISBN: 978-90-5892-161-1, `http://hdl.handle.net/1765/12461`

Puvanasvari Ratnasingam, P., *Interorganizational Trust in Business to Business E-Commerce*, Promotors: Prof.dr. K. Kumar & Prof.dr. H.G. van Dissel, EPS-2001-009-LIS, ISBN: 90-5892-017-8, `http://hdl.handle.net/1765/356`

Quak, H.J., *Sustainability of Urban Freight Transport: Retail Distribution and Local Regulation in Cities*, Promotor: Prof.dr. M.B.M. de Koster, EPS-2008-124-LIS, ISBN: 978-90-5892-154-3, `http://hdl.handle.net/1765/11990`

Quariguasi Frota Neto, J., *Eco-efficient Supply Chains for Electrical and Electronic Products*, Promotors: Prof.dr.ir. J.A.E.E. van Nunen & Prof.dr.ir. H.W.G.M. van Heck, EPS-2008-152-LIS, ISBN: 978-90-5892-192-5, `http://hdl.handle.net/1765/14785`

Radkevitch, U.L., *Online Reverse Auction for Procurement of Services*, Promotor: Prof.dr.ir. H.W.G.M. van Heck, EPS-2008-137-LIS, ISBN: 978-90-5892-171-0, `http://hdl.handle.net/1765/13497`

Rinsum, M. van, *Performance Measurement and Managerial Time Orientation*, Promotor: Prof.dr. F.G.H. Hartmann, EPS-2006-088-F&A, ISBN: 90-5892-121-2, `http://hdl.handle.net/1765/7993`

Roelofsen, E.M., *The Role of Analyst Conference Calls in Capital Markets*, Promotors: Prof.dr. G.M.H. Mertens & Prof.dr. L.G. van der Tas RA, EPS-2010-190-F&A, ISBN: 978-90-5892-228-1, `http://hdl.handle.net/1765/1`

Romero Morales, D., *Optimization Problems in Supply Chain Management*, Promotor: Prof.dr.ir. J.A.E.E. van Nunen & Dr. H.E. Romeijn, EPS-2000-003-LIS, ISBN: 90-9014078-6, `http://hdl.handle.net/1765/865`

Roodbergen, K.J., *Layout and Routing Methods for Warehouses*, Promotors: Prof.dr. M.B.M. de Koster & Prof.dr.ir. J.A.E.E. van Nunen, EPS-2001-004-LIS, ISBN: 90-5892-005-4, `http://hdl.handle.net/1765/861`

Rook, L., *Imitation in Creative Task Performance*, Promotor: Prof.dr. D.L. van Knippenberg, EPS-2008-125-ORG, `http://hdl.handle.net/1765/11555`

Rosmalen, J. van, *Segmentation and Dimension Reduction: Exploratory and Model-Based Approaches*, Promotor: Prof.dr. P.J.F. Groenen, EPS-2009-165-MKT, ISBN: 978-90-5892-201-4, `http://hdl.handle.net/1765/15536`

Rus, D., *The Dark Side of Leadership: Exploring the Psychology of Leader Self-serving Behavior*, Promotor: Prof.dr. D.L. van Knippenberg, EPS-2009-178-ORG, `http://hdl.handle.net/1765/1`

Samii, R., *Leveraging Logistics Partnerships: Lessons from Humanitarian Organizations*, Promotors: Prof.dr.ir. J.A.E.E. van Nunen & Prof.dr.ir. L.N. Van Wassenhove, EPS-2008-153-LIS, ISBN: 978-90-5892-186-4, `http://hdl.handle.net/1765/14519`

Schaik, D. van, *M&A in Japan: An Analysis of Merger Waves and Hostile Takeovers*, Promotors: Prof.dr. J. Spronk & Prof.dr. J.P.M. Groenewegen, EPS-2008-141-F&A, ISBN: 978-90-5892-169-7, http://hdl.handle.net/1765/13693

Schauten, M.B.J., *Valuation, Capital Structure Decisions and the Cost of Capital*, Promotors: Prof.dr. J. Spronk & Prof.dr. D. van Dijk, EPS-2008-134-F&A, ISBN: 978-90-5892-172-7, http://hdl.handle.net/1765/13480

Schellekens, G.A.C., *Language Abstraction in Word of Mouth*, Promotor: Prof.dr.ir. A. Smidts, EPS-2010-218-MKT, ISBN: 978-90-5892-252-6, http://hdl.handle.net/1765/21580

