Mälardalen University Press Licentiate Theses
No. 196

MATHEMATICAL MODELS FOR OPTIMISING DECISION SUPPORT SYSTEMS IN THE RAILWAY INDUSTRY

Sara Gestrelius

2015

MÄLARDALEN UNIVERSITY
SWEDEN

School of Innovation, Design and Engineering

ISBN 978-91-7485-196-0
ISSN 1651-9256
Printed by Arkitektkopia, Västerås, Sweden

Abstract

After the deregulation of the Swedish railway industry, train operating companies compete for and on the same infrastructure. This makes the allocation of rail capacity a most delicate problem, and for a well-functioning railway system the allocation must be fair, efficient and functional.

The capacity allocation tasks include e.g. constructing the yearly timetable and making track allocation plans for rail yards. The state of practice is that experienced planners construct the schedules manually with little or no decision support. However, as the planners are often faced with large combinatorial problems that are notoriously hard to solve there is a great potential in implementing optimising decision support systems. The research presented in this licentiate thesis aims at developing and examining mathematical models and methods that could be part of such support systems. The thesis focuses on two planning problems in particular, and the presented methods have been developed especially for the Swedish railway system.

First of all, a model for optimising a train timetable with respect to robustness is presented. The model tries to increase the number of alternative meeting locations that can be used in a disturbed traffic situation and has an execution time of less than 5 minutes when solving the problem for the track section between Boden and Vännäs. Secondly, the problem of generating efficient classification bowl schedules for shunting yards is examined. The aim is to find the track allocation that minimises the number of required shunting movements while still respecting all operational, physical and time constraints imposed by the yard. Three optimisation models are presented, and simple planning rules are also investigated. The methods are tested on historic data from Hallsberg, the largest shunting yard in Sweden, and the results show that while the simple planning rules are not adequate for planning the classification bowl, two of the optimisation models consistently return an optimal solution within an acceptable execution time.

Populärvetenskaplig sammanfattning

Efter avregleringen av den svenska järnvägsindustrin konkurrerar tågoperatörer om och på samma infrastruktur. Detta gör tilldelning av järnvägskapacitet till ett delikat problem och för en välfungerande järnväg krävs en rättvis, effektiv, och ändamålsenlig tilldelning.

I Sverige är det Trafikverkets ansvar att dela ut järnvägskapacitet, och i uppgiften ingår bl.a. att konstruera den årliga tågplanen och att allokera spår till tåg på bangårdar. Dagens tilldelningsprocess går till så att erfarna planerare konstruerar scheman manuellt med lite eller inget beslutsstöd. Både tidtabellsläggning och rangerbangårdsplanering genererar dessvärre schemaläggningsproblem som är svåra att lösa om det finns mycket trafik. Det är därför troligt att den slutgiltiga tilldelningen skulle bli mer effektiv och ändamålsenlig om beslutsstödssystem introducerades. Forskningen som presenteras i denna licentiatavhandling syftar till att ta fram och undersöka matematiska modeller och metoder som kan ligga till grund för sådana beslutsstödssystem för två olika planeringsproblem. Alla matematiska modeller som presenteras i avhandlingen är framtagna speciellt för den svenska järnvägsindustrin, och målet har varit att modellerna ska vara användbara i verkligheten.

Det första planeringsproblemet som behandlas i avhandlingen är hur man kan lägga en mer robust tidtabell, och en modell som kan öka omplaneringsrobustheten i en tidtabell presenteras. Modellen syftar till att öka antalet alternativa mötesplatser som kan användas i en situation med förseningar utan att ändra ankomst- eller avgångstiderna för tåg på viktiga orter. Modellen returnerar en produktionstidtabell och en säkerhetstidtabell som tryckts isär antingen i tid eller i antalet länkar vars startpunkt och slutpunkt är alternativa mötesplatser. Detta visualiserar och maximerar utrymmet mellan produktionstidtabellen och säkerhetstidtabellen och ökar därmed chansen att det finns en alternativ mötesplats som kan användas om ett tåg skulle bli sent till ett möte. Exekveringstiderna är under 5 minuter för våra fallstudier på enkelspårsbanan mellan Boden och Vännäs.

Det andra problem som ingår i avhandlingen är tilldelning av spår på riktningsgruppen på en rangerbangård. Målet är att givet de operationella, fysiska och tidskrav som finns hitta det schema som kräver minst omflyttning av vagnar. Ett flertal optimeringsmodeller med allt kortare exekveringstider introduceras, och vi gör också en jämförelse mellan dessa optimeringsmodeller och enklare planeringstumregler. Resultaten visar att de undersökta tumreglerna inte är lämpliga för planering av riktningsgruppen, men att två av optimeringsmodellerna konsekvent returnerar optimala lösningar inom en acceptabel exekveringstid för våra testfall från Hallsbergs rangerbangård. Modellerna utökas också för att klara av rullande horisont planering och sortering av vagnar inom avgående tåg utan att exekveringstiderna blir för långa.

Acknowledgments

First of all I would like to heartfully thank my supervisors: Markus Bohlin, Martin Aronsson and Mikael Sjödin. I greatly appreciate all knowledge and advice that you have shared! I also want to thank everyone involved in ITS-EASY where I was given an academic belonging and opportunities for experience sharing.

My colleagues at SICS deserve a massive thank you, especially the people whom I've worked with in the railway projects: Malin, PD and Victoria. It's been great! I also want to thank the people who do not have the fortune of working at SICS but whom I still count as colleagues, namely all my co-authors: Hasse, Florian, Fahimeh, Matúš and Holger, and also all the KAJT PhD students.

Every single person at Trafikverket and Green Cargo AB who have shared their knowledge and discussed and explained how everything works also deserve a huge thank you! Without you I would never ever have gained the knowledge necessary for this thesis. I would also like to thank Trafikverket for funding the research.

