Adding human aspects to Planning Assistance Tools for complex logistical processes: a case-study of railway node planning in the Netherlands

M.P. Baurichter

TIL 5060 Thesis

Bridging the gap between Planning Assistance Tools and humans

by

M.P. Baurichter

Project duration:	October 8, 2019 – June 4, 2020	
Committee:	Prof. Dr. R.M.P. Goverde	TU Delft
	Dr. W.W. Veeneman	TU Delft
	Dr. N. Bešinović	TU Delft
	Ir. W. Fleuren	NS

Delft University of Technology

Contents

Preface

This thesis is the result of my graduation research concerning the applicability of Planning Assistance Tools within complex logistical processes and concludes my studies in Transport, Infrastructure Logistics at the Delft University of Technology. From the 8th of October 2019 till the 4th of June 2020, I worked on this thesis as an intern at the department of Research & Development Node Logistics at Netherlands Railways (NS).

As is tradition, there are some people I would like to thank as they have been a great asset during my research. Lets start with all the support I have received from my supervisors of the TU Delft. I would like to thank Nikola Bešinović and Wijnand Veeneman for their intensive support and the detailed feedback on my work from an academical perspective. The meeting we had were always very constructive and meaningful. Additional gratitude goes to Rob Goverde for feedback and structuring the complete graduation process, even during the unexpected COVID-19 situation.

Another large thanks goes to Wouter Fleuren for his intensive support and trust in my research. The weekly progress meetings were extremely helpful. Next to Wouter, I would like to thank the complete R&D team for being very supportive and helpful. Especially the support from Joris Den ouden, Pim van den Bogaerdt and Demian de Ruijter was essential for this research. I would also like to thank Bob Huisman for introducing me to the subject and providing academical insights within the research scope.

But this research was not possible without the help and cooperation of the many planners that I have interviewed and that were part of my research group. I want to especially thank the V%B planning departments of Eindhoven and Rotterdam and Wietske van Heusden for giving me the freedom to perform the research with help of these two departments. Planners that I would like to thank in special are Marie-Jose Delnoye van Velthoven, Paul Notenboom, Linda Rousseva, Aboubaker Borleh and John Kleuskens.

Last but not least, I would like to thank my parents, family and girlfriend for their everlasting love and support. They have always been there when I needed them, not only during my graduation project. Thank you!

Enjoy reading!

Matthijs Baurichter
Rotterdam, June 2020

Executive summary

Introduction

Within the Netherlands, the amount of train passengers is growing every year. The NS is warning the government about the rapid growth of train passengers and the limited capacity of the rail system (Verlaan, 2019) (Treinreiziger, 2018). To respond to this problem, NS (the biggest railway undertaking of the Netherlands) is expanding its fleet (NS, 2016) in order to move more passengers. In order to maintain the fleet, daily services are executed to trains at Service Locations (SL) close to major train stations. An expansion of the fleet can be directly linked to a necessity of more parking and service area needed by the NS. Physical enlargements of SLs are planned but more efficient planning can already be used in order to increase the capacity of SL in the near future. A Planning Assistance Tool (PAT) should take over parts of the planning process that can be easily solved by a computer. However, human planners will always be in control of the complete process.

Introducing, implementing and working with a PAT comes with problems for actors in planning processes (McKay and Wiers, 2006). Whereas in theory a PAT should result in greater efficiency, lower workload and fewer human errors (Lee and Seppelt, 2009), the practical examples show something different. Experience from simulation projects has shown that plans from a PAT are difficult to interpret, difficult to trust, difficult to compare, mutually contradicting and difficult to act upon for human planners (Lorentzon and Fredlund, 2017). There is a lack of fit between the human aspects and the formal or systematized aspects found in PAT (McKay and Wiers, 2006). It is interesting to find out if a PAT will be better accepted if human aspects are known and integrated into the tool based on the qualitative differences between the human and computerized way of solving planning problems.

Goal and scope of research

This research has two objectives. The first objective is **to determine which aspects are missing in Planning Assistance Tools for complex planning processes**. The term *aspects* refers to all objective and rational reasons why a human solution of a planning assignment is different compared to a computer-based solution. The second objective is **to find out what the implementation of aspects based on human planning observations in the PAT does to the solution of the PAT and the acceptance of human planners**. Does the solution of the PAT with aspects resemble the solution of a human planner? In order to structure the research, the following research question is proposed: **what does the addition of missing aspects based on human planning observations do to the solution of a Planning Assistance Tool?**

The case study that is used for the empirical experiment and the PAT is a railway node planning process. The NS is currently working on a PAT that can be used. The focus of this research will be based on the operational phase of the planning process. This means that no long term (strategical) planning process is involved, just as real-time planning or control by dispatchers.

Research methodology

The complete reseearch methodology consist out of an empirical experiment (Figure 1), the creation, selection and implementation of extensions to cover human aspects (Figure 2) and a comparison between the adjusted Planning assistance tool and the human planning solution and original PAT (Figure 3. The goal of the empirical experiment is to gather aspects that planners think are important to include in planning options that the current algorithmic planning is not including. A real-life planning assignment is used, and for the PAT an already developed tool is used called HIP. Next to it, experienced planners of the node of Eindhoven are the members of the research group.

Based on the list of differences and aspects from the empirical experiment, it is possible to create extensions that have to be added to the PAT. A selection of the most important aspects is made based on the number of appearances during the experiment and the level of importance by human planners. Based on the level of

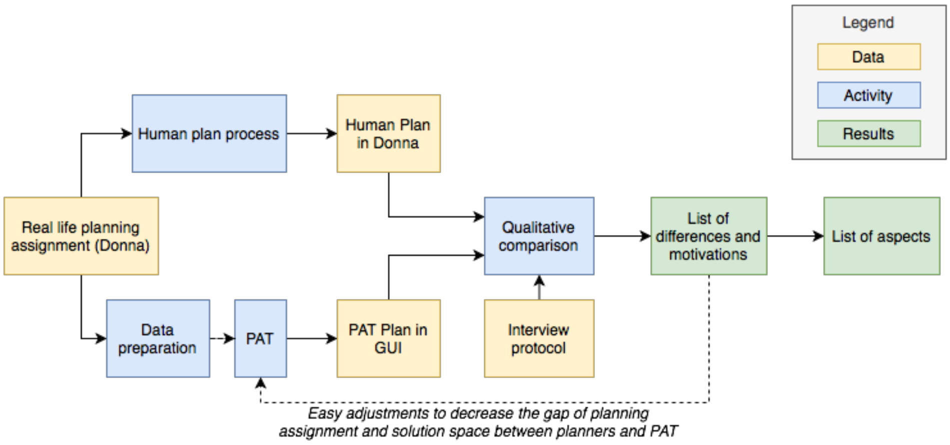

Figure 1: Conceptual model of the empirical experiment performed with human planners in order to compare human solutions of planning assignments with solutions created by the Planning Assistance Tool.

difficulty and the level of importance of the aspect, extensions are implemented to the original PAT. The PAT with the implementations is called the adjusted PAT.

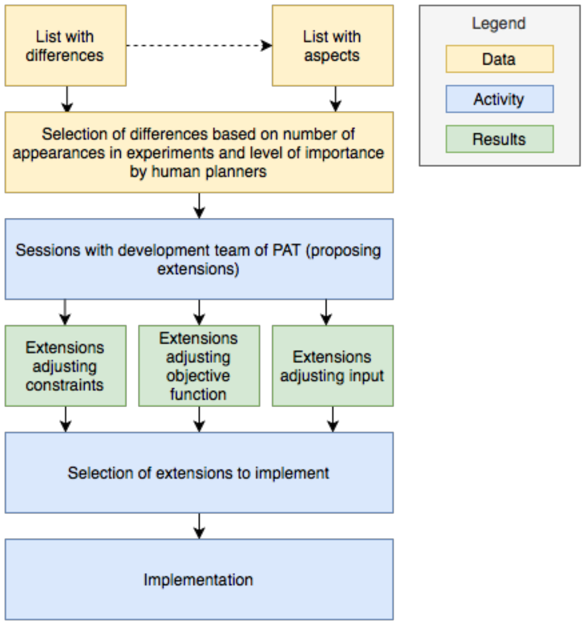

Figure 2: The process of defining, selecting and implementing extensions based on the list of differences and aspects defined by the empirical experiment

In order to find out what the impact of the extensions to the PAT is and what the reaction of the human planners is to the plan of the adjusted PAT, three comparisons are performed. A qualitative comparison is performed in order to capture the reaction of the planners of the solution of the adjusted PAT. A quantitative comparison could compare the performance of the PAT plan in relation to the original PAT for two different planning assignments: real-life and theoretical. The theoretical planning assignments are fictional planning assignments that are created based upon a real-life planning assignment. The theoretical assignments differ between themselves in a level of complexity.

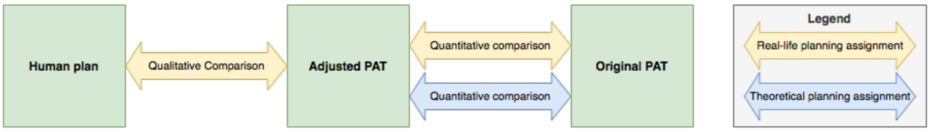

Figure 3: The comparisons made in order to test the adjusted PAT (PAT with extensions). The comparison between the human plan and the adjusted PAT is based on real-life planning assignments and has the goal to capture the qualitative differences. The two comparisons between the adjusted PAT and the original HIP have the goal to capture the quantitative differences based on real-life planning assignments and theoretical planning assignments.

Results

From the experiments 59 differences were identified between the human plan and the PAT plan. The differences could be categorized in six aspects: simplification of the planning assignment, capacity optimization, increasing flexibility, optimizing plan for KPIs, norms and agreements and safety. It is possible that certain differences relate to multiple aspects. Based on the six aspects and the list of differences, 37 extensions were proposed in cooperation with the development team of the PAT. From these 37 extensions, 20 were selected to be implemented.

The qualitative comparison proved that the adapted PAT produced a better solution compared to the original PAT in two ways. The first way was the fact that there were less irrational choices made in the solution of the adjusted PAT. The second way was a higher level of resemblance between the human solution and the solution of the PAT. The quantitative comparisons proved that the adjusted PAT performs better compared to the original PAT when looking to conflict costs.