Schramade, W.L.J., *Corporate Bonds Issuers*, Promotor: Prof.dr. A. De Jong, EPS-2006-092-F&A, ISBN: 90-5892-125-5, http://hdl.handle.net/1765/8191

Schweizer, T.S., *An Individual Psychology of Novelty-Seeking, Creativity and Innovation*, Promotor: Prof.dr. R.J.M. van Tulder, EPS-2004-048-ORG, ISBN: 90-5892-077-1, http://hdl.handle.net/1765/1818

Six, F.E., *Trust and Trouble: Building Interpersonal Trust Within Organizations*, Promotors: Prof.dr. B. Nooteboom & Prof.dr. A.M. Sorge, EPS-2004-040-ORG, ISBN: 90-5892-064-X, http://hdl.handle.net/1765/1271

Slager, A.M.H., *Banking across Borders*, Promotors: Prof.dr. R.J.M. van Tulder & Prof.dr. D.M.N. van Wensveen, EPS-2004-041-ORG, ISBN: 90-5892-066-6, http://hdl.handle.net/1765/1301

Sloot, L., *Understanding Consumer Reactions to Assortment Unavailability*, Promotors: Prof.dr. H.R. Commandeur, Prof.dr. E. Peelen & Prof.dr. P.C. Verhoef, EPS-2006-074-MKT, ISBN: 90-5892-102-6, http://hdl.handle.net/1765/7438

Smit, W., *Market Information Sharing in Channel Relationships: Its Nature, Antecedents and Consequences*, Promotors: Prof.dr.ir. G.H. van Bruggen & Prof.dr.ir. B. Wierenga, EPS-2006-076-MKT, ISBN: 90-5892-106-9, http://hdl.handle.net/1765/7327

Sonnenberg, M., *The Signalling Effect of HRM on Psychological Contracts of Employees: A Multi-level Perspective*, Promotor: Prof.dr. J. Paauwe, EPS-2006-086-ORG, ISBN: 90-5892-119-0, http://hdl.handle.net/1765/7995

Sotgiu, F., *Not All Promotions are Made Equal: From the Effects of a Price War to Cross-chain Cannibalization*, Promotors: Prof.dr. M.G. Dekimpe & Prof.dr.ir. B. Wierenga, EPS-2010-203-MKT, ISBN: 978-90-5892-238-0, http://hdl.handle.net/1765/1

Speklé, R.F., *Beyond Generics: A closer Look at Hybrid and Hierarchical Governance*, Promotor: Prof.dr. M.A. van Hoepen RA, EPS-2001-008-F&A, ISBN: 90-5892-011-9, http://hdl.handle.net/1765/357

Srour, F.J., *Dissecting Drayage: An Examination of Structure, Information, and Control in Drayage Operations*, Promotor: Prof.dr. S.L. van de Velde, EPS-2010-186-LIS, http://hdl.handle.net/1765/1

Stam, D.A., *Managing Dreams and Ambitions: A Psychological Analysis of Vision Communication*, Promotor: Prof.dr. D.L. van Knippenberg, EPS-2008-149-ORG, http://hdl.handle.net/1765/14080

Stienstra, M., *Strategic Renewal in Regulatory Environments: How Inter- and Intra-organisational Institutional Forces Influence European Energy Incumbent Firms*, Promotors: Prof.dr.ir. F.A.J. Van den Bosch & Prof.dr. H.W. Volberda, EPS-2008-145-STR, ISBN: 978-90-5892-184-0, http://hdl.handle.net/1765/13943

Sweldens, S.T.L.R., *Evaluative Conditioning 2.0: Direct versus Associative Transfer of Affect to Brands*, Promotor: Prof.dr. S.M.J. van Osselaer, EPS-2009-167-MKT, ISBN: 978-90-5892-204-5, http://hdl.handle.net/1765/1

Szkudlarek, B.A., *Spinning the Web of Reentry: [Re]connecting reentry training theory and practice*, Promotor: Prof.dr. S.J. Magala, EPS-2008-143-ORG, ISBN: 978-90-5892-177-2, http://hdl.handle.net/1765/13695

Tempelaar, M.P., *Organizing for Ambidexterity: Studies on the Pursuit of Exploration and Exploitation through Differentiation, Integration, Contextual and Individual Attributes*, Promotors: Prof.dr.ing. F.A.J. van den Bosch & Prof.dr. H.W. Volberda, EPS-2010-191-STR, ISBN: 978-90-5892-231-1, http://hdl.handle.net/1765/1