Finally, I want to thank my friends and family who've provided an awesome mix of motivation and distraction. Emelie, Pernilla, Jesper and Katarina, thanks for the fun! Mamma and Pappa and Erik, thanks for everything!

Sara Gestrelius
20:36, April 1st, 2015
Kista, Sweden

List of publications

Papers included in the licentiate thesis[1]

Paper A: S. Gestrelius, M. Aronsson, M. Forsgren, and H. Dahlberg. On the delivery robustness of train timetables with respect to production replanning possibilities. In *Proceedings of the The 2nd International Conference on Road and Rail Infrastructure (CETRA)*, Dubrovnik, Croatia, 2012.

Paper B: M. Bohlin, F. Dahms, H. Flier, and S. Gestrelius. Optimal Freight Train Classification using Column Generation. In D. Delling and L. Liberti, editors, *12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems, volume 25 of OpenAccess Series in Informatics (OASIcs)*, pages 10-22, Dagstuhl, Germany, 2012. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Paper C: M. Bohlin, S. Gestrelius, and F. Khoshniyat. Simulation of planning strategies for track allocation at marshalling yards. In G. M. Carlomagno, C. A. Brebbia, S. Hernández, editors, *Computational Methods and experimental measurements XVI, volume 55 of WIT Transactions on Modelling and Simulation*, pages 465-475, Southampton, UK, 2013. WITPress.

Paper D: S. Gestrelius, F. Dahms, and M. Bohlin. Optimisation of simultaneous train formation and car sorting at marshalling yards. In *Proceedings of the 5th International Seminar on Railway Operations Modelling and Analysis (RailCopenhagen)*, Copenhagen, Denmark, 2013.

Paper E: M. Bohlin, S. Gestrelius, F. Dahms, M. Mihalák and H. Flier. Optimization Methods for Multistage Freight Train Formation. *Transportation Science*, Articles in Advance, 2015.

---

[1] The included papers have been reformatted to comply with the licentiate thesis layout.

Related publications not included in the thesis

1. M. Forsgren, M. Aronsson and S. Gestrelius. Maintaining tracks and traffic flow at the same time. *Journal of Rail Transport Planning & Management*, 3(3):111–123, 2013.
2. M. Forsgren, M. Aronsson, and S. Gestrelius. Towards Shorter Lead Times in Railway Timetabling in Sweden. In *Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation*, pages 1053-1058, Hague, Netherlands, 2013.
3. M. Forsgren, M. Aronsson, S. Gestrelius and H. Dahlberg. Using timetabling optimization prototype tools in new ways to support decision making. *Computers in Railways XIII: Computer System Design and Operation in the Railway and Other Transit Systems*, 127:439-450, 2012.
4. M. Aronsson, M. Forsgren and S. Gestrelius. *The Road to Incremental Allocation & Incremental Planning*. SICS Technical Report T2012:09, 2012.
5. M. Forsgren, M. Aronsson, S. Gestrelius and H. Dahlberg. *Opportunities and challenges with new railway planning approach in Sweden*. SICS Technical Report T2012:11, 2012.
6. S. Gestrelius, M. Bohlin, P. Danielsson and M. Aronsson. *Teknisk slutrapport för RANPLAN - Beräkningstöd för planering och resursallokering på rangerbangården*. SICS Technical Report T2011:11, 2011.

Contents

II Included Papers 29

I

Thesis

Chapter 1

Introduction

In 1988 the Swedish national railway company was split into an infrastructure manager (called *Banverket* then and *Trafikverket* now) and a transportation service provider (*Statens Järnvägar*). This was the start of a market deregulation and in 2014 there were 45 organisations applying for train paths in Sweden [1]. Getting the required track capacity is a decisive business component for train operating companies (TOCs), and the robustness and effectiveness of the timetable and shunting plans highly influence the quality of the train services that can be provided. An effective and fair rail capacity allocation process is therefore a prerequisite for a well-functioning railway system.

The Swedish infrastructure manager, *Trafikverket*, is responsible for allocating capacity to TOCs and maintenance entrepreneurs. Each year the planners at Trafikverket allocate capacity to approximately 1 500 000 trains and 2200 maintenance works [2]. TOCs apply for rail capacity in April each year. The yearly timetable is then constructed by planners at Trafikverket and it is finalised in September. The first day of operation is in the middle of December. Track allocation at shunting yards is normally done on a daily basis by planners at the shunting yard.

Train timetabling and shunting yard scheduling are combinatorial optimisation problems that are very hard to solve. Despite this the planners at Trafikverket currently have limited to no access to decision support tools. This licentiate aims at developing and examining mathematical models and methods that could support and improve the capacity allocation processes at Trafikverket. More precisely, one model for increasing the robustness of a train timetable and multiple models for generating efficient classification bowl schedules at shunting yards are presented. These two research areas will be introduced in more detail in section 1.1.

1.1 Background

1.1.1 Robust timetabling

Delayed trains is a problem that dispatchers have to deal with daily and a timetable should preferably be able to cope with small delays. A commonly used tactic for making timetables more robust is to include time supplements that can absorb small delays caused by unforeseeable events, see e.g. [3, 4, 5]. However, adding time supplements to a train path inherently prolongs the train's travel time, and also increases the capacity required by the train.

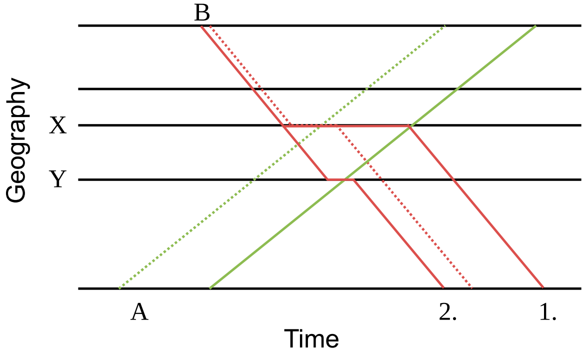

Figure 1.1: Example of a train meeting in a delayed situation.