Conclusion

First of all, the empirical experiment in a field study set-up using real-life planning assignments and experienced planners proved to be a successful way to determine which aspects are missing in a PAT. The extensions made to the PAT in order to cover the missing aspects identified by the empirical experiments do have a positive impact on the solution of the PAT, both in a qualitative and quantitative manner. From a qualitative perspective, the proposed solution of the adjusted PAT better resembled the human solution. The human planners had the feeling to understand the solution and the choices made by the adjusted PAT better compared to the original PAT solution. This indirectly provides a base of trust that is needed for planners to accept the solutions of the Planning Assistance Tool. The increase of interest to the way the PAT is solving planning assignments and the increase of resemblance between the human solution and the PAT solution can be directly related to a higher level of acceptance of the PAT by human planners. From a quantitative perspective, the differences in performance of the adjusted PAT are positive compared to the original PAT regarding real-life scenarios. In summary, it can be stated that the implementation of aspects based on human planning observations in the PAT does have a positive impact on the acceptance of the PAT by human planners and does increase the performance of the PAT.

List of abbreviations

Abbreviations

BD	Basic-Day
BDu	Basic-Day update
CG	Column Generation
CBG	Centraal Bediend Gebied (Centrally Operated Area)
CP	Constraint Programming
DCR-trains	Daily Check-up and Interior Cleaning trains
EMU	Electrical Multiple Unit
FAM	Function Allocation Model
FILO	First In Last Out method
GA	Genetic Algorithms
GUI	Graphical User Interface
HIP	Hybrid Integral Planning
HPP	Hierarchical Production Planning
IDEF0	Integrated DEFinition for Function (model)
IM	Infrastructure Manager
KPI	Key Performance Indicator
MIP	Mixed-Integer Programming
MR	Medewerker Rangeren
NCBG	Niet Centraal Bediend Gebied (Not Centrally Operated Area)
NO	Netwerk Ontwerp
NS	Nederlandse Spoorwegen (Dutch Railways)
NSR	NS Reizigers
OPG	Opstel Plan Generator
O&S	Onderhoud & Service (Maintenance & Service)
PAT	Planning Assistance Tool
RGCH	Randomized Greedy Construction Heuristics
RTC	Rail Traffic Controller
RU	Railway Undertaking
SD	Specific-Day
SL	Service Location
TAP	Thinking Aloud Protocol
TLV	Track length violation
TUSP	Train Unit Shunting Problem
TWI	Train Washing Installation
V&B	Voorbereiden & Bijstellen
VL	Verkeer Leiding (Traffic Management)
WvB	Werkvoorbereider

1

Introduction

1.1. Assignment introduction

The Dutch Railways (in Dutch: Nederlandse Spoorwegen, NS) is the biggest Railway Undertaking (RU) in the Netherlands. In order to execute services on more than 6000 km of rail, the NS uses over 600 train units. With these services, more than 1 million passengers are being transported every working day. These services are planned in a timetable. Depending on the estimated passenger demand, NS plans rolling stock in order to fulfil the services in the timetable and to optimize the available capacity. If rolling stock is not used, it needs to be parked on a shunting yard in order to keep the infrastructure available for other services. During the nights almost no passenger transportation services are executed, which results in a large amount of rolling stock on the shunting yards. In order to maintain the fleet, services are executed on train units that are parked at a shunting yard overnight. The complete shunting program is planned on beforehand since rail infrastructure is scarce and robustness of the passenger services is crucial for NS. Customers expect clean trains that drive according to the timetable. The planning process that focuses on the different shunting activities is called the node planning process. Simplistically, this planning determines what happens to a train unit when it ends a passenger moving service at a station (part of a node) till it needs to start a passenger moving service from the same node.

The CEO of the NS, Roger van Boxtel, is warning the government about the rapid growth of train passengers and the capacity of the complete rail system (Verlaan, 2019) (Treinreiziger, 2018). To respond to this problem, the NS is expanding their fleet (NS, 2016) but there are also critical capacity problems concerning the railway infrastructure. Investments in the rail infrastructure are necessary to expand the capacity of the Dutch rail system. ProRail, the Infrastructure Manager (IM) of the Netherlands, agrees with the NS and is planning to renovate and expand the critical points in the network: stations, specific connections between cities and the parking and service yards for trains (NRC, 2018) (Dagblad van het Noorden, 2019). An increase in the fleet of NS results in a necessity of more parking space and service locations. Investments and renovations of infrastructure are expensive. Therefore the availability of the infrastructure must be well utilized. A mix of new infrastructure and better use through better planning and control processes should increase the capacity of nodes.

The *Research & Development Node Logistics* department of the NS is researching to increase the capacity of nodes by improving the planning process of nodes. Part of this research is to create a Planning Assistance Tool (PAT) that can support planners to make faster and better decisions. The idea is to generate a Hybrid Integral Planning tool (HIP). The tool should combine multiple different planning processes into one overall planning process (integral) using different kinds of planning methods (hybrid) (Den Ouden, 2019). The reason why an integrated planning tool is necessary is that at this moment, the planning of a node contains several different planning processes that take place simultaneously and successively (de Ruijter et al., 2019). The current plan process does not function as expected on specific nodes since not all the produced plans are always feasible (van Uden et al., 2017). By integrating the planning process, NS expects an increase of feasible plans. Next to that, NS expects that the plans are better measurable (within and between nodes), provide more control, generate more understanding of the complete planning process, increase the transparency and improve solutions regarding parking and service capacity (van Uden et al., 2017).

Introducing, implementing and working with a PAT comes with problems for actors in planning and production processes in general (McKay and Wiers, 2006). Whereas in theory PAT should result in greater efficiency, lower workload and fewer human errors (Lee and Seppelt, 2009), the practical examples show something different. Experience from simulation projects has shown that plans from a PAT are difficult to interpret, difficult to trust, difficult to compare, mutually contradicting and difficult to act upon for human planners (Lorentzon and Fredlund, 2017). Computers have strict rules and a large computing power to generate new plans and restructure existing ones fast while humans have the advantage of using soft skills for planning processes (Sanderson, 1989). It is because of these characteristics that terms like computer-based decision making and human-centred design are solutions that are based on the advantageous characteristics of computers and humans. However, there is a lack of fit between the human element and the formal or systematized element found in PAT (McKay and Wiers, 2006). McKay and Wiers (2006) call this lack of fit in computer-generated plannings the *human factor*. However, the term *human factor* is comprehensive and does not describe clear aspects. Besides, these human factors differentiate per planning process and planning structure. Also a lot of research is performed to the quantitative comparisons between humans and computer-generated planning processes, but almost no research is performed to the qualitative comparisons between these two different planning techniques. These qualitative differences can also influence the acceptance of Planning Assistance Tools by humans.

Summarizing this content to one question, it is interesting to find out if a PAT will be better accepted if human factors are known and integrated into the tool and if the qualitative differences between the human and computerized way of solving planning problems are known? This research describes how this question will be answered using a clear methodology. First, extended problem identification is necessary to explain the details of the problem.

1.2. Node planning and its definitions

Before any objectives or research questions are proposed, it is necessary to understand what the node planning process all includes. This information will be elaborated on in Chapter 3, where a complete analysis of the node planning process of NS will be executed. This section is an introduction to the process and an explanation employing a few essential definitions.

Train

The word 'train' is a collective name for a composition of one or more rail carriages. In this research, the following terms will be used, which are made visible in Figure 1.1:

- Carriage: one individual wagon that is part of a (train) unit;
- Electric Multiple Unit (EMU): a composition of one or multiple carriages in a fixed order that are self-propelled, from now on called a (train) unit;
- Composition: a composition of one or multiple (train) units for executing a passenger service.

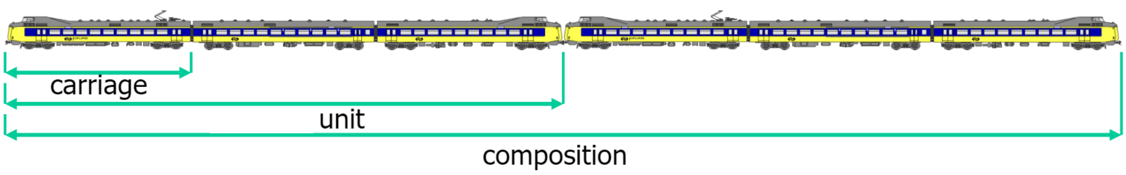

Figure 1.1: Definitions of different parts of a train (Goverde, 2018)

Node

What is node planning exactly, and what activities does it all contain? It is vital to have a clear definition of a railway node. According to the network declaration of ProRail, a node is a network connection point or a point where the network can be joined where some logistical- and planning processes of train services are concentrated and dealt with as a whole (ProRail, 2019). The NS adds to this description the functional area of a node. The area of a node can consist of interconnected yards and station(s), whether or not separated by line sections (in Dutch: vrije baan, the arcs connecting all the nodes) (NS, 2018b) (Rail Net Europe, 2019).

For this research, a more straightforward definition for the nodes of NS is created: A node consists of a station and the surrounding shunting yards. These shunting yards are called Service Locations (SLs) because services (including parking) can be applied to train units on these shunting yards.

The node planning process can be described by three types of resource planning processes (ProRail, 2019):

- Infrastructure process: planning, distribution and release of available infrastructure;
- Equipment process: planning and execution of equipment treatments and shunting;
- Personnel process: planning and control of the personnel services.