Teunter, L.H., *Analysis of Sales Promotion Effects on Household Purchase Behavior*, Promotors: Prof.dr.ir. B. Wierenga & Prof.dr. T. Kloek, EPS-2002-016-MKT, ISBN: 90-5892-029-1, http://hdl.handle.net/1765/868

Tims, B., *Empirical Studies on Exchange Rate Puzzles and Volatility*, Promotor: Prof.dr. C.G. Koedijk, EPS-2006-089-F&A, ISBN: 90-5892-113-1, http://hdl.handle.net/1765/8066

Tiwari, V., *Transition Process and Performance in IT Outsourcing: Evidence from a Field Study and Laboratory Experiments*, Promotors: Prof.dr.ir. H.W.G.M. van Heck & Prof.dr. P.H.M. Vervest, EPS-2010-201-LIS, ISBN: 978-90-5892-241-0, http://hdl.handle.net/1765/1

Tuk, M.A., *Is Friendship Silent When Money Talks? How Consumers Respond to Word-of-Mouth Marketing*, Promotors: Prof.dr.ir. A. Smidts & Prof.dr. D.H.J. Wigboldus, EPS-2008-130-MKT, ISBN: 978-90-5892-164-2, http://hdl.handle.net/1765/12702

Tzioti, S., *Let Me Give You a Piece of Advice: Empirical Papers about Advice Taking in Marketing*, Promotors: Prof.dr. S.M.J. van Osselaer & Prof.dr.ir. B. Wierenga, EPS-2010-211-MKT, ISBN: 978-90-5892-251-9, http://hdl.handle.net/1765/21149

Vaccaro, I.G., *Management Innovation: Studies on the Role of Internal Change Agents*, Promotors: Prof.dr. F.A.J. van den Bosch & Prof.dr. H.W. Volberda, EPS-2010-212-STR, ISBN: 978-90-5892-253-3, http://hdl.handle.net/1765/21150

Valck, K. de, *Virtual Communities of Consumption: Networks of Consumer Knowledge and Companionship*, Promotors: Prof.dr.ir. G.H. van Bruggen & Prof.dr.ir. B. Wierenga, EPS-2005-050-MKT, ISBN: 90-5892-078-X, http://hdl.handle.net/1765/6663

Valk, W. van der, *Buyer-Seller Interaction Patterns During Ongoing Service Exchange*, Promotors: Prof.dr. J.Y.F. Wynstra & Prof.dr.ir. B. Axelsson, EPS-2007-116-MKT, ISBN: 90-5892-146-8, http://hdl.handle.net/1765/10856

Verheijen, H.J.J., *Vendor-Buyer Coordination in Supply Chains*, Promotor: Prof.dr.ir. J.A.E.E. van Nunen, EPS-2010-194-LIS, ISBN: 90-5892-234-2, http://hdl.handle.net/1765/1

Verheul, I., *Is There a (Fe)male Approach? Understanding Gender Differences in Entrepreneurship*, Promotor: Prof.dr. A.R. Thurik, EPS-2005-054-ORG, ISBN: 90-5892-080-1, http://hdl.handle.net/1765/2005

Verwijmeren, P., *Empirical Essays on Debt, Equity, and Convertible Securities*, Promotors: Prof.dr. A. de Jong & Prof.dr. M.J.C.M. Verbeek, EPS-2009-154-F&A, ISBN: 978-90-5892-187-1, http://hdl.handle.net/1765/14312

Vis, I.F.A., *Planning and Control Concepts for Material Handling Systems*, Promotors: Prof.dr. M.B.M. de Koster & Prof.dr.ir. R. Dekker, EPS-2002-014-LIS, ISBN: 90-5892-021-6, http://hdl.handle.net/1765/866

Vlaar, P.W.L., *Making Sense of Formalization in Interorganizational Relationships: Beyond Coordination and Control*, Promotors: Prof.dr.ir. F.A.J. Van den Bosch & Prof.dr. H.W. Volberda, EPS-2006-075-STR, ISBN: 90-5892-103-4, http://hdl.handle.net/1765/7326

Vliet, P. van, *Downside Risk and Empirical Asset Pricing*, Promotor: Prof.dr. G.T. Post, EPS-2004-049-F&A, ISBN: 90-5892-07-55, http://hdl.handle.net/1765/1819

Vlist, P. van der, *Synchronizing the Retail Supply Chain*, Promotors: Prof.dr.ir. J.A.E.E. van Nunen & Prof.dr. A.G. de Kok, EPS-2007-110-LIS, ISBN: 90-5892-142-0, http://hdl.handle.net/1765/10418