One of the main problems in a delayed situation is that train meetings can not be carried out as planned. Figure 1.1 shows an example of a meeting in a delayed situation in a time-geography graph with time on the x-axis and geography on the y-axis. The planned timetable is shown in dotted lines, and the delayed situation in solid lines. In the example two trains A and B moving in opposite directions on a single track line are to meet at a certain meeting location X. During operation train B is running slightly early while train A is running late. The dispatcher has two options; to let train B wait at meeting location X until train A arrives (1.), or to move the meeting to another meeting location Y (2.). The second option only exists if there is a suitable alternative meeting location (like Y). If this is not the case then train B has to wait for train A and the delay will spread to train B (a *knock-on* delay).

Our research in the area of timetable robustness focuses on increasing the number of alternative meeting locations, and thereby constructing a timetable that increases the dispatchers ability to deal with delays. We call this *replanning robustness*. The research builds on the idea that arriving/departing at exactly the same time each day is only important at a few geographical points during a train's trip (passenger exchanges, origins and destinations for freight, associations etc.), i.e. that there exists certain *delivery commitments*. The challenge is then to construct a timetable with the maximum number of alternative meeting locations that can be used to fulfil the delivery commitments.

1.1.2 Shunting yard planning

Freight companies operating a carload system offer the service of transporting a car from an origin to a destination. It is often not cost-effective to run direct trains between every single origin-destination pair, but rather the cars are combined into trains operating in a hub-and-spoke network. The hubs in the network are shunting yards, where cars from inbound trains are sorted into new outbound trains. Shunting yards constitute bottlenecks in the transportation network, and if a car misses its outbound train the delay incurred can be very large. Therefore it is important that shunting yards are operated as efficiently and robust as possible and to this aim we have developed mathematical models for classification bowl scheduling at hump yards.

Shunting yards normally consist of three sub-yards: an arrival yard, a classification bowl and a departure yard (see Fig 1.2). Inbound trains arrive to the arrival yard where the cars are decoupled and undergo inspection. The cars are then rolled into the classification bowl where the outbound trains are built. The outbound trains are then rolled to the departure yard where they wait until their departure time. Many shunting yards have a hump between the arrival yard and the classification bowl. The cars can then be pushed over the hump and are rolled to their allocated classification bowl track by means of gravity. If there is no hump an engine must move all cars to their tracks.

In Hallsberg, the largest shunting yard in Sweden, the classification bowl tracks are divided into two sets: *train formation tracks* and *mixing tracks*. Train formation tracks are used for compounding outbound trains and a train formation track may only contain cars for one and only one train at any point in time. Mixing tracks may contain cars for many outbound trains and are used

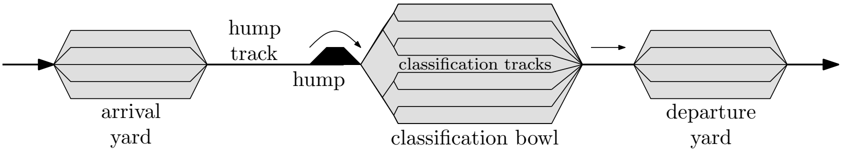

Figure 1.2: Physical layout of a typical hump yard.

to temporarily store cars whose outbound trains still haven't been allocated a train formation track. At given time points the cars on the mixing tracks are coupled and pulled back over the hump to once again be rolled into the classification bowl. Cars belonging to trains that are now being compounded on a train formation track can then be rolled to that track, while cars whose trains still do not have an allocated train formation track will be rolled back to the mixing track. In general any car order is acceptable in the outbound trains, but for some trains, called *blocked trains*, the car order matters. For blocked trains the cars must be rolled to the train formation track in a correct order.

In our research, we aim at developing optimising models that find a schedule that fulfils all the operational, physical and time constraints imposed by the shunting yard and the train timetable while minimising the number of shunting movements. Given the Swedish operational practice this is equivalent to minimising the number of car pull-backs.

1.2 Outline

The thesis is a compilation thesis consisting of two parts; Part I (chapters 1 to 5) contains an introduction to the licentiate research, while Part II (chapters 6 to 10) contains all included papers. Chapter 1 provided an introduction to the research work and area, and chapter 2 contains some important related works. In chapter 3 the motivation and research vision are presented, and the limitations and research questions are specified. Chapter 4 provides a brief summary of all included papers, outlines how the papers relate to the research questions and also specifies my contribution to the papers. Finally, chapter 5 summarises the results of the included papers and discusses future work.

Chapter 2

Related Work

This chapter presents some important related works for the two different research areas. It also explains how the research in the included papers relate to and differ from the previous work.

2.1 Robust timetabling

In real operation the actual timings of events (trips, station stops etc.) often vary slightly from the estimates used when constructing the timetable. To ensure smooth operation the timetable must therefore be robust to such variations in timings. Kroon, Huisman and Marti [6] lists three different types of timetable robustness. First of all there may be time supplements in the timing estimations that can absorb disturbances. Secondly, the timetable could be constructed such that there are few knock-on delays and finally the timetable could be constructed such that the delays are recovered quickly, possibly by some light dispatching actions.

Cacchiani and Toth [7] calls the optimisation problem without any robustness requirements the *nominal problem*. The characteristic optimised in the nominal problem may vary (train trip time, socio-economic value, number of schedules trains etc.) but the objective function value defines the *efficiency* of the timetable [7]. When adding the requirement that a timetable solution should be robust the efficiency of the timetable will often decrease. That is, there is a penalty to pay for robustness, aptly called the *price of robustness* by Bertsimas and Sim [8]. The various approaches to robust timetable optimisation presented in the literature provide different flavours of how to weight the efficiency, robustness and computability.

