"Soms moet je het eerst moeilijker maken voor je eruit kunt komen"

The key in solving the Rolling Stock Circulation

Business Mathematics & Informatics
Vrije Universiteit, Amsterdam

Author: *Chi Hung Mok*
Supervisor: *Sandjai Bhulai*

November 2007

"Soms moet je het eerst moeilijker maken voor je eruit kunt komen"

The key in solving the Rolling Stock Circulation

Author:

Chi Hung Mok (1333267)

Supervisor:

Dr. Sandjai Bhulai

Study:

Business Mathematics & Informatics
Vrije Universiteit
Faculty of Sciences

November 2007

Picture on the title page is taken from the webpage:
http://www.haarlemmermeerspoor.nl/willemuithoorn.htm

Preface

This document is a so called "BWI Werkstuk"; a literature study on a subject within the scope of the study Business Mathematics & Informatics (BMI).

The "BWI Werkstuk" has often been the final hurdle known to the majority of BMI students when they approach their graduation. In order to prevent this from happening to me, I planned to finish this assignment prior to my internship, which would start in March of 2006. Due to several unforeseen reasons I had to start my internship without finishing this document and thus I have joined the "majority" at that moment.

The reason why I chose this subject is due to the new timetable introduced by the NS and an article written in *de Volkskrant*. The article (given in the appendix) was about Alexander Schrijver who solved a very difficult problem for the NS. His solution was to add variables to the model instead of only removing and/or relaxing constraints, which really caught my interest. It is amazing how such a 'simple' operation can make such a big difference.

I want to thank my supervisor Sandjai Bhulai for attending me on this interesting subject and his encouragement during the whole process. In spite of his busy schedule, he always makes for me when I needed to speak to him.

Furthermore, I want to thank Angelique Mak, who was a great support during the writing of this document. And of course other fellow student colleagues for being around for a quick talk at moments when the writing began to become unbearable. In random order: Gitman, Paul, Christel, Jade, Natasia, Aicha, and Suman. Last but not least, Alicia. She is not around much, but it was always nice when she did.

Summary

The NS is the largest railway company in the Netherlands and transports over a million passengers on a regular workday. Naturally this requires a huge amount of planning in advance to ensure that the passengers reach their destinations within acceptable travel durations. It is rather difficult to analyze this large problem at once, therefore the problem is split into three smaller, yet still very complicated, subproblems:

- Construction of the time table
- Rolling stock circulation
- Crew planning

The crew planning problem is a very interesting subject, it describes how the crew of the NS needs to be scheduled to ensure the two other planning processes (timetable, rolling stock circulation) will be carried out as planned. However, this subject is not within the scope of this document and will therefore not extensively explained as the other two problems. A short introduction to the Crew planning problem and recommended literature are given in its respective chapter.

The first problem is probably also the most well known problem. The time table describes when a train departs and arrives at a station, given that the train network is known beforehand. The forecast of the number of passengers to be transported at each station is a requirement too; the number of trains arriving at and departing from a central station of a big city is different compared to the number of a little station of a town, this is due to the difference in number of passengers at each station.

Using this information, the NS can use the DONS system to construct a feasible time table. The first step is to model the network in the program environment. The CADANS module will then investigate if it is possible to construct a feasible timetable given the restrictions and characteristics of the network. If it is possible, a timetable is stored in the database module of the system, the same place where the model of the network is stored. It might occur that the system says that the network is constructed in such a way that it is impossible to make a feasible timetable. CADANS provides a list of conflicting restrictions which have caused this conclusion. This enables the user of the system to change the restrictions in order to create a network which has a feasible solution.
The next step within the DONS system requires more detailed information with respect to the stations. While CADANS routes the train between stations in a feasible way, the module STATIONS routes specific trains within each station to a specific platform. In this step, it is also possible to conclude that given the timetable from CADANS and the user preferences that it is not possible to find a feasible solution. If a feasible solution does exist, this is stored in the database.

The DONS system is by all means not an optimizer. Its primary function is to give a feasible solution. The user can use this solution as a starting point to improve the initial solution. This can be done by using different methods. Two methods used by the NS are the timetable evaluation tool PETER (based on Max-Plus Algebra) and SIMONE (a simulation program) to evaluate the quality of the railway tables obtained from the DONS system.

When the desired timetable is constructed, it is time to move on to the next problem; the rolling stock circulation. The solution to this problem denotes the precise composition to be used on each route described by the timetable. This may not seem to be a very difficult task, but until recently it was not possible to accurately solve this problem due to its large number of variables and constraints. It turns out that by adding a variable to the model, the problem becomes less complicated and even possible to solve. The 'trick' is to represent the train compositions in a different way; the addition of this extra variable eliminates many other variables in the whole model and ultimately makes the model solvable.
The solution is very powerful and practical. It is possible to pursue different solutions by adjusting the objective function accordingly; the described model used the following criteria in the objective function: the *number of carriage kilometers*, the *number of seat shortage* and the *number of shunting movements*. The time needed to construct a feasible solution is very reasonable; for a representative real-life example a solution was found within 2 hours and the solution was very close to the optimal solution. The man behind all this is prof. dr. Alexander Schrijver, who has been working closely with the NS for many years now. In 2002 he suddenly came to this solution method while he was enjoying his Christmas Holidays. The NS was convinced of the solution and since 2005 the NS is using this model to plan its rolling stock circulation.

Content

1. Introduction

On December 10th 2006 the Nederlandse Spoorwegen, NS (the Dutch Railways) introduced a new timetable. This change is since the *Spoorslag '70* in 1970 the most rigorous change made in the existing timetable in the Dutch railway history.

This change was needed due to several reasons. The number of passengers had increased incredibly since the last change and although the NS still managed to provide adequate service levels, i.e., punctuality, it would be difficult to maintain these service levels in the future using the existing timetable when the number of passengers increases even more. And at that moment, the railway network was also expecting two important new railway lines; the High Speed Line (de Hogesnelheidslijn, HSL) and the Betuweroute, the new timetable takes these changes into account as well.

The new timetable ensures better use of equipment, provides a more aligned structure regarding stopovers, accounts for better connections with Germany and Belgium, etc. But why where there no major changes made after 1970? Why wait 36 years? The main reason is the complexity of the problem. There are many variables in this problem with even more dependencies between them which means that a major change like this needs a huge amount of coordination, effort, and time.

The NS has three major planning problems for its daily operations:

- Construction of the timetable
- Rolling Stock Circulation
- Crew Planning

The timetable describes which lines are needed and at what times the trains should arrive and depart from the stations, which as mentioned before, is a very complicated problem. It does not however, account for the number of trains available. Each line has to transport some number of passengers, which will determine to a large extent how much capacity the operating train should have. But due to the limited number of trains available, different types of trains and the fact that the trains cannot be transported to the required place instantaneously, the *rolling stock circulation* is a very complex problem on its own. Prof. dr. Alexander Schrijver found a sophisticated method to reduce the complexity of this problem and provides a model to solve this problem. But can this method also be used in other not easily solvable problems?

This document starts with some general information on the Dutch Railways to provide some insights into the size of the problem, followed by the procedure needed to construct a feasible timetable. The main focus of this document is the rolling stock circulation. And of course the method introduced by Alexander Schrijver will be treated.

The last planning problem is not treated as detailed as the first two problems, because the real focus of this document is on the second problem, which is heavily influenced by the first planning problem. This document gives a brief introduction to this problem and refers to other papers for more specific information regarding this subject.

2. Nederlandse Spoorwegen

The Nederlandse Spoorwegen (NS) is the largest railway company in the Netherlands. It is the result of a merger between the Hollandsche IJzeren Spoorweg-Maatschappij (HSM) and the Maatschappij tot Exploitatie van Staatsspoorwegen (SS) in 1938.These two companies had been working closely together since 1917, but due to the declining economic situation in the Netherlands as result of World War I, a merge was necessary.

The government acknowledges the importance of a good railway network and thus supported the NS by buying their shares. Although the Dutch government is the sole shareholder, the NS remained a private company.