Vries-van Ketel, E. de, *How Assortment Variety Affects Assortment Attractiveness: A Consumer Perspective*, Promotors: Prof.dr. G.H. van Bruggen & Prof.dr.ir. A. Smidts, EPS-2006-072-MKT, ISBN: 90-5892-101-8, http://hdl.handle.net/1765/7193

Vromans, M.J.C.M., *Reliability of Railway Systems*, Promotors: Prof.dr. L.G. Kroon, Prof.dr.ir. R. Dekker & Prof.dr.ir. J.A.E.E. van Nunen, EPS-2005-062-LIS, ISBN: 90-5892-089-5, http://hdl.handle.net/1765/6773

Vroomen, B.L.K., *The Effects of the Internet, Recommendation Quality and Decision Strategies on Consumer Choice*, Promotor: Prof.dr. Ph.H.B.F. Franses, EPS-2006-090-MKT, ISBN: 90-5892-122-0, http://hdl.handle.net/1765/8067

Waal, T. de, *Processing of Erroneous and Unsafe Data*, Promotor: Prof.dr.ir. R. Dekker, EPS-2003-024-LIS, ISBN: 90-5892-045-3, http://hdl.handle.net/1765/870

Waard, E.J. de, *Engaging Environmental Turbulence: Organizational Determinants for Repetitive Quick and Adequate Responses*, Promotors: Prof.dr. H.W. Volberda & Prof.dr. J. Soeters, EPS-2010-189-STR, ISBN: 978-90-5892-229-8, http://hdl.handle.net/1765/1

Wall, R.S., *Netscape: Cities and Global Corporate Networks*, Promotor: Prof.dr. G.A. van der Knaap, EPS-2009-169-ORG, ISBN: 978-90-5892-207-6, http://hdl.handle.net/1765/1

Watkins Fassler, K., *Macroeconomic Crisis and Firm Performance*, Promotors: Prof.dr. J. Spronk & Prof.dr. D.J. van Dijk, EPS-2007-103-F&A, ISBN: 90-5892-138-3, http://hdl.handle.net/1765/10065

Weerdt, N.P. van der, *Organizational Flexibility for Hypercompetitive Markets: Empirical Evidence of the Composition and Context Specificity of Dynamic Capabilities and Organization Design Parameters*, Promotor: Prof.dr. H.W. Volberda, EPS-2009-173-STR, ISBN: 978-90-5892-215-1, http://hdl.handle.net/1765/1

Wennekers, A.R.M., *Entrepreneurship at Country Level: Economic and Non-Economic Determinants*, Promotor: Prof.dr. A.R. Thurik, EPS-2006-81-ORG, ISBN: 90-5892-115-8, http://hdl.handle.net/1765/7982

Wielemaker, M.W., *Managing Initiatives: A Synthesis of the Conditioning and Knowledge-Creating View*, Promotors: Prof.dr. H.W. Volberda & Prof.dr. C.W.F. Baden-Fuller, EPS-2003-28-STR, ISBN: 90-5892-050-X, http://hdl.handle.net/1765/1042

Wijk, R.A.J.L. van, *Organizing Knowledge in Internal Networks: A Multilevel Study*, Promotor: Prof.dr.ir. F.A.J. van den Bosch, EPS-2003-021-STR, ISBN: 90-5892-039-9, http://hdl.handle.net/1765/347

Wolters, M.J.J., *The Business of Modularity and the Modularity of Business*, Promotor: Prof. mr. dr. P.H.M. Vervest & Prof.dr.ir. H.W.G.M. van Heck, EPS-2002-011-LIS, ISBN: 90-5892-020-8, http://hdl.handle.net/1765/920

Wubben, M.J.J., *Social Functions of Emotions in Social Dilemmas*, Promotors: Prof.dr. D. De Cremer & Prof.dr. E. van Dijk, EPS-2009-187-ORG, http://hdl.handle.net/1765/1

Xu, Y., *Empirical Essays on the Stock Returns, Risk Management, and Liquidity Creation of Banks*, Promotor: Prof.dr. M.J.C.M. Verbeek, EPS-2010-188-F&A, http://hdl.handle.net/1765/1

Yang, J., *Towards the Restructuring and Co-ordination Mechanisms for the Architecture of Chinese Transport Logistics*, Promotor: Prof.dr. H.E. Harlambides, EPS-2009-157-LIS, ISBN: 978-90-5892-198-7, http://hdl.handle.net/1765/14527