Papers on robust optimisation often define a set of scenarios $\omega \in \Omega$ that the timetable (to some extent) should be robust for. An example of a scenario is that the coefficients of the constraint coefficient matrix $\mathbf{A}$ can take a value from a symmetric distribution in a bounded range, i.e. $\tilde{a}_{ij} \in [a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$. Bertsimas and Sim [8] calls this model of data uncertainty U, and we will also use this naming convention.

Soyster [9] introduced a framework for *robust optimisation* in 1973 where the returned solution will be feasible under all scenario realisations. However, this high degree of robustness comes at a high price in terms of efficiency, and the approach is considered too conservative for most practical situations. Ben-Tal and Nemirovski [10] present a method similar to Soyster [9], but where the probability that a constraint i is violated is at most $e^{\frac{\Omega_i^2}{2}}$ under the model of data uncertainty U and where Ω_i is a “safety parameter” that can be used to set the reliability level. Unfortunately, this model is non-linear and hence suffers from poor computability. Bertsimas and Sim [8] introduce yet another robust optimisation approach where at most Γ_i coefficients of row i may change. The model will return a solution that is deterministically robust when at most Γ_i coefficients change in each row i and Bertsimas and Sim [8] also derive probability bounds for a solution remaining feasible when data change according to model U regardless of how many coefficients that change.

Fischetti and Monaci [11, 12] introduce even more flexibility to the Bertsimas and Sim [8] approach by including slack variables in the rows with uncertain coefficients, thereby making the constraints soft. The slack variables are then minimised in the objective function while a constraint is included that forbids the efficiency of the solution to deteriorate too much. Fischetti and Monaci [11, 12] calls this approach *Light Robustness*.

A different approach to generating robust solutions is *stochastic optimisation*. Stochastic optimisation is more flexible than robust optimisation as it doesn’t include hard constraints that restrict the solution space. Kroon, Dekker and Vromans [13] present a stochastic optimisation model that reallocate buffer time and time supplements in an existing cyclic timetable such that the average weighted delay for the trains is minimised. Fischetti and Monaci [11, 12] implement a slightly different stochastic programming approach where a constraint forbidding the robust timetable to worsen the nominal objective value by more than δ is included and then the cumulative delay is minimised. A problem with stochastic programming approaches is that they include a copy of the original timing variable for every scenario, which makes them very large and time consuming to solve [11, 14]. In an attempt to overcome the size-problem Fischetti and Monaci [11, 12] introduce a "slim" stochastic model that only contains one copy of the original variables and recourse variables s_{ij}^{ω} for every event $(ij) \in P$ and scenario $\omega \in \Omega$. Events are e.g. link trips or passenger transfers. The recourse variables measure the unabsorbed extra delay for a scenario and are minimised in the objective function.

Liebchen et al. [14] include the recovery strategies available when trains are running late in their model. But from an original optimisation method and a set of scenarios the *Recoverable Robustness* framework of Liebchen et al. [14] also include a set of admissible recovery algorithms, $A \in \mathcal{A}$. The cost of the recovery actions is constrained and the robust time table problem solved. Liebchen et al. [14] model the timetable problem with two types of recovery actions; delay propagation and cancelled passenger transfers. Further, models for two different recovery algorithms are presented; a linear program and a rule-based algorithm. The models are scenario-expansions and thereby too large to solve readily, but Liebchen et al. [14] show that during certain conditions the problem can be reformulated as a compact and solvable linear program.

Many of the robustness models above tries to find a timetable solution that do not change in case of delays. That is, the train order should remain the same and the arrival and departure times should be as close as possible to the ones in the nominal timetable. The recoverable robustness framework does not per se require that the timetable solution remains the same, but the simple timetable model used does not include train order. However, changing the train order is one of the most effective tools for avoiding knock-on delays. The robustness method presented in this thesis therefore focusses on alternative meeting locations rather than where to place buffer or supplement time. It does not rely on any scenarios and is a linear program that is readily solvable for our case study. Further, given our definition of efficiency, i.e. that all trains must arrive and depart on time at delivery commitment locations, the method does not allow for any deterioration in the objective value, i.e. in the timetable efficiency.

2.2 Shunting yard planning

The core of the shunting yard planning problem is neat and easy to explain: the cars of the inbound trains should be sorted into new outbound trains in the most efficient way. However, the plethora of operational and physical constraints that could be included in the planning yields many problem variations. For extensive (but non-comprehensive) classification schemes for shunting yard planning problems see e.g. Hansmann and Zimmerman [15] or Di Stefano et al. [16]. Further, Gatto et al [17] and Boysen et al. [18], provide overviews of classification procedures and sorting methods.

Shunting planning methods are either single-stage, where the cars can be pulled-back only once, or multi-stage, where the cars can be pulled-back and pushed over the hump multiple times (re-classification) [17]. As the Swedish situation does not allow for single-stage sorting we focus on multi-stage methods. Some early multi-stage methods are so-called *sorting schemes* that can be used by hand to construct classification schedules. *Sorting by trains*, *Sorting by block*, *Triangular Sorting* and *Geometric Sorting* are all basic sorting schemes (see e.g. [17, 18, 19] for an introduction). They take a set of inbound cars and then sort them into outbound trains. Arrival and departure times of trains are not considered, nor the order of the cars in the inbound trains. The number of required tracks, classification steps and car roll-ins differ depending on the scheme and an overview of the trade-offs can be found Gatto et al. [17]. According to Boysen et al. [18], the Triangular Sorting scheme is being used at the Lausanne Triage Shunting yard.

As stated above the sorting schemes don't take the car order of the arriving trains into consideration. They thereby risk making excessive use of re-classification as they fail to exploit the *pre-sortedness* of the arriving cars. Dahlhaus et al. [20] present an adaptive radix sorting scheme that takes the pre-sortedness of cars into consideration and show that using this scheme reduces the number of re-classification steps.