Services

If trains are not moving passengers, the NS tries to allocate cleaning and maintenance services to the trains. These services are depending on the Service Locations, as some services can only be performed with special equipment or at a special track. The list of services and required service equipment that are included in the node planning are:

- **Parking of trains:** This can be done on allocated tracks. These locations need to have at least room for the driver to enter or exit the train and needs to have the possibility to exit the yard safely.
- **Mechanical checks on trains:** In this case, it is necessary to walk 360 degrees around the train. On both sides of the train, there should be a designated walking path for the mechanic to check the train externally. During this process, the train must be parked.
- **Interior cleaning of trains:** For this activity it is necessary for the cleaning crew to enter and exit the train and to have access to water and electricity. Therefore, special water and electricity boxes are placed next to the track. During the activity, the train must be parked.
- **Exterior washing of trains:** For this activity, a Train Washing Installation (TWI) is necessary. A train moves slowly through the TWI while the train is washed automatically. Next to the TWI, a train driver is necessary for this activity.
- **Minor reparations to trains:** Certain minor reparations can be executed at the yard, and on some yards, there are special facilities like lifting ramps, a covered track or a piece of track where the power of the overhead line can be switched off. Not for every type of repair these facilities are necessary.
- **Shunting:** This activity is necessary to connect all the activities listed above. Shunting is the movement of a train(unit). This can be executed on the central operated area (Centraal Bediend Gebied (CBG)) or not central operated area (Niet centraal bediend gebied (NCBG)). For this activity, a train driver is necessary.
- **Coupling and splitting of train units:** Train compositions can consist of multiple train units. It might be necessary to couple train units to create a combination or to split train unit combinations in order to prepare trains for their next service. It is also possible to change the composition of a train internally, meaning that train units need to be swapped (the front train unit needs to be rear unit and vice versa). For this activity, a train driver is necessary and two tracks are needed for parking of which one for a short amount of time.

Complexity of node planning

Cordeau et al. (1998), who produced an overview of 153 research papers about rail and timetable planning, admitted that rail activities are in general complex because they involve large-scale systems. The approach in the industry has been to separate planning activities in multiple components so that manageable subsystems were created. However, routing and scheduling problems have important interactions that are not taken into account in these subsystems. Therefore, there is a need for models that integrate multiple of these subsystems. The planning process of a node is complex because it is dependent on the timetable planning, different departments are planning it, and it is node-specific. These conditions make it hard to design a PAT that should integrate the complete planning process while at the same time, provide the amount of specific detail that is being produced by planners at this moment. Because of a hierarchical structure, all planners are dealing with partial problems that need to be solved before deadlines after which the complete planning is being forwarded to planners of the next department. The problems that arise differ over time from big and global problems in the early stages of planning to small and detailed problems at the end phase of the planning. One can understand that the goals of planners of different departments differ as well. Implementing and introducing a PAT within this environment is challenging.

1.3. Scope

This problem will be scoped in different fields. First of all, this research will only look at the node planning problem of passenger trains. Freight trains are out of scope. Next to that, the geographical area of the research field is limited to a node. For an exact definition of a node, see Section 1.2, but a node consists of a train station and the surrounding shunting yards, which are also called Service Locations. It means that this research will only focus on node planning, and not on the timetable and rolling stock planning. However, a short introduction will be provided to understand the input of the node planning in Chapter 3.

On the Service Locations, multiple services can be executed to train units. In order to simplify the planning, the planning includes only the interior cleaning of train units and the mechanical checks for train units. There are two types of mechanical checks for train units: A and B checks. The difference is the intensity of the check, which affects the (estimated) duration of the check.

From the planning process perspective, the focus of this research is on the operational phase of the planning. Time-wise, this is from X - 1,5 year until the time of execution (X). We can split up this operational phase of the planning in two parts: the preparing part and the execution part. Dispatchers execute prepared plannings but also have to deal with real-life information and disturbances. If due to disturbances plannings are not feasible any more, dispatchers will create a new plan that will be executed immediately. Because time constraints play a more significant role, the planning technique of dispatchers are different compared to the departments that prepare plannings in the long term. This research focuses on the preparation and creation of plans in the long term, resulting in the exclusion of dispatchers of the research area. The point where the qualified plan is handed over to the dispatchers is X - 56 hours. This deadline means that the planning timeline that will be discussed in this research is from x - 1,5 year until 56 hours before the operation.

To focus on the different types of resource plannings processes that take place at a node described in Section 1.2, the focus within this research will be on the first two: infrastructure process and equipment process. The third type (personnel process) will be out of scope since this part is not seen as the problem within the complete node planning. However, it is important to know how many personnel is available at a specific node, since personnel is necessary to perform certain services.

The last point that needs to be defined is the PAT. For this research, a Hybrid Integral Planning tool (HIP) is used as the PAT. This tool is evolved from the Train Unit Shunting Problem (TUSP), which is an overarching name of all the sub-problems that come into play when planning shunting activities. Although this PAT is still being developed at the moment this research is performed, it is the most advanced TUSP solver at this moment.

1.4. Objectives

This research has two clear objectives. The first objective is **to determine which aspects are missing in Planning Assistance Tools for railway node planning processes**. The term *aspects* refers to all objective and rational reasons why a human solution of a planning assignment is different compared to a computer-based solution. These aspects can be measurable as well as immeasurable. Examples of measurable aspects are the number of shunting movements, driving times for trains or switch changes. Examples of immeasurable aspects are safety rules, preferred routes and robustness rules. It might also be possible that a human aspect is already taken into account by the PAT, but that the weighing of the aspect regarding other aspects does not correspond to the weights of the aspect for a human planner, regarding the fact that humans do not use precise weights in their decision making.

The second objective of this research is **to find out what the implementation of aspects based on human planning observations in the PAT does to the solution of the PAT and the acceptance of human planners**. In other words, does the solution of the PAT with other aspects resembles the solution of a human planner compared to the solution of the original PAT. This objective assumes that it is possible to add (certain) aspects, while this might not be the case. However, it is expected that there have to be some aspects that are possible to implement in the algorithm.

1.5. Research question and sub-questions

The objectives can be seen as a goal, though research is never complete without an overarching research question. The research question that is proposed is: **what does the addition of missing aspects based on human planning observations do to the solution of a Planning Assistance Tool?**

To be able to answer the research question, the following sub-questions are proposed:

1. What is the best method to research the qualitative difference between plans for complex planning assignments like railway node planning?
2. How is the integrated node planning process of NS set up?
3. What are the characteristics of the developed Planning Assistance Tool?
4. How is it possible to identify lacking aspects of a Planning Assistance Tool using an empirical experiment?
5. What is the impact of the integration of missing aspects to the PAT?

1.6. Outline

The methodology proposed to answer the research question is made visible in Figure 1.2. The arrows present information streams. The blue rectangular areas represent the Chapters of the report where the subjects will be described.

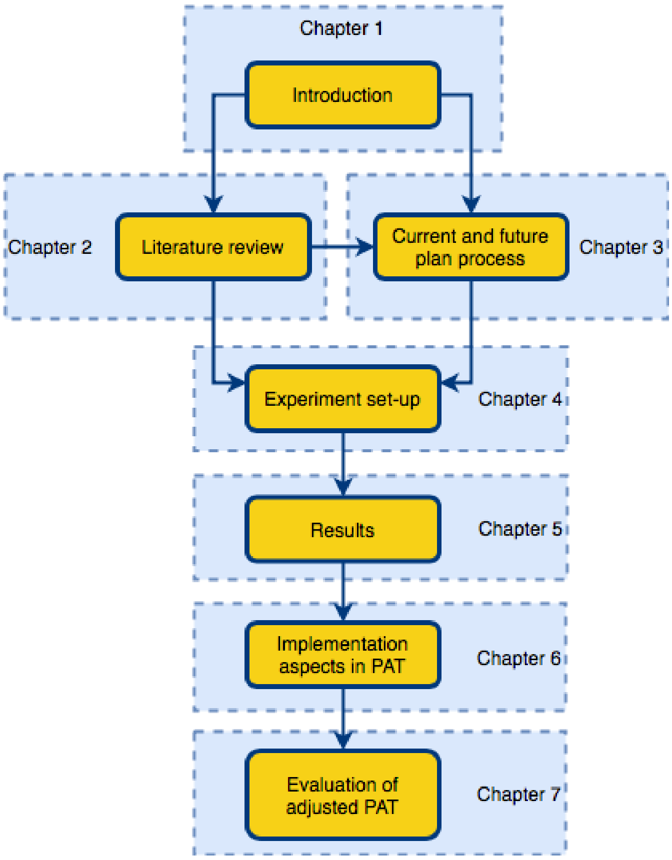

Figure 1.2: Methodology

Literature review

A literature review provides more information about the research that has already been executed. First of all, background information is presented about complex production planning and the implementation of a PAT in complex planning processes. The main matter starts off with an explanation of the algorithmic approach to solve node planning problems within the rail industry. Next, the relation between human and computer optimization is discussed using relevant articles. The fifth subject will be focused on the empirical method to capture human aspects of node planners and the generation of the choice set that is necessary for the experiments. The literature review will be used to answer sub-question 1.

Current planning process and description of PAT

The node planning process as described in Section 1.2 needs to be analyzed in order to collect all the characteristics of the planning process as a whole, the different planning departments and the different layouts of nodes. First of all, the overall planning process is identified and coupled with the hierarchical structure described in the literature review. Secondly, a description is provided what kind of planning tools planners use at this moment. With this information presented, the different plan environments are described per planning department. The planning assignment for every department is defined. In other words: the starting point of a planner and the extra value added are presented. Next to the planning assignment, the solution space per department is described as well. The solution space provides a certain level of freedom in decision making for planners. The complete planning process is summarized with the help of an IDEF0 (Integrated DEFinition for Function model). This model presents all the information streams between the different planning departments. An IDEF0 can be used to describe the data flow and system control of the planning process and is designed to model the decisions, actions and activities of an organization or system (Grover and Kettinger, 1999). Directly related to the planning of nodes are the different layout types of nodes, since the degree of freedom to move and service trains is depending on this. The last subject that needs further analysis is the PAT that is used for this research. The complete idea behind the PAT is explained, just as the necessary input and the visualisations of the solutions of the PAT.

Empirical set-up

The goal of this part is to discover which planning aspects are important for human planners, what the differences are between the solutions of a human and of a PAT, and what the motivations are behind these differences. An empirical experiment with planners is used to collect information. Within the empirical experiment, an experienced planner compares its own solution and the solution of a PAT of the same real-life planning process and motivates the choices that were made. The case-study used is explained in more detail. Based on the outcomes, a list of differences and aspects is created. To identify the differences and categorize the larger aspects, a clear system is created.

Results

The list of differences and aspects is presented in this Chapter. Differences are notified between planners. The level of node-specific or generic differences is presented with the use of a verification experiment using a different node. More information regarding the aspects is provided as well. More information about the aspects is explained as well.