Below are some figures regarding the operations of the NS as of today:

1.100.000	passenger trips per workday
15.000.000.000	passenger kilometers per year
5200	passenger trains per workday
300	cargo trains per workday
2800	kilometers of tracks
377	stations

Table 1: Present operation figures of the NS (Huisman[1])

2.1. *Spoorslag '70*

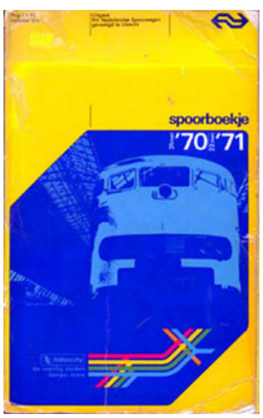

Figure 1: Timetable 1970-1971

In the period 1960-1980 the Netherlands experienced a huge increase in the number of automobiles. This form of transportation was until that period only accessible to the wealthier part of the Dutch population, but due to the great surge in wealth, the automobile was in reach of more people. This development was disastrous for the NS as the growth of passengers stagnated and the goods transportation with the train declined. This in combination with the increasing labor costs, the NS was on the verge of bankruptcy. The NS could have evaluated its network and reject the bad and keep the good performing routes. But this would have only delayed the bankruptcy, because the NS would end up with a smaller train network and as a result attracting even fewer passengers.

The Dutch government did not want to give up on the train network for several reasons. One of the reasons was that the government could not solely rely on auto transport, if

[1] Timetable 2007: Why and How? (2007)

something were to happen to the roads, it would be impossible to travel. And on top of that, the train was invaluable to the citizens who could not afford a car. With the financial help of the government the NS was saved, but in return the NS has to drastically change its service and show to the government that they could be profitable again.

This drastic change was the timetable Spoorslag '70. To attract more passengers the NS needed to offer a significantly better product than before. Some requirements of the Spoorslag '70 are listed below.

- Fixed departure time
- Consistent and improved connections
- Symmetric lines (traffic going one way follows the same way back)
- High frequency of *stoptreinen* (local trains)
- Network of Intercity trains

The combination of Intercity trains and local trains were a success due to the improved connectivity. The NS could provide a better product and passengers were willing to take the train again; in the period 1970-1980 the NS experienced an increase in the number of passengers of 20%. Although the costs went up, the extra traffic generated by the new network covered a lot of it. The NS had introduced other changes to the network in the following years, but the Spoorslag '70 was the most drastic step the NS had made and therefore still the most well known.

2.2. *Timetable 2007*

Since the Spoorslag'70, the network gained a large number of stations. The number of trains on the tracks increased as well; from 4000 trains per day to 5000 trains per day. This was all to cope with the number of passengers that the NS had to transport each day; well over 1 million passenger trips. The timetable, however, did not undergo any major changes. Soon it was clear that the timetable created by the Spoorslag '70 was outdated and not suitable to deal with the scale of the network known today; the punctuality of the trains was declining very quickly and this would only get worse due to the forecasted increase in passengers.
To improve the punctuality both on the short and long term it was needed to change the timetable drastically. This change had a very great impact on the existing timetable and therefore the NS announced the new timetable already on January 8th 2006, almost a year before its start on December 10th 2006, so people could prepare for it.

Timetable 2007 features many changes compared to the old timetable. Stopover times are increased which makes the network more robust to disturbances like failing trains or malfunctioning railway switches. A simple example of the robustness is that the timetable reduces the occurrences of passengers missing a connecting train because its first train was delayed. The new timetable is accompanied with several new stations as well. The overall the punctuality is forecasted to increase with 1.5%. However, the new timetable does not only have positive impacts; 52% of the passengers will need to travel longer to reach their destination due to these increased stopover times. The network has changed so that some passengers might have to loose their direct connection to their destinations and are forced to use a connecting route. Some stations are demoted from Intercity stations to normal stations which increases the travel times as well.

Other changes are, for example, the reduction of types of trains. The fast trains (sneltreinen) and intercity trains are grouped as intercity trains. The stop trains will are replaced by the Sprinter.

The new timetable aligns more with the train connections from the neighboring countries Germany and Belgium. And at the moment of development, the railway network was also expecting two important new lines; the High Speed Line (de Hogesnelheidslijn, HSL) and the Betuweroute, the new timetable takes these changes into account as well.

3. Constructing a feasible timetable

To construct a timetable, the first step is to determine which variables and restrictions are in play and then it might be possible to create a feasible solution using this information. The solution will then be further optimized to come to a final solution which can be implemented.

3.1. Designer Of Network Schedules

The Designer Of Network Schedules or DONS is a software program especially developed for the NS and ProRail (the infrastructure manager). The DONS can, given the necessary information, create feasible timetable constructions. Important to note is that the DONS does not necessarily give the optimal solution. It provides a feasible solution if there is any which can be used to find a better timetable, or it tells the user that given the input no feasible timetable can be constructed. If the latter is true, the DONS will also provide some information regarding the conflicting constraints.

An important feature in practical timetables is the cyclic nature. This means that the timetable will repeat itself in the next period; commonly this period is equal to an hour. Creating timetables of this nature are referred to as Periodic Event Scheduler Problems (PESP). The DONS uses this principle as well and features periods of one or two hours.

The DONS consists of four integrated modules; each part plays a vital role in the determination of a feasible timetable.

3.1.1. Database

The database is the central information storage of DONS. Every piece of information regarding the restrictions of the problem is stored here. But also constructed timetables are stored here and can be approached by the GUI or external processes like the simulation program SIMONE.

3.1.2. Graphical User Interface

The interaction between the program and user is via the Graphical User Interface (GUI). The user uses the GUI to define the infrastructure project which will be used to construct a timetable. Depending on the type of information, the GUI will utilize different representations of the timetable. The network and the nodes for example will be shown in a graphical manner while many timetable characteristics will be shown in a tabular form. This enables the user to quickly and easily alter restrictions of a timetable. The GUI stores and retrieves information from the database.

3.1.3. CADANS-solver

The CADANS (Combinatorisch Algebraïsch Dienstregeling Algoritme voor de Nederlandse Spoorwegen, translated in English: Combinatorial Algebraic Timetable Algorithm for the Duth Railways) is written by Schrijver and Steenbeek in 1994. The DONS consists of two solvers and the CADANS is the first one. Depending on the stored information regarding the restrictions, the CADANS-solver will output a feasible timetable or indicate that a feasible solution could not be found and the reasons for that conclusion, i.e., conflicting restrictions.

The timetable consists of arrival and departure times for all timetable points. Note that the CADANS does not need very detailed information regarding the infrastructure to construct a feasible timetable; defining the tracks that enter and leave a timetable point is enough.

3.1.4. STATIONS-solver

Given the solution found by the CADANS, the STATIONS-solver will regulate the trains within a station on a platform level. This means that individual trains are now assigned to a specific platform by considering all the tracks available in the station. The STATION-solver will indicate if a feasible solution is possible. Naturally a detailed track and platform layout for each station is needed to perform this operation.

3.2. Evaluation of timetables

Timetables constructed by the DONS system are evaluated and further optimized. The NS uses two different approaches to evaluate timetables:

- Max-Plus Algebra
- Simulation

The following chapters will explain the methods and highlight what their advantages and disadvantages are.

3.2.1. Max-Plus Algebra

A train network can be modeled as a *Petri Net*, a precedence graph. These kinds of graphs describe a system where certain events can not be activated before some other events are finished; like the situation where another train cannot arrive at a station before the preceeding train has left that station.

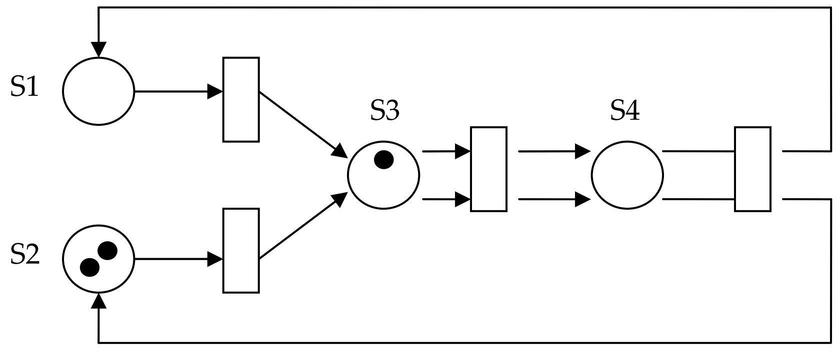

Figure 2: Example of a simple Petri Net

Figure 2 shows a simple train network modeled as a Petri Net. Let *S1*, *S2*, *S3,* and *S4* be train stations. Station *S2* currently has two trains, while station *S3* has one. In this situation station *S2* can send one train to *S3*, but *S3* cannot send its train to *S4*. The network enforces it to send two trains simultaneously to *S4* (combining of trains); the train at station *S3* must wait for the train sent by *S2*.