Jacob et al. [19] introduce a binary representation of the planning problem that can be used to construct an optimal schedule in linear time for the case of unrestricted track lengths. Jacob et al. [19] also use their encoding to analyse the previously mentioned sorting schemes and prove that for a restricted number of tracks and restricted track lengths the problem is NP-hard. Despite the complexity results, Maue and Nunkesser [21] use the binary encoding from Jacob et al. [19] to construct an IP model including both a limited track capacity and a restricted number of classification tracks. The model also includes a rudimentary departure time constraint and a number of practical limitations important in Switzerland such as a time-dependent number of available classification tracks and the use of two humps. The model is tested on both synthetic and real data, and in the real-data test instance the optimal solution is found in 3 minutes. In Márton, Maue and Nunkesser[22] the model from Maue and Nunkesser [21] is further adapted for the real operational practices of Lausanne Triage Shunting

Yard. More precisely, the outbound trains are divided into two sets that are to be built on two different parts of the yard: the north partition and the south partition. Variables corresponding to the north partition are marked with a $\hat{}$ and variables corresponding to the south partition with a $\check{}$. Deciding which trains that should be in which set is part of the optimisation. Further, there is a two-stage process that dictates that first all cars should be rolled into one of $\hat{W}/\check{W}$ sorting tracks, exactly which track a car should be rolled to is decided by the optimisation model. After this the outbound trains can be built. The number of tracks available for building outbound trains vary with time, and the track capacity at time point t is given by $\hat{N}_t/\check{N}_t$. This new model finds the optimal schedule for a one day example case from Lausanne Triage Shunting Yard after 5.75 hours.

The models presented by Maue and Nunkesser [21] and Márton, Maue and Nunkesser[22] are not compatible with the Swedish use-case we are investigating as they do not respect a given pull-back schedule nor exact arrival and departure times for the trains. Further, they model the track capacity as uniform, which is not the case for Swedish shunting yards.

The Swedish practice with mixing tracks and a booking system is first introduced and modelled in Bohlin et al. [23, 24]. However, the models fail to return optimal solutions for some of the problem instances in the test-case. The papers included in this licentiate present more efficient integer programming models respecting the same real-world constraints as in Bohlin et al. [23, 24]. The new models successfully return the optimal solution for all test-instances. Further, the column generation model is extended so that the car ordering in the outbound trains can be specified, which is not included in Bohlin et al. [23, 24]. Finally, a rolling-horizon framework is implemented allowing for new schedules to be generated as time progresses.

Chapter 3

Research Framework

This chapter provides motivation for the research and presents the research vision. Further, the limitations are specified and the research questions and methods are briefly introduced.

3.1 Motivation and research vision

As already stated in chapter 1 finding optimal railway timetables and shunting yard schedules are hard combinatorial optimisation problems. Despite this the planners are currently referred to solving these problems by hand, with little or no decision support. Even though the planners have a rigorous training and are very skilled at their job, there is reason to believe that providing them with a decision support tool could improve the infrastructure allocation process. By using a decision support tool the timetable and yard schedules could be generated much faster, allowing for a more flexible and adaptive planning process. Further, by using optimisation the efficiency of the timetables and schedules could be improved and also mathematically proven. The research vision is a flexible process that is supported by optimising computerised decision support systems that allow for consistent and efficient planning of all parts of the railway network, and a fair and effective allocation of infrastructure.

The research of this licentiate is carried out in an industrial setting. Industry research often aims at strengthening the industry at hand by discovering or examining new processes or technical solutions. Working in an industrial setting also means that there is a framework that the research has to relate to.

First of all, the methods presented should be suitable for use in the industry even if some adaptation and development work is to be expected before real implementation. A large number of mathematical models for timetabling (see e.g. the overview in [25]) and shunting yard planning (e.g. [20, 17, 21, 19]) have been developed in the past. However, these models are rarely put into real use, maybe because they are too theoretical [26]. Further, the models don't take Swedish operational practices into consideration. Developing models that are suitable for the Swedish railway industry is therefore a goal and also a requirement for the research presented in this thesis. Ensuring acceptable execution times is another important aspect when trying to achieve real world feasibility. Often a "good enough" solution that is returned quickly is more useful to the industry than an optimal solution that it takes a very long time to compute.

Despite the fact that the research results have to be applicable in the current industrial setting, an inherent consequence of implementing new solutions proposed by research is that processes and ways of working have to be changed. Therefore another important part of the research is to estimate and/or explain the gains/losses from the proposed solution. After all, not all methods evaluated may be suitable for implementation. This aspect has been important when formulating the research questions and planning the research.

3.2 Limitations

To reach the overall vision of a flexible planning process supported by optimising computerised decision support systems there is a large number of problems that need to be solved, ranging from how to prioritise between trains and maintenance works in the long-term process to assessing and deciding which timetable characteristics that gives a wanted result with respect to e.g. robustness and efficiency. This licentiate thesis does not aim to answer all of these questions but is focused on two in particular: *how to increase the replanning robustness of a timetable* and *how to generate a schedule with the minimum shunting effort for a classification bowl in a shunting yard.* The research focus is on mathematical methods that can be used to generate schedules, and to some extent the processes in which they may be used, but does not cover e.g. user interface issues or system design. Further, although some of the models are easy to adapt, they have been developed specifically for the Swedish railway system. Finally, even though care has been taken to ensure that the models respect the requirements and demands of the Swedish railways, individual specific requirements of e.g.

stations with rare track configurations have not been modelled. The aim has rather been to model the official rules and the level of detail generally used in the various planning stages.

3.3 Research questions

The research questions for the two chosen areas were formulated with the industry requirements in mind and therefore focus on two aspects: usability and potential.

Q1: Is it feasible to use a mathematical programming model to generate a timetable that provides replanning robustness?

Q2: Can an optimising mathematical programming model be used to construct a classification bowl schedule with minimum shunting effort?

- **a)** How effective are simple dispatching rules for constructing a classification bowl schedule with minimum shunting effort?
- **b)** Is it feasible to extend the optimisation model to include also the car order in outbound trains? How is the execution time affected? How is the shunting effort affected?
- **c)** Is rolling horizon planning a feasible method for generating shunting schedules for longer periods of time? How is the execution time affected?