Extending HIP

A creation and selection process is presented in order to adjust the original PAT with extensions based on the aspects found by the empirical experiment. With the help of the development team of the PAT, extensions are created that are easy to implement. Based on the priority of the aspect and the level of difficulty of the implementation, the extensions are added to the PAT. The way how the extensions are added is presented as well. The original PAT with the implemented extensions is called the adjusted PAT.

Evaluation of adjusted PAT

The solutions of the adjusted PAT model is compared with human solutions and original PAT solutions, using different planning assignments. First, a qualitative comparison is performed between the human solution and the solution of the adjusted PAT. In order to do so, multiple real-life planning assignments are used together with the same experienced planners of the empirical experiment. Secondly, comparisons between the adjusted PAT and original PAT are executed. These quantitative comparisons are used to decide whether the performance of the adjusted PAT differs compared to the original PAT. Different plannings assignments are used in order to test the adjusted PAT for different levels of complexity.

2
Literature review

This research is not the first time that complex planning processes and the involvement of Planning Assistance Tools are analyzed. The research earlier performed on this topic is summarized in this Chapter. The question that is answered in this Chapter is: **What is the best method to research the qualitative difference between plans for complex planning assignments like railway node planning?** In order to answer this question, this Chapter is divided into two parts: background information and a real literature review. **The background information is presented in Sections 2.1 and 2.2. The literature review is presented in Sections 2.3, 2.4 and 2.5.**

2.1. Complex production planning

As described in the introduction (Chapter 1), the planning process of nodes is a complex problem that includes much information, and that is often prepared a long time before the actual execution. Over time, the planning undergoes many changes due to exterior changes and the increase of available information on one hand. On the other hand, node planning processes are defined as production planning processes: although no new products are created, trains are maintained at Service Locations. Different tactics are used to maintain these trains that are familiar to real production processes like job shop and flow shop processes. Because of these two characteristics, the organization of most of the planning departments that are responsible for the shunting problem worldwide are based on the Hierarchical Production Planning (HPP) method. In order to understand the planning organization of NS, it is helpful to know the characteristics of this method, the advantages, the disadvantages, and for what kind of problems this method can be used. In Section 3.1.1, the current planning process of the NS is related to the HPP method. Relations between the HPP method and node planning for railway transport, in general, are part of this Section.

HPP is oriented on capacity, whereas other planning methods (e.g. materials requirement planning and just in time planning) are oriented on material availability (Meijer, 2012). The method is based on the philosophy that in the hierarchical nature, in which production decisions are made, global production quantities are specified at an aggregate level first, and these quantities are decomposed later to detailed item production lots (Zijm, 2000). The planning is created with aggregated global numbers on the highest level and will be supplemented by more specific numbers and details on lower levels. HPP determines for which product types capacity has to be reserved. At a later point, these capacity reservations are disaggregated to time slots reserved for particular product families within each type. Lastly, the determination is made how much time in each slot should be spent to the production of particular items within each family (Hax and Meal, 1975) (Hayes-Roth and Hayes-Roth, 1979). Figure 2.1 shows an overview of the conceptual hierarchical planning system. Visible are the three steps of aggregation hierarchy (boxes (1), (2) and (3)). The first (1) is the level where product types are determined (Bitran et al., 1981). The planning horizon is usually one year, and the level is concerned with tactical decisions. The second level (box (2)) describes the run quantity of each type is disaggregated to obtain production quantities of each family. These are passed to the item level (box (3)) where they are further disaggregated to determine the amount of each item to be produced in the first period.

But for what production process is HPP a good planning method? HPP is often used for processes at which material complexity is low, and resources are expensive and have to be highly utilized (Zijm, 2000). Next to that, HPP is often used for plannings dealing with capacity allocation decisions across several plants or locations (Steinrücke, 2012). HPP includes long- (strategic), medium- (tactical) and short-term (operational) horizons (O'Reilly et al., 2015). The HPP approach follows the organization's hierarchy with decision making on the higher levels imposing constraints on the lower levels, and the lower levels providing feedback in order to evaluate the higher-level decisions. The arrows at the left side of Figure 2.1 show the feedback to higher levels.

Now that the HPP method is explained, it is interesting to find out in what way the node planning process relates to this method. From literature, Node plannings are most of the time defined as a (flexible) hierarchical planning (Van Wezel and Jorna, 2009). The process of servicing trains has a lot of the characteristics described above. The resources where these services can be executed are scarce (during the night): the shunting areas are full, the number of cleaning teams, mechanics and train drivers are limited, and services can often only be applied at a limited amount of dedicated tracks. From a bigger perspective, there are multiple service locations scattered over the country where these services can be executed. This characteristic relates to the HPP method in a way that capacity allocation decisions have to be made across several plants or locations. The fact that multiple departments are responsible for smaller sub-problems of node planning is also derived from the HPP method.

A description of the advantages of HPP is also necessary to understand why this method is ideal for complex planning problems. First of all, a key strength of the HPP approach is that it breaks down large and complex tasks into various phases, where each phase provides a basis for the decision-making area of the next phase (O'Reilly et al., 2015). The planning method aligns timescale to data availability. Next to this, the plans generated by the HPP tend to be quite stable in the rolling horizon process, because an aggregated forecast tends to be more accurate than a detailed forecast (Bitran et al., 1981). The aggregate forecast also has the advantage that the data collection required is considerably smaller than for a detailed forecast.

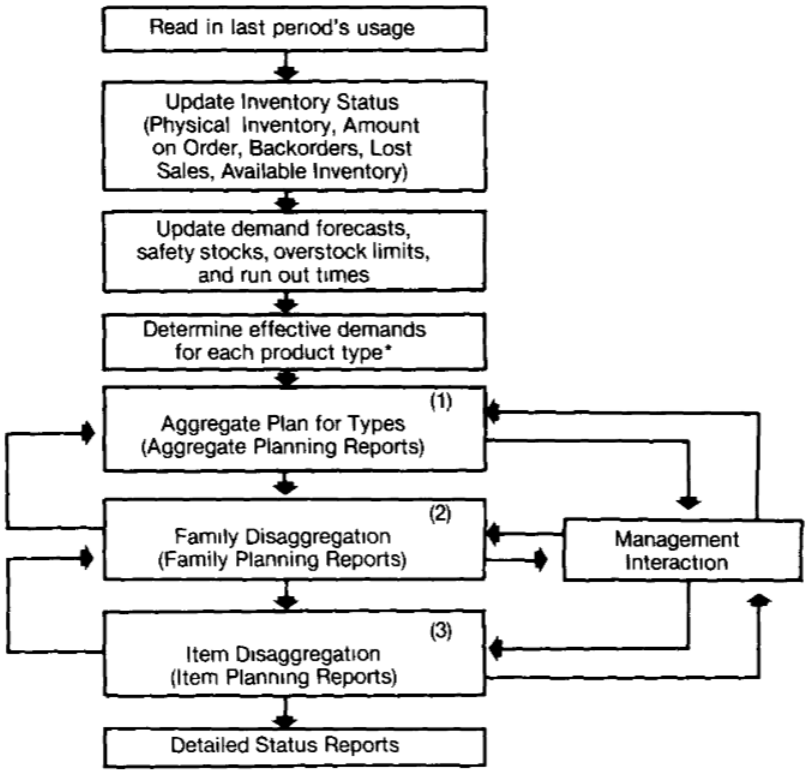

Figure 2.1: Conceptual overview of hierarchical planning system (Bitran et al., 1981)

There are also disadvantages to hierarchical production planning. Riezebos et al. (2011) describes that the amount of human and organizational noise in the planning environment is present because the input for one planner depends on other planners [from a higher level]. In other words: bad performance of planners of a high level causes noise for planners on a lower level.

Summarizing: the essential characteristic of HPP is that the production process is planned on beforehand with a low level of detail, and when more information is available, the aggregate numbers are specified. This process is often used for processes where resources are expensive and are highly utilized. Because of this, HPP fits perfectly to the servicing/maintenance process of trains. The advantage of using HPP is the scale down of large and complex problems into smaller problems that align the level of detail to data availability, resulting in stable plans. The disadvantage is that human and organizational noise is present, which is hard to find and hard to determine who is responsible for this noise.

2.2. Implementing PAT in complex planning process

The development of algorithms and optimization methods is a great innovation, but it has to be noticed that the implementation of the PAT faces a lot of difficulties. Within the rail industry, the problem of implementing a PAT is present as well. Cordeau et al. (1998) created an overview of 153 papers on train routing and scheduling problems and concluded that even though the proposed models are tested on realistic data instances, very few are actually implemented and used in operation. Cordeau et al. state in 1998 that: "Effort must be made to bridge the gap between theory and practice". Over time, not a lot of improvement has been made to fill this gap. In order to face the important reasons behind the implementation challenge of advanced scheduling solutions in practice, De Man et al. (2016) mention that: "recent studies indicate the need for scheduling tools that focus on the actual work routine, characteristics of the scheduling environment and cognitive tasks of the scheduler". Within this Section, the information will be provided about the experience of implementing Planning Assistance Tools for complex planning problems.

In general, the PATs that are developed include different types of software: from interactive systems that allow an automatic check of the feasibility of plannings and schedules to systems where optimal schedules are suggested (Dios and Framinan, 2016). Introducing, implementing and working with a Planning Assistance Tool provides problems for actors in production processes (McKay and Wiers, 2006). A PAT is generally designed independently of the end-user, and psychological relevance is considered to be low (Cegarra and Wezel, 2011). Experience from simulation projects has shown that automatic generated (and optimal) plans by PATs are difficult to interpret, difficult to trust, difficult to compare, mutually contradicting and difficult to act upon for humans (Lorentzon and Fredlund, 2017). Sufficient knowledge has not been gathered about what planning really means and what people do when they plan (McKay and Wiers, 2003). Without knowing the key variables and attributes of this problem, it is difficult to craft solutions that consistently and clearly assist with problem-solving.