When a Petri Net includes the time factor, this graph is called a *timed event graph* (TEG). TEGs are very suitable to model stochastic systems and to be able to analyze them the Max-Plus Algebra is developed. This algebra only features two operations: "taking the maximum of a collection" and "addition". The notation $a \oplus b$ is equivalent to $\max\{a, b\}$ and $a \otimes b$ is equivalent to $a + b$. Max-Plus Algebra operations are also applicable on matrices, which is needed to analyze TEGs.

PETER (Performance Evaluation of Timed Events in Railways) is used to evaluate and compare railway networks based on Max-Plus Algebra. The great advantage of this evaluation method is that it enables the user to obtain some network performance indicators regarding the capacity and stability of the timetable. These indicators can be used to compare different timetables with each other. The disadvantage of this approach is that the user does not have full control on which part of the network to test for. It is, for example, possible to introduce disturbance in the network to evaluate how the system reacts to it, but it is not possible to introduce random disturbances or add any disturbances during the evaluation period.

3.2.2. Simulation

Simulation is a versatile evaluation method often used to analyze complicated (stochastic) systems. The main difference between a simulation method and a conventional analytical model is that a simulation can deal with systems regardless of its complexity, but note that the simulation time needed to produce reliable results will increase greatly as the complexity of the system increases. To be able to utilize an analytical model, one must remove all the details from the system which are not compatible with the model and naturally the resulting model will therefore differ from the reality.

To further optimize the solution given by the DONS system, ProRail and Incontrol Enterprise Dynamics developed a simulation program which can reveal the strengths and weaknesses of a timetable. Using this knowledge it is possible to improve the solution. This simulation program is called the SImulation MOdel for NEtworks (SIMONE). This program is used by ProRail, the Dutch railway infrastructure capacity planner, and the NS self.

The main disadvantage is that simulation is used purely for evaluation purposes. It will not show to the user what a good timetable looks like or give indicators on how to improve it. It is up to the user to regard the results presented by the simulation and to decide which improvements are needed to optimize the timetable. A simulation program can deal with complicated systems, but at the same time it will require a lot of work to produce such a program. Knowledge on both the real-life railway network and programming expertise are needed to develop a correct simulation program.

3.3. Summary

In order to create a feasible timetable different steps are needed. First, the network is modeled as a set of restrictions. The DONS system uses these restrictions and user preferences to create a conflict free timetable. This timetable is most likely not a practical solution, but it is a very important base solution which can be further optimized by evaluating this solution. Where are the bottlenecks? What is the shortest time needed to complete the longest route? How much disturbance can the timetable handle? By solving these issues a significantly improved timetable is within reach.

Two evaluation methods are shown in this chapter, an analytical model based on Max-Plus Algebra called PETER and a simulation program SIMONE. Both methods are used by ProRail and the NS to evaluate and to improve the railway network.

4. Rolling stock circulation

Chapter 3 gave a brief overview on how NS creates its timetable. But this process is mainly focused on the infrastructure of the problem. There are still two problems that need to be solved for the timetable to be a success. Rolling stock and personnel are needed to perform the schedule, but how must they be scheduled? This chapter describes how the rolling stock circulation is done.
The rolling stock circulation is a difficult subject due to the different available train types. Therefore the exact configuration of the trains must be tracked to perform the correct combining and splitting of trains.

4.1. Problem description

We describe a real life network in the Netherlands called the Noord-Oost (North-East) case which is taken from Fioole en al [1]. This network connects the cities in the Western part to the Northern and Eastern part of the country. This is the list of cities:

Western part
- Amsterdam (Asd)
- Schiphol (Shl)
- Rotterdam (Rtd)
- The Hague (Gvc)

Northern part
- Leeuwarden (Lw)
- Groningen (Gn)
- Enschede (Es)

Eastern part
- Utrecht (Ut)
- Amersfoort (Amf)
- Deventer (Dv)
- Zwolle (Zl)

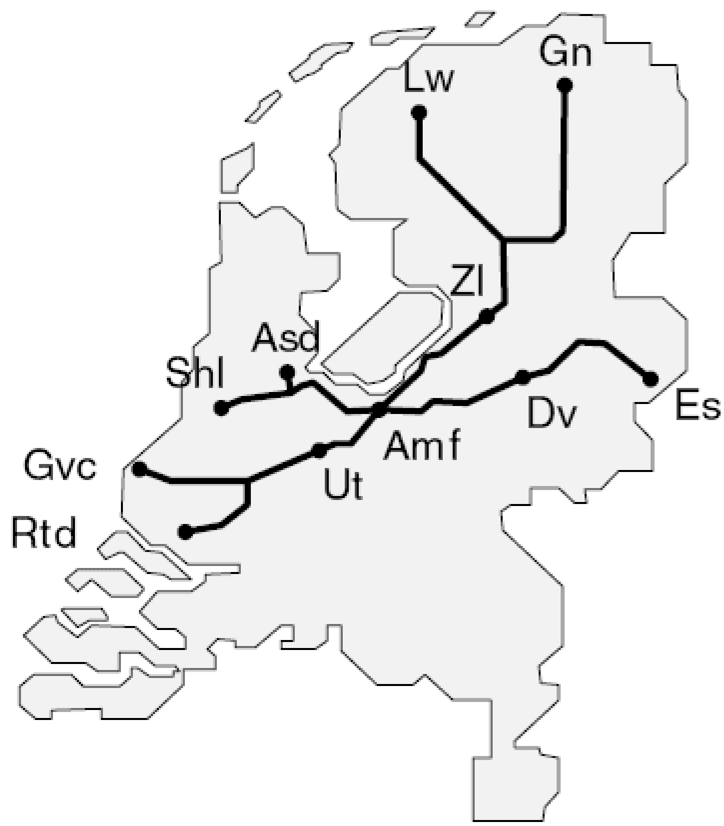

Figure 3: Noord-Oost situation (Picture taken from Fioole et al.[2])

The trains consist of a number of self-propelled trains that have three or four carriages each. The size of the platform and the forecasted number of passengers usually determine the length of a train.

Figure 4: Train with three carriages (Picture taken from Alfieri et al.[3])

Due to the limited time and space available at stations only small scale changes to the composition of the trains are possible; usually coupling and uncoupling of train units at the desired side of the train. For example, composition '343' can be changed to '34' by decoupling the last train unit '3'. The transition to '33' is not possible, because this will involve more than one (un)coupling operation.

At some stations trains are split or combined. This means that two trains are connected to each other or separated from each other, respectively. Regard the following situations:

[2] A rolling stock circulation model for combining and splitting of passenger trains (2005)
[3] Efficient Circulation of Railway Rolling Stock (2006)

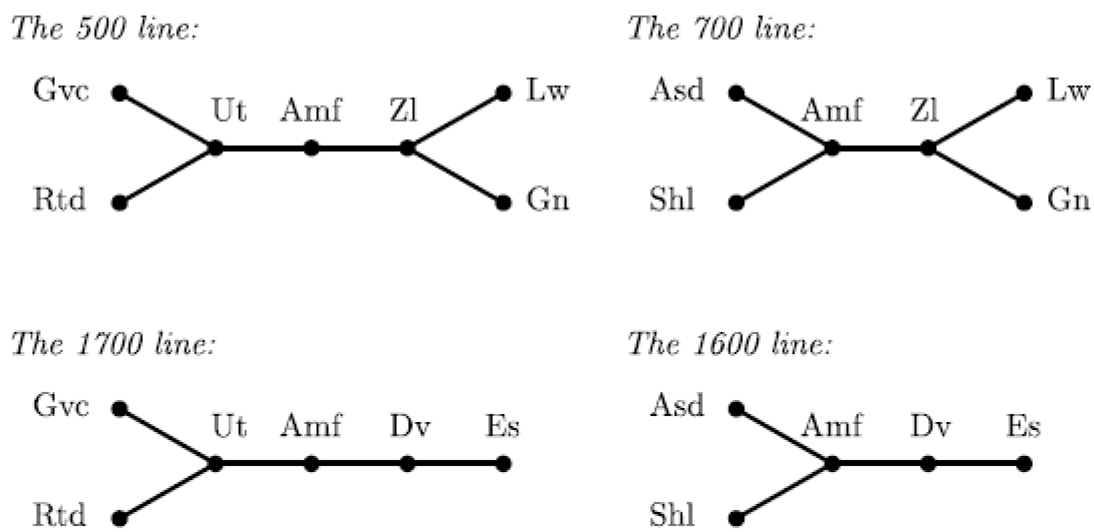

Figure 5: Noord-Oost lines (Picture taken from Fioole et al.[4])

Trains are combined in the 500 line at Ut; the train from Gvc and Rtd are joined. The train is split again in Zl, where one train heads to Lw and the other to Gn.