3.4 Methodology

The research in this licentiate thesis builds on work previously presented in Forsgren et al. [27] for the robust timetable research, and work by Bohlin et al. [28, 29] for shunting yard planning. The research is applied and focused on the Swedish railway industry.

The work started with background reading and informal interviews with railway personnel in order to understand the real world scheduling problems and processes, and a literature study to gauge the state of art. Based on this understanding and knowledge optimisation models and sometimes heuristic algorithms were developed for the scheduling problem at hand. An inductive analysis of the chosen method's ability to return an optimal solution and of the execution times was made using data from test cases based on real data. As long as the results did not disqualify a method from being a suitable candidate to solve the problem at hand, it was often further changed and developed in an iterative manner to more effectively and accurately provide the functionalities required.

More precisely, Q1 was answered by developing a mathematical programming model and analysing its performance using real traffic data provided by Trafikverket in a proof of concept tool. Real data was also used for the geography and train characteristics. Q2 was answered by developing mathematical programming models and heuristic algorithms, and analysing their performance in a proof of concept tool and in a simulation. The tests used real traffic data provided by Trafikverket. The geographical constraints of Hallsberg shunting yard and an empirical investigation of the timings of different shunting tasks were available in a technical report by Averstad [30]. Also, the simulated delays were sampled from an empirical distribution provided in Lindfeldt[31]. However, when it comes to testing the new blocking model for Q2 b) the block data had to be simulated.

Chapter 4

Summary of included papers

This chapter provides a brief summary of all included papers and explains how the papers answer the research questions. My contribution to the papers is also identified.

4.1 Summary of Paper A

As stated in Chapter 2 changing the train order is one of few recourse actions available to dispatchers, and by increasing the number of alternative meeting locations for trains the robustness of the system could be increased as well.

In Paper A a method for increasing the number of alternative meeting locations between two timetable solutions is developed. The method is based on generating two feasible timetables: one that is to be used as the operational timetable and one that is used as a “safety-net”. Both timetables fulfil all delivery commitments and hence the efficiency of the timetable is not decreased. The distance between the two timetables is maximised, either in terms of time or in terms of number of alternative meeting locations. The replanning robustness of a timetable is decided both by the number of alternative meeting locations and by how much time that can be saved by moving a meeting, and an objective function that combines the two measures is therefore likely to provide the most robustness. The method is tested on a single track section between Boden and Vännäs, and as expected the two objective functions return different solutions, indicating that devising and testing a robustness measure including both time and number of meeting locations is indeed desirable.

Paper A answers to research question Q1 and shows that for small track sections it is feasible to use mathematical programming to generate a timetable that provides replanning robustness, although an exact measure for replanning robustness is not devised. The presented method visualises the available replanning space and increases it. The execution times of the model is below 5 minutes when solving our case-study example on a Thinkpad T430S.

For Paper A I co-developed the explanatory framework of the replanning robustness idea with the other authors. I also wrote the entire paper but from the related work section and presented the paper at the conference.

4.2 Summary of Papers B-E

In Bohlin et al. [23, 24] the Swedish shunting yard operational practices are introduced, and optimisation models for generating a classification bowl schedules with minimum shunting effort are introduced. However, the models presented fail to return an optimal schedule within an acceptable execution time for some test-instances.

In Paper B a new extended integer programming formulation (CG-IP) for the classification bowl problem is introduced. This new model can be solved using column generation, and successfully returns a proven optimal schedule for all test-cases within an acceptable execution time. In Paper E yet another integer programming model (AI-IP) is presented. AI-IP is a flow-based model that returns optimal results within an acceptable time and is easier to implement than CG-IP. Both CG-IP and AI-IP are tested on real-life traffic data from Hallsberg shunting yard, the largest shunting yard in Sweden. The results of Papers B and E show that it is feasible to generate classification bowl schedules with minimum shunting effort using mathematical programming, and thereby answer research question Q2.

In Paper C we investigate if simple planning rules can be used to generate a classification bowl schedule. The rules are tested in both a deterministic and a stochastic simulation and are compared with the optimal schedule generated by the CG-IP model presented in Paper B. The results show that simple planning rules are not adequate for scheduling the classification bowl as the schedules returned are often infeasible. Further, the shunting effort required in the schedules generated by the simple rules was generally at least 5 times the shunting effort of the optimal schedule. The simple planning rules tested are therefore judged to be ineffective for planning the classification bowl, and the paper provides an answer to Q2 a).

Sometimes the cars should be sorted according to drop-off location in the outbound trains at shunting yards. We call such outbound trains *blocked trains*, and a block is a set of cars with the same drop-off location. In Paper D the classification bowl planning framework is extended for CG-IP to sort the cars not only based on their outbound trains, but also based on which block they belong to. The framework contains new methods for calculating feasible sequences and costs that respect the car-order of the outbound trains. Further, as no real block data was available a simulation model for generating feasible and realistic blocks was developed and implemented. The extra shunting effort required for sorting cars according to blocks is analysed and the results show that although the total shunting effort inevitably increases as more sorting is needed, the extra effort caused by yard capacity limitations does not seem to be affected as much. The new model is solved equally fast as the previous model that excludes sorting by blocks. This paper thereby answers Q2 b).

In Paper E a rolling horizon planning framework is developed and implemented for CG-IP and AI-IP. The results show that while the execution times are slightly increased with rolling horizon planning they do not become unacceptably long. AI-IP in particular remains a very fast method and always return an optimal solution within 6 minutes. Paper E thereby answers question Q2 c).

For Papers B and E I implemented parts of the test frame-work, ran the experiments, analysed the results, and wrote the Experiments and Results sections. For Paper E I also developed and implemented the rolling horizon planning framework. When it comes to Paper C me and Ms. Khoshniyat developed the simulation together and we also co-operated closely when writing the paper. Finally, for Paper D I developed the method used for planning with blocked trains, implemented the code changes needed, developed and implemented the block-data simulation, ran the experiments and wrote and presented the paper.