But why is the implementation of planning tools in practice hard? What are the problems that can be identified from earlier research? Since the early 1960s, the gap between theory and practice of scheduling and planning has been discussed by a lot of researchers (McKay and Wiers, 2006). The lack of fit between the human element and the formal or systematized element found in the planning methods and systems is the biggest problem. This lack is what McKay and Wiers (2006) call the human factor in the success or failure of planning methodology. According to them, the role and contribution of the human element have been largely ignored and under-researched compared to the effort placed on the mathematical and software aspect. Luckily, a lot of new research is available at the moment that describes the human element as depicted by McKay and Wiers (2006). A selection of aspects will be described underneath, collected from different research papers:

- **Difference in the interpretation of constraints.** Part of the missing human element can be explained by the constraints that are used in algorithms. One of the problems that arise in the development of the tools is the fact that the constraints implemented in the tool do not resemble the same constraints that human planners have (Dios and Framinan, 2016) (Arica et al., 2014). Constraints are added in tools as 'hard-' or 'soft-' constraints. The hard constraints are clear, while the soft constraints are vague and hard to add in an algorithm. The weight of all these constraints to each other differ from planner to planner. Dios and Framinan (2016) state that a planning tool should provide the user a system to handle the soft constraints. Another problem that is remarked by Framinan and Ruiz (2012) are 'hidden' constraints. These constraints are not recognized during the development of the tool, but are considered by human planners nevertheless. The problem regarding the constraints result in a perceived problem and solution space for the model (Arica et al., 2014), while in real life it is reasonable to assume that there is more planning flexibility than any system can model (Moscoso et al., 2010).

- **The lack of human expertise in PAT.** Humans often contain expertise that is hard to implement in planning tools. The incorporation of human expertise into planning tools has been controversial for a long time (Dios and Framinan, 2016). From a review of many manufacturing scheduling tools, a limited number of tools were found that reported this functionality. In most cases the human expertise is confined to a sort parametrization. In some cases, the expertise is translated to a pure knowledge-based tool.

- **Misinterpretation of reality and the difference in solutions space.** Reality is different compared to the description or assumption of reality by a computer. Key data inserted in an algorithm does not always resemble the real data or the estimated data (Dios and Framinan, 2016) (De Man et al., 2016). Next to that, the solution space of a computer is different and less flexible compared to the solution space of planners. This results in solutions that are not realistic or not optimized for real-life applicability. This can result in two extremes: the planning is unrealistic and unfeasible (input data is to strict), or the planning is too robust, minimizing the utilization of resources (De Man et al., 2016). These scenario's restrict the usability of the tool for the user, who has a better view of the reality and more flexibility in the solution space and therefore can act with greater flexibility and adaptability (McKay et al., 1989).

- **The lack of user input and judgement in PATs.** Rescheduling is a key activity of human planners. Although rescheduling is considered an integral part of planning tools, algorithms are simple or not driven by the user judgement (Dios and Framinan, 2016). This problem is again part of the human aspect. It results in users only using the planning tool for a initial planning and instead decides to reschedule by hand, while algorithms, especially in these days, can be advanced in rescheduling.

- **A Graphical User Interface (GUI) that is not created together with end-users.** For planners, it is important to work easily with a planning tool. This last point of the human aspect considers the Graphical User Interface. Visualization plays an important role in the successful implementation of systems (Arica et al., 2014). The review of many production scheduling tools showed that a lot of tools were not implemented because of a malfunctioning or not desired GUIs (Dios and Framinan, 2016) (De Man et al., 2016).

To conclude, the problems of implementing a PAT in complex planning processes are two-sided: the missing human element (the lack of human factors) and problems that are not human-based. Scoping on the human element the problem can be found on five different aspects: the difference of the interpretation of constraints, the lack of human expertise in PAT, the misinterpretation of reality (problem space) and solution space, the lack of human input and preferences when using the PAT and problems with the Graphical User Interface.

2.3. Algorithmic approach to solve node planning

In this Section, a general overview is provided of the existing literature of solving the passenger train shunting problem with algorithms and optimization methods. The Train Unit Shunting Problem (TUSP) is the overall name for the planning problem of Service Locations (or shunting yards). The development of algorithms and optimization tools to solve this problem is important for this research because the PAT that is going to be used for this research (HIP) is partly based on these developments. Next to that, the literature clearly shows that the complex TUSP can be divided into smaller problems. Understanding the sub-problems gives a better view of the complexity of node planning processes. These sub-problems are (Den Ouden, 2018):

1. Matching train units into combinations: self-propelled train units can drive in combinations, which means that train units may start the shunting process combined and may need to leave the process in a different combination. An example is that an ICM III + IV + III enters the SL, while the combination ICM IV and ICM III + III need to leave the SL. A match between the incoming and outgoing combinations needs to be found, where each incoming unit is matched to a component of an outgoing combination.

2. Job shop scheduling for non-human resources (servicing): the services that need to be performed on a train unit are distributed over the non-human resources. For every service, specific tracks are allocated, meaning that a train unit can only undergo a certain service if it is at a certain track where that service is allowed.

3. Parking: the node also serves as a parking space for (combined) train units when no service is performed. For this, there is a set of stabling tracks. A unit needs to be assigned to these stabling tracks in order to make room on service tracks.

4. Routing: train units need to be routed over the node in order to move from the platforms to the stabling tracks and service tracks. In order to do this safely, strict rules apply that should be taken into account.

5. Splitting and combining: in order to change the combinations of the train units, combinations have to be split into separate units and have to be combined later. This action can be done during the complete process. The problem is when and where this should happen for a unit.

6. Crew assignment: In order to perform services and to move trains, crew needs to be assigned to the tasks. This process includes maintenance and cleaning tasks as well as shunting (the movement of trains over the node) tasks. There are specific constraints concerning this problem like workloads, mandated break times, walking times and others.

In order to structure the information provided in this Section, Table 2.1 shows which sub-problems are integrated into the research and the proposed models of the different authors noticed in this literature review.

Table 2.1: The sub-problems incorporated per literature that is reviewed in this Section

	Sub-problems					
Source	**Matching**	**Servicing**	**Parking**	**Routing**	**Splitting & combining**	**Crew**
Freling et al. (2003)	x		x			
Lentink et al. (2003)	x		x	x		
Kroon et al. (2008)	x		x	x		
Van Den Akker et al. (2008)	x		x	x		
Jekkers (2009)	x		x	x		
Lentink (2006)	x		x	x		x
Jacobsen and Pisinger (2011)	x	x	x	x		
Opstel Plan Generator (OPG) (NS, 2018a)	x		x	x	x	
Van Dommelen (2015)	x	x	x	x		
Van Den Broek (2016)	x	x	x	x	x	
Haahr et al. (2017)	x		x			
Den Ouden (2018)	x	x	x	x	x	x
Hybrid Integral Plan method (HIP) (Den Ouden, 2019)	x	x	x	x	x	
Adapted HIP (Baurichter, 2020)	x	x	x	x	x	

The Train Unit Shunting Problem (TUSP) was introduced by Freling et al. (2003). The problems consist of the matching of arriving trains to departing trains and the parking of these train units on a track. The train units can be combined to form a longer train, and vice versa. The matching problem assigns parts of arriving trains to corresponding blocks of train units of departing trains. The objective of the matching problem is to minimize the number of actions of splitting arriving trains in order to position a train unit in a departing train. The problems are solved using a standard MIP (Mixed-Integer Programming) solver (CPLEX). The parking problem is solved after the matching problem. The problem is solved using a column generation approach in order to find a feasible park plan. The feasibility lies in the fact that the train should fit on the track and the train should be able to leave this track on a determined time without being blocked by other trains. A dynamic programming algorithm solves this problem.

The problem environment of Freling et al. (2003) does not capture the real-life shunting problem. The most important thing that is missing is the routing to the stabling tracks. Within the model, routes may be crossing, which is not possible in real-life. Lentink et al. (2003) proceeded with the idea by including the routing problem. With the use of a graph representation of the shunting yard, routing costs are estimated from and to stabling tracks. This information is added to the column generation approach. The last step is to compute the actual routes taking into regard the track occupation resulting from the column generation approach. Trains can move simultaneously, but the entire path of a movement of a train is reserved for the duration of the movement. The routes are computed sequentially. The sequence of train movements are evaluated by a local search approach that can swap the order of two movements in order to improve the results.

Whereas Freling et al. (2003) clearly separates the matching and parking tasks and solves it sequentially, Kroon et al. (2008) created a model that solves the two sub-problems simultaneously. This more complex environment resulted in a large increase of constraints. To reduce the number of constraints, the authors grouped them in clique constraints. Although this allowed them to find feasible solutions for a test case, for larger cases the computation time increased rapidly, making the tool not usable for real-life application.

The simultaneous execution of tasks have also been researched by Van Den Akker et al. (2008). The authors include waiting time for arrival and departing trains at the platforms, which is extra flexibility that is not included in earlier algorithms. The shunting of parked trains to other tracks is also possible, which also allows more flexibility in shunting plans. Because of these new properties, there is an exponential increase in variables and constraints, which makes it harder to include them in the linear programming approaches that are proposed by earlier research. To solve this problem, Van Den Akker et al. (2008) used a greedy heuristic and an exact dynamic programming algorithm to solve the matching and parking problem. This combination decreased the solving time, but the model was not able to find feasible solutions for complex problems.

Jekkers (2009) proceeded with the integral solving possibility by presenting two Genetic Algorithms (GA) for the matching, parking and routing problems, including waiting times at arrival and departure platforms. This approach is applied successfully to generate shunting plans for real yards within a reasonable time.

Where Van Den Akker et al. and Jekkers tried to solve the problem integral, Lentink (2006) studied a practical extension of the TUSP by adding the cleaning task to the planning problem. The matching, parking and routing problems where solved using methods proposed in earlier work (Freling et al., 2003) (Lentink et al., 2003). The cleaning problem only focused on the crew necessary (so not yet the track resources) and had the objective to minimize the sum of the completion times. This machine scheduling problem, where the crew were seen as machines, could be solved by CPLEX.

Jacobsen and Pisinger (2011) combined the integral planning idea with the fact that certain services can only be applied at certain tracks (or locations) on the shunting yard. Within their plan sequence, each train needed maintenance at one facility or workshop but also needed to be parked before and after their service task. Using a mix of different algorithms (three Meta-Heuristics, Guided Local Search, Guided Fast Local Search and Simulated Annealing), the authors tried to make plans such that no trains were blocked by other trains, no departure delays occurred, and the makespan of the service tasks was minimized. Although the model was faster compared to the MIP solver, the absence of the routing and matching problem made it not directly applicable to be used in real life.