4.2. *Model formulation*

To understand the model, we need to know some of the basic notation used in this document.

Notation:

S	Set of stations
M	Set of train types
n_m	Number of available train units of type $m \in M$
P_t	Set of compositions p allowed on trip t
$n_{p,m}$	Number of train units of type m in composition p
$\beta(p)$	Vector of $n_{m,p}$, where $\beta(p)_m = n_{m,p}$ where $m \in M$
T	Set of t trips
$s_d(t)$	Departure station of trip t
$s_a(t)$	Arrival station of trip t
$\tau_d(t)$	Departure time of trip t
$\tau_a(t)$	Arrival time of trip t
$v(t)$	The successor trip of trip t, the train unit which performed trip t, immediately performs trip $v(t)$. (De)coupling is allowed.
T_0	Set of trips with no defined previous trips
T_1	Set of trips with no defined successor trips

[4] A rolling stock circulation model for combining and splitting of passenger trains (2005)

Γ_t	Set of pairs of compositions (p, p') such that $p \in P_t$, $p' \in P_{v(t)}$ and the composition change $p \to p'$ after trip t is allowed. This means that Γ_t is the set of possible changes to the composition between two consecutive trips.
$X_{t,p}$	$\begin{cases} 1, & \text{composition } p \text{ is used for trip } t \\ 0, & \text{otherwise} \end{cases}$
$Z_{t,p,p'}$	$\begin{cases} 1, & \text{composition } p \text{ is used for trip } t \text{ and } p' \text{ is used for } v(t) \\ 0, & \text{otherwise} \end{cases}$
$N_{t,m}$	Number of train units of type m used on trip t
$C_{t,m}$	Number of train units of type m that have been coupled to the train composition right before trip t
$U_{t,m}$	Number of train units of type m that have been decoupled from the train composition right after trip t
$I_{t,m}$	Number of train units of type m stored and available at station $s_d(t)$, immediately after the departure of trip t
$I_{t,m}^0$	Number of train units of type m stored that starts at station
$\rho(s)$	Re-allocation time at station s for decoupled train units

The basic model is as follows:

$$\min \quad F(X, Z, N) \tag{1}$$

subject to

$$\sum_{p \in P_t} X_{t,p} = 1 \qquad \forall t \in T \setminus T_1; \quad p \in P_t, \tag{2}$$

$$X_{t,p} = \sum_{p' \in P_{v(t)};(p,p') \in \Gamma_t} Z_{t,p,p'} \qquad \forall t \in T \setminus T_1; \quad p \in P_t, \tag{3}$$

$$X_{v(t),p'} = \sum_{p \in P_t;(p,p') \in \Gamma_t} Z_{t,p,p'} \qquad \forall t \in T \setminus T_1; \quad p' \in P_{v(t)}, \tag{4}$$

$$N_{t,m} = \sum_{p \in P_t} n_{p,m} X_{t,p} \qquad \forall t \in T \setminus T_1; \quad m \in M, \tag{5}$$

$$C_{v(t),m} = \sum_{(p,p') \in \Gamma_t : n_{p',m} > n_{p,m}} \left(n_{p',m} - n_{p,m}\right) \cdot Z_{t,p,p'} \qquad \forall t \in T \setminus T_1; \quad m \in M, \tag{6}$$

$$U_{t,m} = \sum_{(p,p') \in \Gamma_t : n_{p,m} > n_{p',m}} \left(n_{p,m} - n_{p',m}\right) \cdot Z_{t,p,p'} \qquad \forall t \in T \setminus T_1; \quad m \in M, \tag{7}$$

$$C_{t,m} = N_{t,m} \text{ and } U_{t,m} = 0 \qquad \forall t \in T_0; \quad m \in M, \tag{8}$$

$$U_{t,m} = N_{t,m} \text{ and } \mathrm{C}_{t,m} = 0 \qquad \forall t \in T_1; \quad m \in M, \tag{9}$$

$$\sum_{s \in S} I^0_{s,m} = n_m \qquad m \in M, \tag{10}$$

$$I_{t,m} = I^0_{s(t),m} - \sum_{\substack{t' \in T : s_d(t') = s_d(t), \\ \tau_d(t') \le \tau_d(t)}} C_{t',m} + \sum_{\substack{t' \in T : s_a(t') = s_a(t), \\ \tau_a(t') \le \tau_a(t)}} U_{t',m} \qquad \forall t \in T; \quad m \in M, \tag{11}$$

$$X_{t,p} \in \{0,1\} \qquad \forall t \in T; \quad p \in P_t, \tag{12}$$

$$N_{t,m}, C_{t,m}, U_{t,m}, I_{t,m} \in \mathbb{R}_+ \qquad \forall t \in T; \quad m \in M, \tag{13}$$

$$I^0_{s,m} \in \mathbb{Z}_+ \qquad \forall s \in S; \quad m \in M, \tag{14}$$

$$Z_{t,p,p'} \in \mathbb{R}_+ \qquad \forall t \in T; \quad (p,p') \in \Gamma_t. \tag{15}$$

The following sections describe each equation in the model.

4.2.1. Objective function

Equation (1) is the objective function that must be minimized and as described above, this function represents the following performance indicators:

- Carriage kilometers (CKM)
- Seat-shortage kilometers (SKM)
- Number of shunting movements (SHM)

Regard the following equation:

$$\mathrm{CKM} = \sum_{t \in T} \sum_{m \in M} \ell_t \cdot c_m \cdot N_{t,m}, \tag{16}$$

where ℓ_t is equal to the length of trip t in kilometers and c_m is the number of carriages in a train unit of type m. Within the scope of this particular problem $c_m \in \{3,4\}$. Recall that $N_{t,m}$ represents the number of train units of type m used on trip t. Therefore, equation (16) does indeed represent the total number of kilometers operated by all the carriages.

The seat shortage kilometers is denoted by the following equation:

$$\text{SKM} = \sum_{t \in T} \sum_{p \in P_t} \ell_t \cdot s_{t,p} \cdot X_{t,p}, \tag{17}$$

where $s_{t,p}$ represents the expected number of seat shortages when composition p is used for trip t. Take for example a trip with the following forecast and chosen composition:

Trip t_e		
	Forecast	**Composition p_e**
First class	30	28
Second class	100	120

Table 2: Example trip with corresponding forecast and composition choice

In this example s_{t_e,p_e} is equal to 2, which means that it is expected that trip t_e will have a shortage of two seats due to the fact that composition p_e does not feature enough first class seats with respect to the forecast.

SKM regards the chosen compositions ($X_{t,p}$) and multiplies the corresponding seat shortages with the length of the trip.

$$\text{SHM} = \sum_{t \in T} \sum_{\substack{(p,p') \in \Gamma_t: \\ \beta(p) \neq \beta(p')}} Z_{t,p,p'} \cdot \tag{18}$$

$Z_{t,p,p'}$ denotes the transition of compositions between trip t and $v(t)$. By only regarding transitions where $\beta(p) \neq \beta(p')$, the number of changes in the compositions of the trains is recorded.

The function F is a non-negative linear combination of these criteria. By adjusting the weight factors the relative importance of the criteria can be expressed.

4.2.2. Constraints

Constraint (2)

- Note that $X_{t,p} \in \{0,1\}$. If the sum over all the compositions is equal to 1 for all trips t, this means that for each trip only one composition is used. This is needed because during a trip it is not possible to change the composition.

Constraints (3) and (4)

- These constraints make sure that the transitions between consecutive trips are correct by only regarding sets of (p, p') that are within the set of allowed composition changes Γ_t.

Constraint (5)

- This constraint defines the link between the composition used in a trip and the number of different train unit types used on that trip.

Constraint (6)

- The difference in the number of train units is regarded after an allowed composition change, where the number of train units of type m after the change is greater than before the change. (Coupling)

Constraint (7)

- The difference in the number of train units is regarded after an allowed composition change, where the number of train units of type m before the change is greater than after the change. (Decoupling)

Constraint (8)

- The number of train units coupled before the first trip is equal to the number of train units used in the first trip and the number of uncoupling should be 0, regardless of the train unit type m.