Chapter 5

Conclusion

This chapter provides a summary of the most important results and also discusses future work.

5.1 Summary of results

This thesis presents models that could be used in future decision support systems for planning in the Swedish railway industry.

Paper A presents a method that can be used to generate timetables with replanning robustness. The proposed method does not guarantee that the timetable is optimal with respect to replanning robustness as an exact measure of replanning robustness is not proposed, but the method does return timetables that have some degree of replanning robustness and it provides a visualisation of the replanning possibilities available to the dispatchers. The execution times for the test cases are less than 5 minutes on a Thinkpad T430S.

Papers B-E present various methods for generating a shunting yard classification bowl schedule with minimum shunting effort. Three optimisation models and some simple rules are tested. The results from the test cases show that the simple rules are inadequate for planning the classification bowl. Even the best rule results in infeasible schedules in 1 (6) out of the 21 periods for the static (stochastic) simulation test. Further, in general the best rule generates schedules with 5 times the optimal shunting effort.

When it comes to optimising methods the first direct model, D-IP (originally presented in Bohlin et al. [24]), finds feasible schedules for most separate test instances within the execution time limit of 20 minutes, but often fails to prove optimality. It performs better than the simple rules. The two other optimisation models, CG-IP and AI-IP, always prove optimality within 18 and 8 minutes respectively for separate problem instances. As AI-IP is easier to implement than CG-IP it is the suggested method.

The CG-IP model was extended to include sorting by block, i.e. sorting the cars in the outbound train depending on which block they belong to. For this model and a planning horizon of 3 days an optimal schedule was always found within 3 minutes for all 50 separate problem instances. An increase in shunting effort is unavoidable as more sorting will be required when the cars' block associations are to be respected as well. However, the extra shunting movements enforced by capacity constraints is comparable to the base case without block sorting.

A rolling horizon framework was developed for both the CG-IP and the AI-IP model. When using rolling horizon planning the execution times for the use-case were generally slightly increased but there is no indication that rolling horizon planning should be infeasible for real world implementation.

5.2 Future work

The licentiate thesis is mainly focused on shunting yard planning while the chosen research area includes both line and yard planning. When it comes to line planning a research question that we are currently investigating is how to decide on *delivery commitments*. A future planning process, *Incremental Allocation (IA)*, is envisioned where delivery commitments rather than the entire timetable are published once a year, and where the exact train paths may vary from day to day. However, a reliable and fast process for extracting such delivery commitments from the capacity applications has yet to be developed. The problem is hard as the scheduling problem defined by the delivery commitments should be solvable for all operation days. Also, there will most likely be instances when not all train operating companies can be granted the delivery commitments originally applied for, in which case a method must be developed for deciding which commitments to make and which ones to change. This process should be competitive neutral and prioritise trains with high benefits to society in case of unsolvable conflicts.

Developing the replanning robustness model further and examining how the time and number of meeting locations should be weighted in the objective function is also an area where more research is needed. Testing if the timetables are more robust with respect to delivery commitments in a stochastic simulation would also be interesting.

There are also many interesting questions left for shunting yard planning. For example, currently the arrival and departure yard schedules, as well as the hump-schedule, are planned heuristically in our proposed solution. It would be interesting to develop a more rigorous approach and investigate if the various sub-yard scheduling problems could be solved iteratively to arrive at a near-optimal solution for the entire shunting yard. Including robustness in the shunting yard planning framework, and developing more detailed mixing track modelling, would also be beneficial. Better heuristics and more detailed mixing track modelling is particularly important for the model that respect blocks, as the heuristics were not adapted for this model and more mixing is required.

The interface between yard planning and line planning in an IA setting is also an area for future research. Currently, the shunting yards are planned on a daily basis which is in line with the short-term planning process of IA. However, there is no process for ensuring that the yard capacity allows for granting a set of delivery commitments in the long-term planning process. Finally, as cars often have some buffer time at yards it would be interesting to investigate how this time could be used when re-planning in case of e.g. traffic disturbances.

Bibliography

[1] Trafikverket. Fastställd tågplan 2015 TRV 2013/76110. `http://www.trafikverket.se/PageFiles/159824/faststalld_tagplan_2015.pdf` (accessed online 11-03-15), 2014.

[2] Jonas Gustafsson, Carin Rosell, and Camilla Fredriksson. Att skapa tidtabeller för tåg - nu och i framtiden ver 2. `http://www.trafikverket.se/PageFiles/150114/broschyr_skapa_tidtabeller_tag_nov_2014.pdf` (accessed online 11-03-15), 2014.

[3] Emma Andersson, Anders Peterson, and Johanna Törnquist Krasemann. Robustness in Swedish railway traffic timetables. In *Railrome 2011: Book of Abstracts 4th International Seminar on Railway Operations Modelling and Analysis*, 2011.

[4] Leo G. Kroon, Rommert Dekker, and Michiel J.C.M. Vromans. Cyclic railway timetabling: A stochastic optimization approach. In Frank Geraets, Leo Kroon, Anita Schoebel, Dorothea Wagner, and Christos D. Zaroliagis, editors, *Algorithmic Methods for Railway Optimization*, volume 4359 of *Lecture Notes in Computer Science*, pages 41–66. Springer Berlin Heidelberg, 2007.

[5] Christian Liebchen, Michael Schachtebeck, Anita Schöbel, Sebastian Stiller, and André Prigge. Computing delay resistant railway timetables. *Computers & Operations Research*, 37(5):857–868, 2010.

[6] L. Kroon, D. Huisman, and G. Marti. Optimisation models for railway timetabling. In I.A. Hansen and J. Pachl, editors, *Railway Timetable & Traffic*. Eurail Press, Hamburg, 2008.