The NS has created a tool internally that is known as the Opstel Plan Generator (OPG, 'stabling plan generator' in English) which is based on the model created by Kroon et al. (2008). The OPG can determine the matching for arriving and departing train units and the track assignment for parking and the routing (NS, 2018a). The OPG also decides when and where trains are split and combined. The OPG does not take into account service tasks (BMI Projectteam, 2018). The solving method is based on the shortest path algorithm by Dijkstra and MIP.

Van Dommelen (2015) tried to generate feasible shunt plans that include servicing, parking and routing given a fixed matching. By modelling the order of services as a flow shop problem using CPLEX, a sequence of movements was created. This sequence, which consisted of parking intervals per train unit, was used as input for the OPG. It determined parking locations and routes. A feasible schedule was found for a test case, though without guarantee of routing conflicts.

Van Den Broek (2016) used a Simulated Annealing (Stochastic Local Search) approach to solve the TUSP, with some aspects being solved by heuristics such as tabu search. The model starts with the generation of an initial shunting plan which is optimized by the cost function that includes weighted factors. A single change is made to the plan and the plan is optimized again, resulting in a new value of the cost function. If the cost function performs better compared to the last one, the new solution is accepted. Otherwise, the new solution is rejected. This process runs until the runtime is over. The outcomes of the model outperformed the OPG. The model is not able to cope with disturbances in the input, so for the short-term planning, it is not applicable.

Haahr et al. (2017) focused on the core matching and parking problem and tried to solve the problem with three different methods: Constraint Programming (CP), Column Generation (CG) and Randomized Greedy Construction Heuristic (RGCH). The CP method is only available to solve the problem for small instances. The CG method proved to be outperformed by the other two methods. The RGCH selects a track based on three criteria: the same type of train, empty track or track with sufficient capacity. This method was able to solve almost all instances within one second. However, routing is not part of this solution.

Den Ouden (2018) wanted to expand the work produced so far with the final sub-problem (6) since this sub-problem is not taken into account with the combination of sub-problems 1 -4 being answered. Lentink (2006) provided a crew assignment problem, but did not integrate the job-shop scheduling for non-human resources, minimizing the walking times and required personnel. Sub-problem 5 was never taken into account (except for the OPG), but assumptions were made by Van Den Broek (2016) that all splitting of train units will happen immediately on entry and combining will happen just before exiting the SL. Next to that, an assumption was made that there was always a sufficient number of human-resources, while this might not be the case. Goal for Den Ouden (2018) is to solve all the problems, but with sub-problem 5 still being an assumption. Den Ouden used a greedy heuristic to obtain an initial solution of the problem, after which a local search improvement step optimizes the combination of the flexibility, fairness and walking distances. The method created can consistently find successful results for realistic scenarios.

At the moment, the NS is working on a Hybrid Integral Planning tool (HIP). More information about HIP is provided in Section 3.2, but the big differences compared to other tools described in this Section will be provided here. Earlier described Planning Assistance Tools all solved the sub-problems sequentially, HIP solves the sub-problems parallel (BMI Projectteam, 2018). To provide an example: earlier planning tools first solved sub-problem 1, after which sub-problem 2 and 3 were solved. By doing this, results of sub-problem 3 depend on results of sub-problem 2 and 1, but results of sub-problem 1 do not depend on result of sub-problems 2 and 3. Integrating the solving of the different sub-problems has the advantage that all the sub-problems are depending on each other. A more practical example is that first the capacity of the infrastructure is being checked, after which the capacity of the crew is being checked. With HIP, these two capacities are checked at the same moment, providing feedback to each other. If the crew capacity is reached, this information will be taken into account when solving the infrastructure capacity and vice versa. In order to be able to plan this, multiple planning methods are applied to find an initial plan (Den Ouden, 2019). These planning methods communicate with each other in order to form a hybrid system that is able to produce one initial solution (Snijders, 2019). Local Search Functions perform the optimization of this initial solution. Local Search Functions make small changes in order to find improvements.

Next to the fact that HIP integrates the sub-problems, the scope of HIP is bigger compared to the tools described above. Where the tools developed earlier are only focused on the Service Location, HIP takes into account the complete node. This difference means that HIP also plans movements of trains between the Service Location and station. Trains that end their service will be scheduled with a specific route and time towards the SL, and trains that need to start a passenger moving service will be planned a route from the SL towards the station. An extra factor that needs to be taken into account is the (timetable) traffic that also uses this infrastructure.

The last point about HIP for this research is to describe the current state of development. The goal of the developers of HIP is to integrate all the sub-problems into the tool. That is why Table 2.1 shows an X at every sub-problem. However, at this moment, the crew assignment planning is not yet integrated into HIP (BMI Projectteam, 2018). Because of this the adapted HIP method, which is the result of this research, will not take into account the crew assignment problem.

This research provides the adapted HIP tool, which is an adaptation of HIP that includes aspects that are derived from real-life planning solutions. At this moment, HIP is only tested in (hypothetical) testing environments and not yet in real life. By testing HIP in real life and comparing its solutions to real human generated solutions, differences can be identified which lead to aspects in HIP that are missing, overvalued or undervalued. Adaptations will be made to the algorithm to take into account these aspects.

To summarize this Section, it is clear that the complete shunting problem can be divided into 6 sub-problems: matching, job-shop scheduling for non-human resources, parking, routing, splitting and combining and crew assignment. Over the years, researchers have tried to solve these sub-problems sequentially, but there has never been an attempt to solve the sub-problems integral. Especially the combining and splitting problem (when and where a train needs to be combined or split up) is a sub-problem that is lacking research since it is based on assumptions except for the OPG. The way the sub-problems were solved differs from researcher to researcher, which has a direct relation to the solving time. For some algorithms, the solving time can be set, meaning that a longer solving time will produce a better result. For other solvers, a small problem can be solved within a reasonable time, but to solve more significant problems the solving time increase rapidly. Some algorithms are able to solve simple problems but are infeasible to solve complex problems. HIP is the first tool that can solve the sub-problems parallel instead of sequential by using a hybrid system that uses different planning methods to solve the sub-problems. HIP also takes into account the complete node and does not only focus on the Service Locations. Because HIP is still under development, HIP does not yet solve the crew assignment problem. The adapted HIP tool takes into account aspects that are derived from real-life testing. These aspects found are missing, undervalued or overvalued in the original HIP tool.

2.4. Computer and human cooperation

Although planning tools can solve plannings problems on their own, it is not wanted to let these systems take over the task of human planners. Within a lot of operational industries, human actors remain an essential link in the chain (e.g. pilots in passenger aircraft). The most important reason is the fact that computers might fail or might not work in which the human needs to take over the task (Hoc, 2000). A loss of skills of humans in unwanted in these situations. There are also human characteristics that are beneficial and needed to solve complex problems (Sanderson, 1989). Sanderson is one of the first researchers that acknowledged that humans and algorithms seem to have complementary strengths which could be combined. In this Section, these strengths are explained in order to understand why cooperation between humans and computers will result in the best planning possible.

Humans do have qualities that are necessary for real-life planning operations that computers are lacking. Humans are flexible and adaptable, giving them the power to cope with stated, not-stated, incomplete, erroneous and outdated goals and constraints (McKay et al., 1989). Human planners can also communicate and negotiate within their process, but also with third parties. At last, humans have the intuition that can help to fill the blanks of missing information to schedule. Since production and logistical processes often cope with these situations, humans must take part in these planning processes.

However, humans are not good at everything. For planning purposes information processing is essential. Unfortunately, this is not a strength of humans. The limitations of information processing capabilities of humans can be categorized into two factors: bounded rationality and incomplete problem representation. Bounded rationality is the cause of the limited mental capacity of humans, like information processing capabilities (Sanderson, 1989). Human memory is lacking compared to the memory of a computer, which is therefore ideal to use when information needs to be processed. The second point is caused by the fact that for human beings, it is impossible to have all the needed information on hand (McKay and Wiers, 2006). This missing knowledge leads to an incomplete problem representation, resulting in longer solving times or possible infeasible solution options compared to computers. The fast computing time and the large storage capacity are the advantages of computers, that can be beneficial when solving complex problems.

Concluding, beneficial characteristics of humans are flexibility, adaptability, intuition, the possibility to communicate and to negotiate. With these skills, humans are preferred above computer when cooping with stated, not-stated, incomplete, outdated and missing information (goals, constraints and data). Computers are preferred above humans when dealing with information processing because of its great memory and fast computing times. The mix of these characteristics of humans and computers are beneficial when solving complex problems.

2.5. Empirical experimenting with planners

In order to discover aspects that are missing, undervalued or overvalued in the current PAT, an empirical experiment in which the solution of the PAT will be compared to the results of the human solution is performed in this research. Research on comparisons between the results of computer-generated plannings and human-generated planning is provided in this Section. This information is helpful for the set-up of the empirical experiment and the execution of the experiment. Next to this, the results of the papers might also contain relevant information, although, for many cases, this information cannot be transferred entirely since the domain of the planning is different. First, an overview of the empirical methods is provided. Next, relevant research is summarized in order to understand experiments executed, the results and the experience. At last, information is presented that is applied for the experiment of this research.

There are different reasons why empirical experiments are executed on planners and operators in combination with Planning Assistance Tools. They usually have the goal to formulate and implement rule-based systems or heuristics with human-based rules of thumb (Van Wezel and Jorna, 2009). Sanderson (1989), Crawford et al. (1999) and Lauria and Wagner (2006) have executed extensive overviews of empirical studies on planning and scheduling problems. Since these overviews only include relatively old paper (mostly all before 2000), they are not described in this Section. However, these overviews are used in order to define a framework that describes the different types of empirical experiments executed on this topic. The following types of empirical experiments will be taken into account:

1. Laboratory studies;
2. Reviews of human scheduling behavior;
3. Surveys;
4. Case studies;
5. Longitudinal studies;
6. Field studies.

Table 2.2: Overview of literature described in this Section with important characteristics of experiment.