Yu, M., *Enhancing Warehouse Perfromance by Efficient Order Picking*, Promotor: Prof.dr. M.B.M. de Koster, EPS-2008-139-LIS, ISBN: 978-90-5892-167-3, http://hdl.handle.net/1765/13691

Zhang, X., *Strategizing of Foreign Firms in China: An Institution-based Perspective*, Promotor: Prof.dr. B. Krug, EPS-2007-114-ORG, ISBN: 90-5892-147-5, http://hdl.handle.net/1765/10721

Zhang, X., *Scheduling with Time Lags*, Promotor: Prof.dr. S.L. van de Velde, EPS-2010-206-LIS, ISBN: 978-90-5892-244-1, http://hdl.handle.net/1765/1

Zhou, H., *Knowledge, Entrepreneurship and Performance: Evidence from Country-level and Firm-level Studies*, Promotors: Prof.dr. A.R. Thurik & Prof.dr. L.M. Uhlaner, EPS-2010-207-ORG, ISBN: 90-5892-248-9, http://hdl.handle.net/1765/1

Zhu, Z., *Essays on Chinas Tax System*, Promotors: Prof.dr. B. Krug & Prof.dr. G.W.J. Hendrikse, EPS-2007-112-ORG, ISBN: 90-5892-144-4, http://hdl.handle.net/1765/10502

Zwart, G.J. de, *Empirical Studies on Financial Markets: Private Equity, Corporate Bonds and Emerging Markets*, Promotors: Prof.dr. M.J.C.M. Verbeek & Prof.dr. D.J.C. van Dijk, EPS-2008-131-F&A, ISBN: 978-90-5892-163-5, http://hdl.handle.net/1765/12703

ROLLING STOCK RESCHEDULING IN PASSENGER RAILWAYS
APPLICATIONS IN SHORT-TERM PLANNING AND IN DISRUPTION MANAGEMENT

Modern society is highly dependent on a reliable railway system for workforce mobility and easy access to the cities. However, the daily operations of a large passenger railway system are subject to unexpected disruptions such as rolling stock breakdowns or malfunctioning infrastructure. In a disrupted situation, the railway operator must adapt the timetable, rolling stock and crew to the modified conditions. This adaptation of resource allocations requires the solution of complex combinatorial problems in very short time and thus represents a major challenge for the involved dispatchers.

In this thesis we develop models and solution methods for the rescheduling of the rolling stock during disruptions. The models incorporate service aspects (such as seat capacity), efficiency aspects (such as number of kilometers driven by the rolling stock), and process related aspects (such as the need for night-time relocation of rolling stock).

The thesis contains applications of the developed models in three different contexts. First, we present a framework for applying the rescheduling models in the highly uncertain environment of railway disruption management, and we demonstrate the trade-off between computation time and solution quality. Second, we embed the rolling stock rescheduling models in a simulation framework to account for the dynamic passenger behavior during disruptions. This framework allows us to significantly decrease the delays experienced by passengers. Third, we apply the rescheduling models to real-life planning problems from the short-term planning department of the Netherlands Railways. The models lead to a considerable speed-up of the process and significant savings.

ERIM

The Erasmus Research Institute of Management (ERIM) is the Research School (Onderzoek-school) in the field of management of the Erasmus University Rotterdam. The founding participants of ERIM are Rotterdam School of Management (RSM), and the Erasmus School of Economics (ESE). ERIM was founded in 1999 and is officially accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW). The research undertaken by ERIM is focused on the management of the firm in its environment, its intra- and interfirm relations, and its business processes in their interdependent connections.

The objective of ERIM is to carry out first rate research in management, and to offer an advanced doctoral programme in Research in Management. Within ERIM, over three hundred senior researchers and PhD candidates are active in the different research programmes. From a variety of academic backgrounds and expertises, the ERIM community is united in striving for excellence and working at the forefront of creating new business knowledge.

Erasmus Research Institute of Management - ERIM

ERIM PhD Series
Research in Management

Erasmus Research Institute of Management - **ERIM**
Rotterdam School of Management (RSM)
Erasmus School of Economics (ESE)
P.O. Box 1738, 3000 DR Rotterdam
The Netherlands

Tel. +31 10 408 11 82
Fax +31 10 408 96 40
E-mail info@erim.eur.nl
Internet www.erim.eur.nl

Design & layout: B&T Ontwerp en advies (www.b-en-t.nl) Print: Haveka (www.haveka.nl)