[7] V. Cacchiani and P. Toth. Nominal and robust train timetabling problems. *European Journal of Operational Research*, 219(3):727–737, 2012.

[8] D. Bertsimas and M. Sim. The price of robustness. *Operations research*, 52(1):35–53, 2004.

[9] A. L. Soyster. Technical note - convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations research*, 21(5):1154–1157, 1973.

[10] Aharon Ben-Tal and Arkadi Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical programming*, 88(3):411–424, 2000.

[11] M Fischetti and M. Monaci. Light robustness. Technical Report ARRIVAL-TR-0019, ARRIVAL PROJECT, 2008.

[12] Matteo Fischetti, Domenico Salvagnin, and Arrigo Zanette. Fast approaches to improve the robustness of a railway timetable. *Transportation Science*, 43(3):321–335, 2009.

[13] LeoG. Kroon, Rommert Dekker, and MichielJ.C.M. Vromans. Cyclic railway timetabling: A stochastic optimization approach. In Frank Geraets, Leo Kroon, Anita Schoebel, Dorothea Wagner, and ChristosD. Zaroliagis, editors, *Algorithmic Methods for Railway Optimization*, volume 4359 of *Lecture Notes in Computer Science*, pages 41–66. Springer Berlin Heidelberg, 2007.

[14] Christian Liebchen, Marco Lübbecke, Rolf Möhring, and Sebastian Stiller. The concept of recoverable robustness, linear programming recovery, and railway applications. In *Robust and online large-scale optimization*, pages 1–27. Springer, 2009.

[15] Ronny S. Hansmann and Uwe T. Zimmermann. Optimal sorting of rolling stock at hump yards. In Hans-Joachim Krebs and Willi Jager, editors, *Mathematics Key Technology for the Future*, pages 189–203. Springer Berlin Heidelberg, 2008.

[16] Gabriele Di Stefano, Jens Maue, Maciej Modelski, Alfredo Navarra, Marc Nunkesser, and John van den Broek. Models for rearranging train cars. Technical Report ARRIVAL-TR-0089, ARRIVAL Project, November 2007.

[17] Michael Gatto, Jens Maue, Matús Mihalák, and Peter Widmayer. Shunting for dummies: An introductory algorithmic survey. In Ravindra Ahuja, Rolf Mhring, and Christos Zaroliagis, editors, *Robust and Online Large-Scale Optimization*, volume 5868 of *Lecture Notes in Computer Science*, pages 310–337. Springer Berlin / Heidelberg, 2009.

[18] Nils Boysen, Malte Fliedner, Florian Jaehn, and Erwin Pesch. Shunting yard operations: Theoretical aspects and applications. *European Journal of Operational Research*, 220(1):1 – 14, 2012.

[19] Riko Jacob, Peter Márton, Jens Maue, and Marc Nunkesser. Multistage methods for freight train classification. *Networks*, 57(1):87–105, 2011.

[20] Elias Dahlhaus, Fredrik Manne, Mirka Miller, and Joe Ryan. Algorithms for combinatorial problems related to train marshalling. In *Proceedings of the Eleventh Australasian Workshop on Combinatorial Algorithms (AWOCA)*, pages 7–16, 2000.

[21] Maue Jens and Marc Nunkesser. Evaluation of computational methods for freight train classification schedules. Technical report, ARRIVAL Project, 2009.

[22] Peter Márton, Jens Maue, and Marc Nunkesser. An Improved Train Classification Procedure for the Hump Yard Lausanne Triage. In Jens Clausen and Gabriele Di Stefano, editors, *9th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'09)*, volume 12 of *OpenAccess Series in Informatics (OASIcs)*, Dagstuhl, Germany, 2009. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[23] Markus Bohlin, Holger Flier, Jens Maue, and Matúš Mihalák. Hump Yard Track Allocation with Temporary Car Storage. In *The 4th International Seminar on Railway Operations Modelling and Analysis (RailRome)*, 2011. Available on `http://soda.swedish-ict.se/5089/`.

[24] Markus Bohlin, Holger Flier, Jens Maue, and Matús Mihalák. Track Allocation in Freight-Train Classification with Mixed Tracks. In *11^{th} Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, volume 20 of *OpenAccess Series in Informatics (OASIcs)*, pages 38–51, Dagstuhl, Germany, 2011. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[25] Steven S Harrod. A tutorial on fundamental model structures for railway timetable optimization. *Surveys in Operations Research and Management Science*, 17(2):85–96, 2012.

[26] A. K. Nemani and R. K. Ahuja. OR models in freight railroad industry. In James J. Cochran, editor, *Wiley Encyclopedia of Operations Research and Management Science*. 2011.

[27] M. Forsgren, M. Aronsson, P. Kreuger, and H. Dahlberg. The maraca- a tool for minimizing resource conflicts in a non-periodic timetable. In *The 4th International Seminar on Railway Operations Modelling and Analysis (RailRome)*, 2011.

[28] Markus Bohlin, Holger Flier, Jens Maue, and Matúš Mihalák. Hump Yard Track Allocation with Temporary Car Storage. In *The 4th International Seminar on Railway Operations Modelling and Analysis (RailRome)*, 2011.

[29] Markus Bohlin, Holger Flier, Jens Maue, and Matús Mihalák. Track Allocation in Freight-Train Classification with Mixed Tracks. In Alberto Caprara and Spyros Kontogiannis, editors, *11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, volume 20 of *OpenAccess Series in Informatics (OASIcs)*, pages 38–51, Dagstuhl, Germany, 2011. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[30] K.-Å. Averstad. *Handbok BRÖ 05-35/BA50: Trafikeringsplan Hallsberg rangerbangård*. Banverket, Maj 2006. In Swedish.

[31] A. Lindfeldt. A study of the performance and utilization of the swedish railway network. In *First International Conference on Road and Rail Infrastructure*, Opatija, Croatia, 2010.