Constraint (9)

- The number of train units decoupled after the last trip is equal to the number of train units used in the last trip and the number of coupling should be 0, regardless of the train unit type m.

Constraint (10)

- This constraint ensures that the allocations of train units to the initial station inventories are correct.

Constraint (11)

- During the process many changes are made in the inventories of the different stations. This constraint uses the initial inventory as a base and uses the coupling and decoupling movements to determine the current inventory.

Constraint (12), (13), (14) and (15)

- These constraints define the range of the different variables.

It is also possible to add extra "preferences" to the objective function. An example of such a condition is that the largest occurrence of seat shortage cannot be larger than some arbitrary value (e.g., 20% during rush hours). The same can be done for the other criteria. Models with different preferences are treated later on.

4.2.3. Adding new integer variables

The most interesting part of the model is the addition of an extra variable. The variable $X_{t,p}$ is used to describe the precise composition used in trip t. It appears that by introducing a new variable, which describes the length of a train, it is allowed to drop the integrality constraint on the $X_{t,p}$ variable in the model.

By using this reformulation it is possible to reduce the number of $X_{t,p}$ (approximately 9900) to about 5700 of these new variables. This new variable determines the number of the train units per type used in a trip, disregarding the exact order in the train.

Regard the following:

$$B_t = \{\beta(p) \mid p \in P_t\} \tag{19}$$

Recall that $\beta(p)$ is used to denote the number of each type of train unit used in composition p. This implies that $\beta(p)$ also denotes the length of composition p. Consequentially B_t defines the set of possible train lengths for trip t.

Now let $Y_{t,b} \in \{0,1\}$. This variable is equal to 1 if and only if combination b from the set B_t is used for trip t. The definition of $Y_{t,b}$ is:

$$Y_{t,b} = \sum_{p \in P_t : b = \beta(p)} X_{t,p} \qquad \forall t \in T; \qquad b \in B_t \tag{20}$$

The basic model is extended by the binary variable $Y_{t,b}$ and the corresponding constraint (20). Because the integrality constraint of $X_{t,p}$ is dropped, constraint (12) needs to be rewritten to:

$$X_{t,p} = [0,1] \qquad \forall t \in T; \qquad p \in P_t \tag{21}$$

Now Fioole et al [1] propose that this relaxed mixed integer model has an integral optimal solution whenever it has a feasible solution.

4.2.4. Combining and splitting trains

At this point the model only features coupling and decoupling of train units. The model can easily be extended so that combining and splitting of trains is possible as well. Fioole et al [1] only describe the splitting of trains.

Let T^s be the set of trips that occur before trips where trains are split. The split trips are denoted by $v^1(t)$ and $v^2(t)$. Now let Γ_t^s be the allowed set of split operations at station s for trip t, which is denoted by using the following notation: (p, p_1, p_2), $p_1 \in P_{v^1(t)}, p_2 \in P_{v^2(t)}, p \in P_t$ and p is the concatenation of p_1 and p_2.

$$X_{t,p} = \sum Z^s_{t,p,p_1,p_2} \qquad \forall p \in P_t \tag{22}$$

$$X_{v^1(t),p_1} = \sum_{p,p_2:(p,p_1,p_2)\in\Gamma_t^s} Z^s_{t,p,p_1,p_2} \qquad \forall p_1 \in P_t \tag{23}$$

$$X_{v^2(t),p_2} = \sum_{p,p_1:(p,p_1,p_2)\in\Gamma_t^s} Z^s_{t,p,p_1,p_2} \qquad \forall p_2 \in P_t \tag{24}$$

Although constraints of combining the trains are not described by the article its formulation is very similar to the splitting. One of the minor changes would be another set of trips denoting which two trips are to be combined in one successor trip.

The possibilities in composition changes can be extended even more by allowing coupling and decoupling after combining or splitting of trains.

4.2.5. "Practical" constraints

To ensure that the model is suitable to implement additional conditions are needed. One of these conditions is the cyclic nature of the solution. As mentioned before, the initial inventory at each station should be the same as the inventory at the end of the day. If this is satisfied, the model can be (re)used every day.

The inventory at the end of the day is defined as follows:

$$I^{\infty}_{s,m} = I^0_{s,m} - \sum_{t\in T:s_d(t)=s_t^s} C_{t,m} + \sum_{t\in T:s_d(t)=s_t^s} U_{t,m} \tag{25}$$

The corresponding cyclic constraint is as follows:

$$I_{s,m}^{\infty} = I_{s,m}^{0} \qquad \forall s \in S \qquad m \in M \tag{26}$$

Another practical constraint is to include the storage capacity of a station into the model. This is easily done by introducing an upper limit to the $I_{t,m}$ variable.

$$\sum_{m \in M} c_m \cdot I_{t,m} \leq B\,, \tag{27}$$

Recall that c_m is the number of carriages in a train unit of type m. This restriction regards the number of carriages after the departure of trip t. In this notation, all the stations have the same storage capacity B, by specifying this variable for each trip (a trip automatically determines the begin and end station) it is possible to define boundaries for each station individually.

Due to the combining, splitting, coupling and uncoupling of trains, it might occur that a passenger is traveling in a carriage which will be split before its destination is reached. This means this passenger needs to move to another train unit to reach its destination. Of course this is not favorable and that is why there should be at least one train unit that follows the complete route. This is called the continuity requirement.

Take line 1600 as an example. The trains from Schiphol and Amsterdam are combined in Amersfoort and travel to its final destination in Enschede. In this situation there are two routes: Asd – Es and Shl – Es. Changes on the composition of the train can only be done at Deventer. At this station it is possible to couple a train unit to the front of the train or decouple a train unit from the rear. Now due to the fact that the train units originating from Shl (front part of the combined train) cannot be uncoupled, the constraint will only force the continuity requirement upon the route Asd – Es.

$$\sum_{m \in M} U_{Amf-Dv,m} \leq \sum_{m \in M} N_{Asd-Amf,m} - 1 \tag{28}$$

This constraint ensures that the number of train units uncoupled is less than the number of train units used on the trip Asd – Amf.

4.2.6. Aggregation

To further improve the model in terms of computational feasibility without sacrificing any features, special cases are tackled separately like the situation at Utrecht station. In this situation trip t_1 is split into two trips at Ut. t_2 continues its course to Gvc, while t_3 is heading to Rtd. After arriving at its destinations, both trains will return to Ut (t_4 and t_5) and there they will be combined once again to operate on trip t_6.

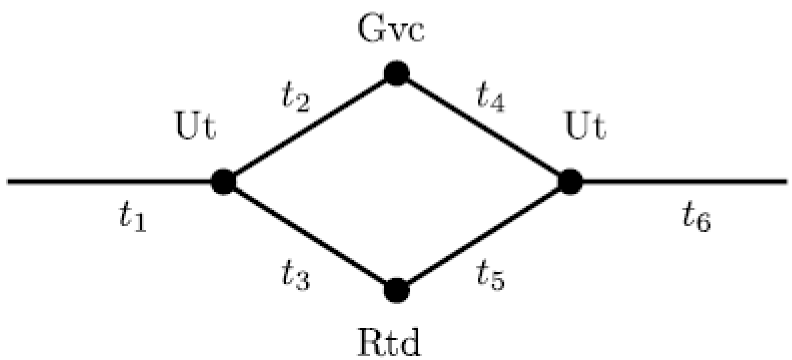

Figure 6: A special case; Utrecht station (Picture taken from Fioole et al.[5])

Note that additional coupling and uncoupling, next to the combining and splitting, is not possible in Ut, but these operations are allowed at both Gvc and Rtd.

This special case requires special variables, that is why some of the conventional variables ($X_{t,p}, Y_{t,p}$, and $Z_{t,p,p'}$) are not used for these specific trips (t_2, t_3, t_4, and t_5) and are replaced by $Z^{\text{Ut}}_{t_1,p_1,p_2,p_3,p_4,p_5,p_6}$. This variable is equal to 1 if for each trip t_i its composition is equal to p_i [$i = 1,\dots,6$]. This reformulation also replaces the variables Z^s and Z^c, which describe the splitting and combining of trains at t_1 and t_6, respectively:

$$X_{t_1,p} = \sum_{p_2,p_3,p_4,p_5,p_6} Z^{\text{Ut}}_{t_1,p,p_2,p_3,p_4,p_5,p_6} \qquad \forall p \in P_{t_1} \tag{29}$$

$$X_{t_6,p} = \sum_{p_1,p_2,p_3,p_4,p_5} Z^{\text{Ut}}_{t_1,p_1,p_2,p_3,p_4,p_5,p} \qquad \forall p \in P_{t_{16}} \tag{30}$$

Because some of the $Z_{t,p,p'}$ variables are removed from the model, it is not possible to track the number of coupling and decoupling occurrences anymore as described in constraints (6) and (7). It is however easy to use the $Z^{\text{Ut}}_{t_1,p_1,p_2,p_3,p_4,p_5,p_6}$ variable instead. These variables are also added to the objective function in such a way to represent the three criteria for the trips $t_2,\dots,t_5$.