	Type of experiment									
Research	**Laboratory study**	**Review**	**Survey**	**Case study**	**Longitudinal study**	**Field study**	**Quant.**	**Qual.**	**Domain-specific**	**Domain**
Riezebos et al. (2011)				x		x	x		x	Shunting problem
Mebarki et al. (2013)	x						x		x	Production process
Käki et al. (2019)			x				x		x	Hydro Power Plant
Myers and Lee (1999)	x							x		
Coman and Muñoz-Avila (2011)				x				x		
Berglund and Karltun (2007)						x		x	x	Production process
Baurichter (2020)				x		x	x	x	x	Shunting problem

Zooming in on the comparison between human versus computer-generated plannings, there are two types: quantitative and qualitative comparisons. Quantitative comparisons are domain-independent and do not require extensive knowledge-engineering effort (Coman and Muñoz-Avila, 2011). However, they can fail to reflect differences that are truly meaningful to users. Quantitative plan comparison, therefore, generally consists of counting the plan elements which plans have, or do not have, in common. Plan elements are most often described as actions which are not interpreted in any domain-specific way. They do not contain any extra information. An example of the cooking domain is the fact that adding lemon juice, vinegar and sugar are all the same actions in the eyes of the planner.

While quantitative comparisons are easy to execute because no domain-specific knowledge is needed, qualitative comparisons are based on domain-specific characteristics which human experts might use to differentiate between plans, thus being able to produce results of greater practical value (Coman and Muñoz-Avila, 2011). In other words: qualitative plan comparison should "see" plans much like human users would. The focus points transfer from numbers to characteristics like robustness, risk and degree of preferences. The downside of qualitative comparisons is that they are knowledge intensive since knowledge about the complete process, planing problem and solution space is necessary. The elements (or actions) that include a planning are all interpreted by the planner. To refer back to the example of the cooking planning: Lemon juice and vinegar are both sour, while sugar is sweet. This gives the planner the possibility to vary the set of features along which he or she would like to see the differences (eg. the lemon juice action can be removed if more vinegar is added to the dish).

In order to create an overview of the literature described in this Section, Table 2.2 is created. It identifies the characteristics of the empirical experiments that are performed in the research. First of all, the type of experiment is noticed. It also states whether the experiment is quantitative or qualitative. The last characteristic is if the experiment is domain-specific or not. If domain-specific, the domain is named as well. First, the quantitative comparisons are described, after which the qualitative comparisons are described.

Riezebos et al. (2011) analyzed the strategy of planners by joining them in solving the problem for a specific case: the shunting problem of the node Zwolle. The planner was requested to think out loud while planning. The Thinking Aloud Protocols (TAP) were analyzed beforehand, after which sessions with planners were held in which the researchers actively asked for and discussed explanations of decisions. The researchers performed a quantitative comparison between the plan from a planning algorithm and the plan generated by a human planner. The aspects of this comparison are summarized in the Table made visible in Figure 2.2. The algorithm was able to solve the problem within 3 minutes, while a human planner used a lot more time. Next to the solving time, the plans differed significantly in two aspects. The first aspect is the number of trains that are located at a trespassing track for more than one hour. In more detail, planners allowed half the trains that needed to be cleaned to stay for a time longer than 1 hour at such a track. The verbal description of human planners of the objectives and constraints of this problem pointed towards avoiding such behaviour, while in real life planners aimed at a balance between decreasing the number of train movements and avoiding usage of these tracks. Referring back to the experience of implementing a PAT (Section 2.2), this is a clear example of a difference of interpretation of constraints. The second aspect is that planners allowed for combined movements of trains, which reduces the total number of movements as well as the number of saw movements. Saw movements are movements of trains in which a change of direction is necessary in order to arrive at a platform or a shunting track (van den Broek and Lentink, 2008). The algorithm only allowed movements of single trains (so no train compositions). After the possibility of combined movements of trains was implemented in the model, the outcomes of the model resembled the decision of the planner much more.

Schedule	Experiment	
	Model	Planner
Number of trains located at a trespassing track for more than 1 h	1	11
Number of movements	95	60
Single train	95	32
Connected trains (2 or 3)	0	29
Number of saw-movements	21	18
CPU Time (s)	187	–

Figure 2.2: Table with characteristics between the plans created by a model and by a human planner for the node Zwolle (Riezebos et al., 2011).

Mebarki et al. (2013) used a laboratory study in order to compare the quantitative differences between 4 groups that need to use a PAT for a production process. For his research students were used as research group instead of experienced planners. The difference between the groups was the number of criteria to focus on and the option to receive good next planning options by the PAT. The fact that a laboratory set up is used in combination with students results in the fact that characteristics such as experience, habit and real-life interaction are not taken into account. In order to perform an experiment that lies as close as possible to real-life situations, these sort of characteristics are essential to take into account.

Käki et al. (2019) developed an intervention process in order to find out why planners do not follow recommendations from model-based decision support systems (like planning algorithms). By comparing the behaviour of planners with the recommendations provided by a PAT, insights can be provided. The behavioural data is collected using a survey. Even though planners make some adjustments that have a positive impact, most of their adjustments have a negative impact. The results are mostly quantitative, but the survey provides some qualitative aspects. However, the simple survey provides superficial data which makes it harder to find underlying reasons that explain the behaviour.

Myers and Lee (1999) have performed research to the possibility if plans can be generated that are qualitative different compared to the original plan. The creation of these plans is rooted by the creation of biases that focus the planner toward solutions with specific attributes (e.g. comfort, affordability, etc.). These biases are derived from analysis of a domain metatheory and enforced through compilation into preferences over planning decisions. A metatheory is a theory to distinguish three types of characteristics of planning based on the planner: features, roles and measures. Features can be seen as a preference of an individual. An example can be the choice of transport mode for specific routes. Roles correspond to the details of action: a movement action has an origin, destination and a carrier. Measures can be seen as a description of the solution space or the goal. For example, if the measure is fast, the relation can be made that air transport will be ranked higher compared to transportation over water. The bias distance, which measures the extent to which a choice satisfies a stated bias, is used to generate different qualitative plans. This algorithm is tested by the use of a laboratory experiment where generated plans ranged at a length from 2 to 14 actions that were not domain-specific. The biasing technique provides a low cost, simple mechanism to generate plans that are qualitatively different reliably. However, the algorithm cannot be used to identify qualitative differences between two plannings. Because of this, the experiment is not used for this research, but the metatheory is interesting to take into account concerning the preferences of planners concerning specific optimization goals.

Coman and Muñoz-Avila (2011) are the first that perform a qualitative comparison to a case-based planning context. In order to do so, they use distance metrics differences between two plannings. The usage of distance metrics has the purpose of minimizing the domain-specific knowledge necessary in order to find qualitative differences between two plannings. The results show that the standard deviation of the qualitative comparisons is higher compared to the standard deviation found for the quantitative comparison. However, the questions arise if this information can be used to answer the difference in qualitative characteristics described earlier like robustness, risk and preferences. The experiment adds only one small element to an action (which weapon is used to attack) on which the complete qualitative comparison is based. Next to that, the case-study is extremely simple: actions can only take place sequentially, the number of possible actions is low, and the solution space is small. This research is interesting from a global perspective, but the goal to not incorporate domain-specific knowledge makes it not useful for the research described in this paper.

Berglund and Karltun (2007) compared real plannings of different production processes on a qualitative manner. The data was collected by interviews and observations (summarized as field study: 6). Although the data includes super precise plannings of these case studies, in order to compare them, they had to be generalized and made quantitative. This generalization and quantification step have a big impact on the results since sources why planners of different production plannings act different cannot be found. However, this research shows that in order to collect detailed information, field studies next to experienced planners are important.

Concerning the contribution of this research (Baurichter, 2020) to the literature, the experiment that is executed is a mixture between a field study and a case study. From Berglund and Karltun (2007) it is clear that field study (sitting next to the planner) is the best way of obtaining the most detailed information. Since the goal is to compare the planning of the PAT with a real-life planning, a field study is the best option to obtain the data. Since the research area is limited, the experiment can also be seen as a case study, just as the experiment of Riezebos et al. (2011). A laboratory study is not chosen, because it lacks aspects of reality (Mebarki et al., 2013) (Myers and Lee, 1999), and surveys do not provide enough information in order to perform a thorough qualitative analysis because the content is too specific and it would require to much time for planners to answer (Käki et al., 2019). In addition to the fact that the qualitative studies described in this Section do not have the depth required, it appears from several sources (Lauria and Wagner, 2006) (McKay and Wiers, 2006) (Coman and Muñoz-Avila, 2011) that qualitative comparisons are lacking within the literature. The experiment of this research is qualitative. The experiment will also be domain-specific, as the non-domain-specific experiments are not covering the complete qualitative requirements necessary (Myers and Lee, 1999) (Coman and Muñoz-Avila, 2011). Other important lessons that are important for the experiment is the fact that experienced planners are needed if comparisons are needed in a real-life situation (Mebarki et al., 2013). Experienced planners have much knowledge that researchers or students do not have.

Within this Section, different empirical experiments are described based on the type of experiment, the approach of the experiment (quantitative or qualitative) and the usage of domain-specific information. The information has been used to shape the characteristics of the experiment of this study. By combining the characteristics and experience of multiple experiments, this study will bridge the gap that is currently existing concerning the application of qualitative comparisons between plannings.

2.6. Answer to sub-question 1

The conclusion of the literature review is used to answer the sub-question: **What is the best method to research the qualitative difference between plans for complex planning assignments like railway node planning?** The Hierarchical Production Planning method describes the most common way of solving complex planning problems. From a mathematical perspective, the conclusion can be drawn that the current developments make it possible to create PATs that are able to solve node planning processes. The complex process is divided into sub-problems that can all be solved. Where earlier TUSP solvers solved the sub-problems sequentially, HIP is solving the sub-problems parallel. Since HIP is still being developed, the crew assignment sub-problem is not yet integrated into the solver. However, HIP is only tested in a testing environment that did not include the (possible) end-user: node planners. Background information concludes that there are many difficulties in implementing PAT in real-life planning systems.

In order to increase the cooperation between HIP and node planners, humans need to understand, interpret and trust the solutions proposed by HIP. HIP can solve the planning problem mathematically, but a lot of (human) aspects are not taken into account. These aspects must be integrated into HIP to ease the understanding of the solution of HIP by planners. In order to identify aspects that are undervalued, overvalued or missing, an empirical experiment is needed that focuses on qualitative data. Many experiments provide quantitative results of comparisons between plannings, but experiments that are focused on qualitative differences between plannings are limited. Besides, the data used is generalized or superficial, resulting in limited results that are not case dependent. Based on the experience of executed comparison experiments, it is clear that a mixture between a case-study and a field study is necessary in order to collect the data needed to create an adapted version of HIP that includes more information about aspects.