This aggregation adds more variables to the model because there are many combinations possible in the set $(p_1,\dots,p_6)$; in this case over 700 combinations. But at the same time this reformulation eliminates a number of dummy variables $Z_{t,p,p'}$, Z^s, and Z^c corresponding to the trips $t_2,\dots,t_5$, and also $X_{t,p}$, which describes the composition on these trips.

[5] A rolling stock circulation model for combining and splitting of passenger trains (2005)

4.2.7. Summary

The Noord-Oost case consists of 167 routes, divided into 665 trips. The routes are operated by 50 train units of 3 carriages and 35 train units of 4 carriages long. Fioole et al used different models (replacing variables, cyclic constraints, continuation requirement, and aggregation) in an attempt to make the model both more practical and better solvable in a reasonable amount of time.

The next section shows the results of the implementation.

4.3. Computational results

The model formulation is at this point complete for the Noord-Oost case and is ready for the implementation. Fioole en al [1] implemented the model using the modeling software ILOG OPL Studio 3.7. The model is solved by using the Mixed Integer Programming (MIP) solver ILOG CPLEX 9.0 on a Pentuim IV 3.0 GHz processor in combination with 512 MB of internal memory.

The model is run by using different objective functions which are described in Table 2:

	CKM	**SKM**	**SHM**
Value in practice	**317,853**	**555,215**	**135**
Weight in Obj1	1	1	5000
Weight in Obj2	1	1	1000
Weight in Obj3	1	1	100
Weight in Obj4	1	10	1000
Weight in Obj5	1	10	100

Table 3: Objective criteria in the practical solution and the chosen weight factors

It is impossible to minimize all three criteria to their minimum values due to their contradicting nature. If the model wants to minimize the CKM, then the number of train units used to operate the routes will most likely decrease. At the same time, the SKM will most likely to rise, because less carriages also means fewer seats available and thus increasing the expected seat shortage. The SHM will most likely increase as well due to the fact that the model will decrease the number of train units whenever possible which will lead to many extra shunting movements and even more extra shunting movements are needed when the next station has more passengers.

The question is how the user values each criterion, which criterion is more important than the other two? The NS has no strict preferences with regard to the CKM, which is why this criterion has a low weight factor. There is however a constraint to this value: $CKM \leq 318{,}000$. This is the value found from practice). Different weights are used on the SKM and SHM to investigate how these influence the solution.

	Without aggregation		With aggregation	
	MIP	Reduced MIP	MIP	Reduced MIP
Variables	51,677	34,228	56,757	39,251
Constraints	32,308	17,896	28,149	15,348
Non-zeros	195,184	136,105	287,286	177,622

Table 4: Characteristics of the MIPs created by CPLEX

Table 3 shows the dimension of the different models, not counting the different weight factors in the objective function. First, the models are divided in a group with and without aggregation as described in 4.2.6. Then, within each group there are again two models; with and without variable reduction as described in 4.2.3.

An important property of the solution method is the required time to find the solution which is described in Table 5.

		Without aggregation		With aggregation	
Obj	**SHM**	**Optimality gap after 7200s**	**Best solution found in**	**Optimality gap after 7200s**	**Best solution found in**
	-	2.56%	5400 s	1.69%	3000 s
1	120	2.06%	4800 s	2.30%	3500 s
	90	4.90%	5700 s	4.80%	6000 s
	-	0.76%	5000 s	0.22%	3200 s
2	120	0.67%	6400 s	0.50%	7100 s
	90	16.11%	5300 s	2.67%	7100 s
	-	0%	5000 s	0%	2000 s
3	120	-	-	1.24%	3400 s
	90	-	-	-	-
	-	0.12%	4300 s	0%	1900 s
4	120	11.20%	6300 s	2.22%	6500 s
	90	-	-	11.43%	6400 s
	-	0%	4400 s	0%	2700 s
5	120	13.16%	5000 s	2.32%	5100 s
	90	-	-	-	-

Table 5: Overview of objective functions and the corresponding optimality gap after 7200 seconds and the amount of time needed to achieve the best solution in seconds

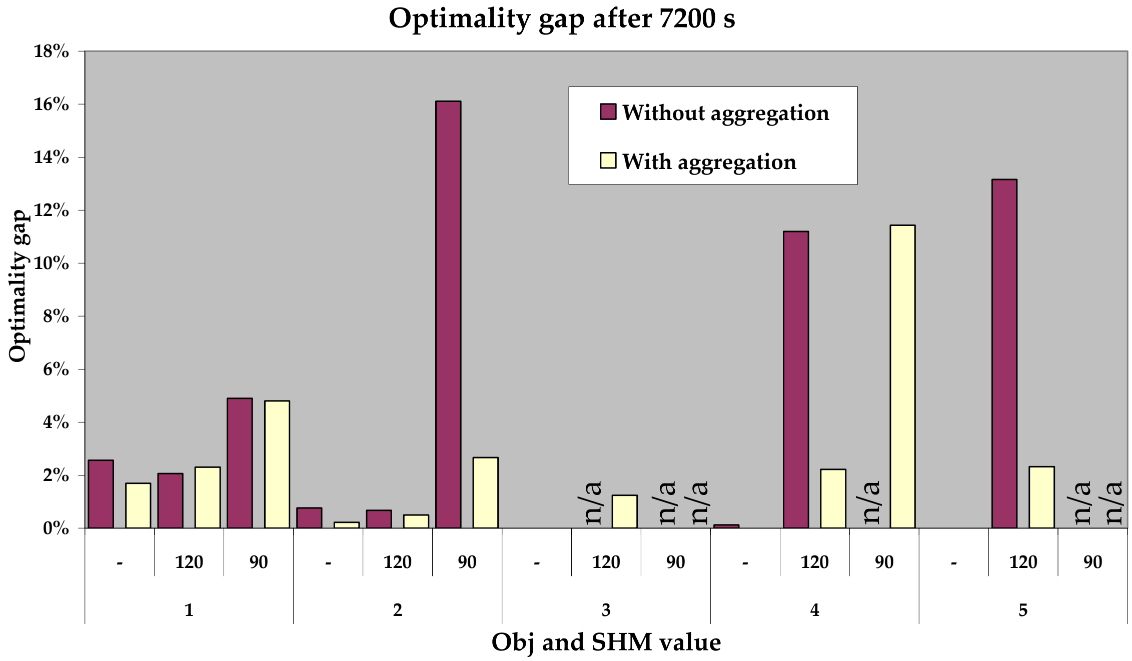

Figure 7: The gap between the optimal solution and the solution obtained after running the model for 7200 seconds

Figure 7 shows the optimality gap (gap in percentage between the optimal solution and the obtained solution) after 7200 seconds. "n/a" in the figure denotes that no feasible solution was found within 7200 seconds of running the model. The first observation is the fact that including the "aggregation" step, the model finds in almost all cases a feasible solution after 7200 seconds. On top of that, in most cases the corresponding optimality gap is lower than its counterpart "without aggregation".

Figure 8 shows how long it takes to obtain the solution described above. Note that this solution is not necessarily equal to the optimal solution. Due to the relaxation and rewriting of constraints, it might be possible that the resulting model will not be able to reach the optimal solution at all.

If no additional constraints are placed on the SHM, the models with aggregation are considerably faster in reaching their best solution. When the restriction on SHM increases, this difference becomes less and, in some cases, is even beaten by the model without aggregation.

The most important observation is that the model with aggregation gives better results and this is achieved in a respectable amount of time.

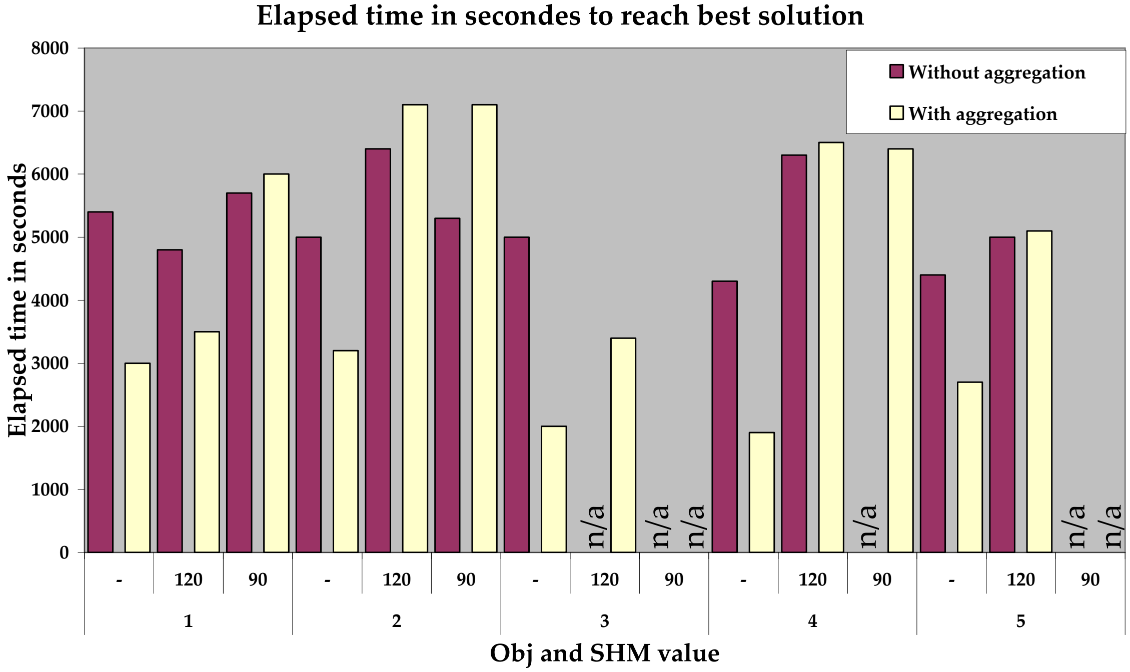

Figure 8: The amount of seconds needed to obtain the best solution, per objective function and SHM value combination

5. Crew planning

Given the timetable and the rolling stock circulation, the only ingredient left are the people who will carry out the service. This also requires a great amount of planning due to the different restrictions and requirements on this problem. This planning problem consists of two sub processes:

- Crew scheduling
- Crew rostering

The first process describes the length of each duty and the necessary breaks in each duty. By using these duties to create a roster is called crew rostering, the second process. Crew rostering can be done in various ways:

- A roster for individuals where their specific wishes are taken into account
- A bid line, personnel can bid on their favored roster
- A cyclic roster

Since 2000 the NS is using TURNI, a program based on mathematical programming techniques originating from Operations Research, to make the crew schedules. The paper *Reinventing Crew Scheduling at Netherlands Railways (2004)* is a great source of information on how the NS deals with the crew scheduling problem.

For more information regarding crew rostering, this document refers to the paper *Decision Support for Crew Rostering at NS (2006)*. This paper describes a method to solve the cyclic crew rostering problem (CCRP) and implements this solution for a real life case. The results are very satisfying; not only the speed of generating such a roster was vastly superior to the manual process, the personnel were also more content with the new roster as opposed to the manually made roster.

6. Closing

The Netherlands has one of the world's busiest train networks of the world. The number of passengers is increasing each year, while the capacity of the infrastructure is not growing at the same rate. More improved and/or advanced models are needed which will utilize the available capacity more efficiently in order to ensure that the system will not be overloaded.

The problem that the NS has to face can be described in two steps. The first step is to construct a feasible time table. This time table consists of all the departure and arrival times of the trains at all defined stations. The first part of this document briefly discusses this process. One of the most important constraints in this problem is to avoid conflicts; an example of such a conflict is a train overtaking another train on the same track.

Considering the situation where the NS has successfully constructed a feasible time table, it is possible to regard the next challenge: Rolling Stock Circulation. Just like the former problem, this problem is very difficult to solve due to the large size of the problem. The professionals in this field of study are constantly improving the analyses on these problems to pursue one goal: a practical solution. Such a solution is presented by Alexander Schrijver and this is also the focus of this document. The second part of this document describes how this difficult problem is tackled by applying standard modeling and optimizing steps and a not so standard optimization step; by adding an additional integer variable to represent the composition of the trains it is possible to reduce the total number of variables in the model. On top of that, the original variable to represent the composition can be relaxed to a fractional number without having to give in on the integral solution of the model. Unfortunately, this extraordinary procedure is very specific to this problem and therefore unlikely to be directly applicable to other problems that are difficult to solve.

A problem which was next to unsolvable due to its enormous size turns into a solvable problem that is solvable within a respectable amount of time by using clever modeling and optimization techniques. Although the solution might not be optimal, the difference between the two is not very large. Like said before, the goal is to obtain a practical solution and this is exactly what it is.

7. Appendix

The article about Alexander Schrijver from de Volkskrant:

De ideeën komen uit het niets (de Volkskrant, Kennis, 16 december 2006 (pagina K07))
Michael Persson

Als jongen maakte Lex Schrijver al schema's voor fietsritten. Nu deed hij wiskunde voor de NS. 'Die dienstregeling vergde grensverleggend onderzoek .'

Lex Schrijver: 'Het moet maar eens gezegd, wiskunde is moeilijk.'

Lex Schrijver
1948 Geboren in Amsterdam
1977 Promotie wiskunde aan de Vrije Universiteit Amsterdam
1983 Hoogleraar Universiteit van Tilburg
1989-nu Hoogleraar Centrum voor Wiskunde en Informatica (CWI)
1990-nu Hoogleraar Discrete Wiskunde en Optimalisering aan de Universiteit van Amsterdam.
2003 Driedelig standaardwerk: Combinatorial Optimization – Polyhedra and Efficiency
2005 Spinozapremie

In een kast in zijn grachtenpand staat een halve eeuw aan spoorboekjes, van de jaren vijftig tot 2006. Netjes in het gelid. Even verderop drie planken volgestouwd met kaarten, op land geordend. 'Ik gooi niet makkelijk iets weg', zegt prof. dr. Lex Schrijver. 'Zeker dat soort dingen niet.'

Het is allemaal studiemateriaal voor de Amsterdamse wiskundige, hoogleraar aan de Universiteit van Amsterdam en het Centrum voor Wiskunde en Informatica. De problemen die hij bestudeert, komen vroeg of laat steeds neer op het maken van dienstregelingen, het uitstippelen van routes en het verdelen van gebieden op plattegronden. Het zoeken naar de kortste, de snelste, de makkelijkste weg. Tussen drie, vier of *n* punten. Iedereen die ooit een krantenwijk heeft gelopen, kent het probleem.

Deze week had Schrijver (1948) weer een van zijn finest hours. De nieuwe dienstregeling van de NS is gemaakt met methodes en computerprogramma's die door hem en zijn collega Adri Steenbeek zijn bedacht en doorgerekend. Een typisch staaltje van combinatorische optimalisering, zoals Schrijvers vakgebied in wiskundige wandelgangen wordt genoemd.

Hij ziet er tevreden uit, in zijn opgeruimde werkkamer aan de Keizersgracht in Amsterdam. Houten vloer, sterke koffie. Slapeloze nachten heeft hij niet gehad, in de aanloop naar de invoering van het nieuwe spoorboekje. 'Ik dacht wel dat het zou werken. En je moet mijn verantwoordelijkheid niet overdrijven. De wiskunde is waardenvrij. De keuzes worden gemaakt door de NS.'

Als hij een presentatie geeft, zet hij op de laatste dia steevast het klachtennummer van de spoorwegen. Aan de andere kant, zegt hij grijnzend, zal hij als de nieuwe dienstregeling de komende maanden zijn waarde bewijst, natuurlijk wel wijzen op de oorsprong van dat succes. De wiskunde.

Als jongetje in Amsterdam-West maakte hij al dienstregelingen voor fietsritten van hem en zijn vier broers, met afspraken waar ze elkaar zouden treffen en hoe laat. Inmiddels denkt hij al vijftien jaar over treinen na. In 1991 kwam de NS in de problemen, omdat er na invoering van de ov-jaarkaart veel studenten in de trein stapten. De dienstregeling, waarvan het basispatroon in 1970 met de hand was uitgetekend, raakte zo vol met nieuwe ad hoc-treinen, dat het steeds moeilijker werd om er nog iets aan toe te voegen.

Tamelijk eindeloos: De NS stapte eerst naar een softwarebedrijf met de vraag de boel aan te passen. Maar zo makkelijk ging dat niet. Het aantal mogelijkheden om de bijna vierhonderd stations in Nederland met elkaar te verbinden is tamelijk eindeloos. Computers kunnen dat met domme rekenkracht niet uitrekenen, voor het einde van het universum.

Dus moest de NS op zoek naar slimme algoritmen. 'Zo kwamen ze bij ons', zegt Schrijver. 'Het was een praktisch probleem, maar het vergde grensverleggend onderzoek.'

De Amsterdamse wiskundigen vertaalden alle vereiste reistijden, aansluitingen en overstaptijden in een verzameling vergelijkingen, Cadans genaamd (Combinatorisch Algebraïsch Dienstregeling Algoritme voor de Nederlandse Spoorwegen). Vervolgens moest de computer een dienstregeling bedenken die aan al die eisen voldeed. 'Dat lukte natuurlijk niet', zegt Schrijver.

Dan moet iemand ingrijpen. Aansluitingen worden geschrapt, treinen verlengd of opgeknipt. Waarna het geheel opnieuw wordt doorgerekend.

In eerste instantie keek Schrijver vooral naar de capaciteit van het spoornet. De problemen in de Randstad bleken vooral te worden veroorzaakt door de oude ijzeren enkelsporige brug over de IJssel bij Zutphen. 'Dat was een grote verrassing', zegt Schrijver. 'Als je het met de hand zou doen, dan zou je toch eerst op zoek gaan naar knelpunten in de Randstad.' De brug over de IJssel is nu dubbelspoor.

Ander overduidelijk knelpunt, bleek jaren geleden al: het spoor tussen Amsterdam en Utrecht. 'Daar hadden al veel eerder vier sporen moeten komen. Dat dat niet gebeurde, was natuurlijk ook een gevolg van politieke keuzes.'

Daar houdt hij zich verre van. Schrijver houdt van de dienstregeling als wiskundig probleem, als ideale prooi voor combinatorische optimalisering. 'Je ziet de mensen, de treinen, de zitplaatstekorten. En denkt: dat moet beter kunnen.'

Het leuke aan elk wiskundig vraagstuk, zegt hij, is dat je precies weet wat het probleem is. Hij houdt van de helderheid – er is geen mist van waaruit ineens nieuwe ijsbergen kunnen opdoemen. 'Je hebt het helemaal in de hand. Hoeft niet te gaan improviseren. Je kunt je in volkomen concentratie wijden aan die ene vraag. Die concentratie maakt een mens gelukkig, volgens de flow-theorie van die man met die moeilijke naam.' De volgende dag, per mail: 'Csikszentmihalyi.'

De beste ideeën komen uit het niets. Op de fiets, van zijn instituut in de Amsterdamse Watergraafsmeer naar zijn huis in het centrum – hij neemt daarom nooit de kortste route.Of tussen Kerst en Oud en Nieuw, als hij even helemaal niets te doen heeft, zoals in 2002. 'Ik weet niet waar de ingeving vandaan kwam. Ik had er al jaren over nagedacht. Ik zat in de kerstboom te staren, en toen kwam ineens de oplossing voor het kilometerzitplaatstekort bij de koplopers.'

Het ging om het aan- en afkoppelen van treinstellen van het koploper-type, die verschillende lengtes kunnen hebben en daardoor een complex logistiek vraagstuk vormen. Schrijver deed iets raars: hij maakte de vergelijkingen ingewikkelder, waarna het hele stelsel ineens oplosbaar bleek. 'Dat is het grensverleggende. Soms moet je het eerst moeilijker maken voor je eruit kunt komen.'

Dat moet ook maar eens gezegd: wiskunde is moeilijk. Hij heeft het niet zo begrepen op de beleidsmakers die, in het landsbelang, het aantal bèta-studenten willen vergroten door te doen alsof wiskunde en andere exacte disciplines voor iedereen zijn weggelegd. Dat werkt averechts, denkt hij. Want daarmee jaag je de slimme jongens en meisjes weg die het juist leuk vinden omdat het moeilijk is. 'Dat was voor mij een reden om wiskunde te gaan studeren', zegt Schrijver. 'Hoe lastiger, hoe groter de voldoening van het oplossen.'

Kleuringsproblemen: Hij heeft een deel van de anderhalf miljoen euro die hij vorig jaar als Spinozaprijswinnaar kreeg gereserveerd voor een programma dat scholieren in wiskunde moet interesseren. Dat project, DisWis (Discrete Wiskunde) geheten, gewijd aan het bewijzen van stellingen en het analyseren van klassieke problemen als het kleuren van landkaarten (hoeveel kleuren heb je minimaal nodig om de verschillende landen op een kaart van elkaar te onderscheiden?), gaat hij richten op 5-vwo scholieren die wiskunde-B in hun pakket hebben. 'Daar zitten mensen die het leuk vinden als het moeilijk wordt.'

Daarbij is Schrijver wel iemand die oog heeft voor meer dan de elegantie van een formule. Hij heeft het over mensen, ontstaansgeschiedenissen, toepassingen. Neem het kortste-pad-algoritme van de Nederlandse informaticus Edsger Dijkstra uit 1957, dat nu in elke TomTom zit. 'Als ik daar in Paradiso een verhaal over vertel, is iedereen mateloos geïnteresseerd.'

Soms is hij zelfs een halve historicus. In zeventien jaar schreef hij het standaardwerk in de combinatorische optimalisering, een knalgeel driedelig werk dat in een stevige kartonnen houder bescheiden tussen de andere boeken bij hem in de kast staat. Het is harde wiskunde, natuurlijk. Maar voor Schrijver staan theorieën nooit op zichzelf. Ze komen ergens vandaan, er zitten verhalen achter.

Neem de vraag hoe je langs zoveel mogelijk routes in een netwerk zoveel mogelijk goederen kan vervoeren, het Maximum-Stroom Probleem. Anders geformuleerd: waar zit de bottleneck?

Het probleem was uit de spoorwegpraktijk afkomstig, stond in een voetnoot bij een wetenschappelijk artikel uit 1955. De noot verwees naar een 'Harris-Ross rapport' van de RAND Corporation, een destijds aan de Amerikaanse luchtmacht gelieerd adviesbureau.

Luchtmacht? Spoorwegpraktijk? Schrijver vroeg RAND om het rapport, maar dat bleek nog classified. Na aandringen van Schrijver kreeg RAND toestemming om het rapport vrij te geven, en stuurde een van de twee bestaande exemplaren naar Amsterdam.

Een van de kaartjes uit het rapport zit nu in het boek van Schrijver. Daarop alle spoorlijnen tussen Moskou en de Oostbloklanden, met tonnages en knooppunten. Vanaf de Oostzee, dwars door Polen en Tsjechoslowakije loopt een stippellijn naar het zuiden met het woord bottleneck.

'Daar was het dus om te doen', zegt Schrijver. 'De Minimale Doorsnede, de keerzijde van de Maximale Stroom. Als je het treinverkeer wilde platleggen, moest je de spoorlijnen kennelijk daar bombarderen. De minimale inspanning voor het maximale resultaat. Gelukkig wordt combinatorische optimalisering ook voor vreedzame doeleinden gebruikt.'

8. Bibliography

1. Fioole, P.J., Kroon, L., Maróti, G., Schrijver, A., *A rolling stock circulation model for combining and splitting of passenger trains* (2005)

2. Vromans, M. J. C. M., *Reliability of Railway Systems* (2005), Thesis

3. Abbink, E., Fischetti, M., Kroon, L., Timmer, G., *Reinventing Crew Scheduling at Netherlands Railways* (2004)

4. Hartog, A., Huisman, D. Abbink, E. J. W., Kroon, L., *Decision Support for Crew Rostering at NS* (2006)

5. Barber F., Abril M., Salido, M. Á., Ingolotti L.P., Tormos, P., Lova, A., *Survey of automated Systems for Railway Management* (2007)

6. Alfieri, A., Groot, R., Kroon, L., Schrijver, A., *Efficient Circulation of Railway Rolling Stock* (2006)

7. Dennis, H., *Timetable 2007: Why and How?* (Powerpoint presentation, 2007)

8. http://www.ns.nl

9. http://www.wikipedia.org

10. http://www.spoorweggeschiedenis.nl

11. http://www.volkskrant.nl