3

Current and future planning process

From the introduction (Chapter 1), it is clear that we are dealing with the planning of a logistical problem that consists of the movement of train units on the node (station and SL) to apply services to the train units at designated tracks. From the literature review, it is found that a domain-specific experiment is needed to gather qualitative data that can be used to adjust the HIP tool. This means that it is essential to analyze the current situation at the NS (Section 3.1). The future planning process includes the integration of HIP within the current structure to assist planners. Section 3.2 describes the possibilities of how HIP can assist planners, what information is necessary to run HIP and information about the output of HIP. The Section also describes the inner-workings of HIP in order to understand how the planning problem is solved. The last Section of this Chapter describes the answer on two subquestions: **How is the integrated planning process of NS set up and what are the characteristics of the developed Planning Assistance Tool?**

3.1. Current planning process

The first part of this Chapter describes the current planning process of the NS. First, the organisation of the planning process is explained with the help of the Hierarchical Production Process (Section 3.1.1). It explains the hierarchical structure of the process and the availability of data over time. This part also introduces the different planning departments that are responsible for solving node plannings. The second Section (3.1.2) describes currently used interactive tools to support planning. The third Section (3.1.3) describes the planning environment of departments introduced in the first Section. Section 3.1.4 presents the relation and dependencies of the different departments. At last, differences between nodes are described by listing different characteristics of nodes in Section 3.1.5.

3.1.1. Hierarchical structure of the train operation planning process

Riezebos et al. (2011) state that node planning problems are a semi-structured problem. That is why the NS and many other RUs want to use a structured method to solve this complex problem. Timetable planning, in general, is hierarchical, and to continue this structure, the node planning process at NS is also of a hierarchical sort. According to the literature review in Chapter 2, HPP is often used for processes at which material complexity is low, resources are expensive and have to be highly utilized (Zijm, 2000). These characteristics fit the operation of trains, which have to be highly utilized, and the capacity of SLs of the NS, which is scarce. This Section defines the layout of the complete node planning process of NS as an HPP.

As Zijm (2000) stated, hierarchical plans are created with aggregated global values on the highest level, which will be supplemented by more specific values and details on lower levels. In order to understand the node planning, it is helpful to have an idea about the other planning parts. In Figure 3.1, the complete planning process of the NS is presented. The complete planning can be divided into four parts: network, rolling stock, crew and nodes. The planning processes of these parts can be divided into the strategic phase, the tactical phase and the operational phase. This research focuses on the operational phase of the node: the node planning (bottom right). Omitting the crew planning (HIP does not cover this part), the Node planning is dependent on the input from the network planning and the rolling stock planning. In other words: the timetable and the rolling stock planning are needed to create a node planning. Within the rail industry, the definition of the timetable is *"an overview listing the times at which rail traffic products (train and rolling-stock movements), such as arrivals, transits and departures at timetable points (passenger or transport stations) are expected to take place"* (ProRail, 2019) (Rail Net Europe, 2019). From a node perspective, the timetable contains information when and where (on which platform) a train has to start and will end its passenger moving service. From the moment a train ends its services, it is the goal for node planners to make sure train units undergo services and are prepared for next passenger moving services. Next to network and rolling stock plannings, node plannings are also dependent on strategic and tactical decisions for nodes. A strategical decision to build an extra switch influences node plannings, just as a tactical decision to change shunting rules.

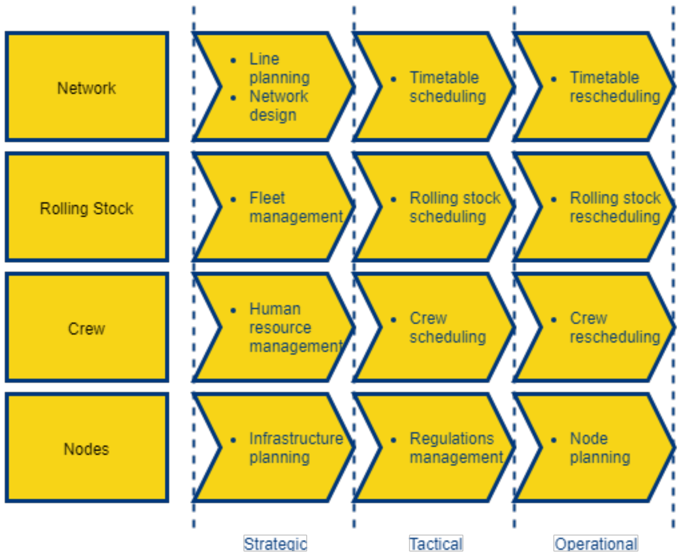

Figure 3.1: Overall planning process of the NS. The flexibility and adaptability of the planning process increases from the network to the node planning. It shows the operational node planning is dependent on a lot of decisions made in earlier phases and in other planning departments. Based on Goverde (2018).

Referring back to Zijm (2000), the node planning process (bottom right block of Figure 3.1) can also be divided in high and lower levels. Figure 3.2 presents the levels of node planning and the corresponding period of these levels. The big blocks represent the different planning phases (levels) of the node, while the smaller blocks represent the departments that are dealing with the planning problem during these phases. The node planning starts from **X - 1 year**, where X is the day of operation.

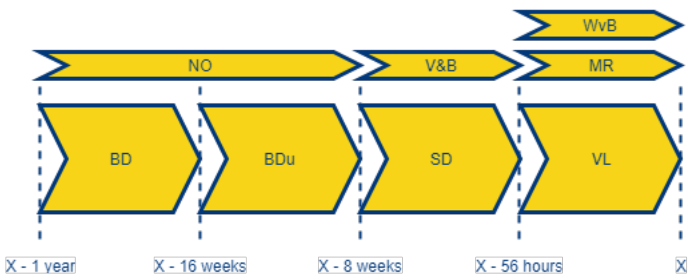

Figure 3.2: Timeline of the NS with actual deadlines

The **BD (Basic-Day) phase** comes with the first document that includes the output of the original timetable and is used as input for the node planning. This document also contains rolling stock planning. A standard planning day for a node planner starts at 7:00 in the morning until 7:00 the next morning. This time is chosen, because at 7:00 in the morning, almost all trains are in service, meaning that the amount of trains on the SL is low, making it an ideal transfer moment because the amount of information is limited. So based on the information of the timetable and material planning, a plan is created for the nodes by the Netwerk Ontwerp (NO, Network Design in English) department. This information consists of the trains ending a service and starting up a service (most probably the next day) and the type of rolling stock of these trains. The rolling stock planning makes sure the balance between the trains ending and the trains starting up from the node is right. The BD information also includes repetitive, or significant maintenance works planned by ProRail, the Infrastructure Manager.

The **BDu (Basic-Day update) phase** starts around 16 weeks before the day of operation. The updates characterize this phase to the Basic-Day schedule that is necessary to make the schedule more robust because of weather and demand changes over the year. Also, extra maintenance works can be added, or planned maintenance works can be removed. Famous examples are the higher demand of travellers during September, which results in longer boarding times, and leaves (autumn) that causes a slippery layer on the tracks which makes it harder for trains to accelerate and brake. The updates to the Basic-Day order are marginal and should not result in substantial differences in node plannings unless significant infrastructure works are planned within this BDu. It is the NO department that performs updates to the planning based on this newly available information.

The next phase is the **SD (Specific-Day) phase**. From 8 weeks before the day of operation, the BDu planning is being adapted to a SD planning to specify a general planning to a planning for a specific day. For big concerts or football games, extra capacity is necessary to transport all passengers. These activities might cause an overload or reduction of trains that have to be serviced and parked. Also, an updated maintenance schedule might results in different infrastructural restrictions compared to the BDu schedule. The changes that are applied during this phase can be relatively significant, and are performed by the Voorbereiden & Bijstellen (V&B, Prepare & Adjust department) department.

From 56 hours before operation, the **VL (Verkeers Leiding / Traffic Management) phase** is active. Before this phase, the plan of NS should be finished and should not contain any problems or conflicts. ProRail is checking if the plan is feasible. From this moment, small and unplanned disturbances can lead to changes in the planning. Dispatchers make these changes. Employees of these departments are called Medewerker Rangeren (MR, Shunt Traffic Manager in English) and Werkvoorbereider (WvB, Dispatcher in English). The difference between these two departments is based on the area they hold responsibility. The MR is responsible for the station and routes toward the SL, while the WvB is responsible for the SL. These responsibilities are not standard for nodes within the country but can differ. More information about this will be provided in the node layout analysis (Section 3.1.5).

3.1.2. Current planning tools

Before characteristics per department are presented, it is necessary to explain the existing tools that are used by planning departments. Planners can use three planning tools. The first one is the most important: Donna. Donna is an application created by ProRail in which all rail operators plan their activities. This tool provides essential information for the planner, such as the material type and information, timetable information, earlier plans created for the node, etc.. The planner plans activities based on a track occupancy graph. A visual of a Track Occupancy Graph (TOG) is provided in Figure 3.3. On the x-axis it shows the time, on the y-axis, the difference tracks are plotted. A white dotted line represents a parked train on the track. An arrow represents a movement on the track. Every parked train starts and ends its line by a small arrow, representing the movement of this train in order to park. An example from Figure 3.3: track 10W is occupied by a first train from 7:02 till 7:14 and by a second train from 10:19 till 10:27.

Next to the visualization of a plan, Donna also helps planners by routing. Planners have to add activities by selecting a train and providing a start track and an end track and a time of departure. Donna picks the shortest route over the network but also gives the planner an option to pick other possible routes. An overview of the node and the selected route will be shown to the planner (Figure 3.4). The routing assistance of Donna helps the planner since it does not have to search for routes itself. The last important function of Donna is to detect conflicts and rate them. This is a crucial part for planners since most planners are replanning based on conflicts. Donna can recognize conflicts like crossings of routes at the same time and lengths of trains exceeding the shunting track.

The second tool is the Opstel Plan Generator (OPG). A small part of the planners only uses this tool because

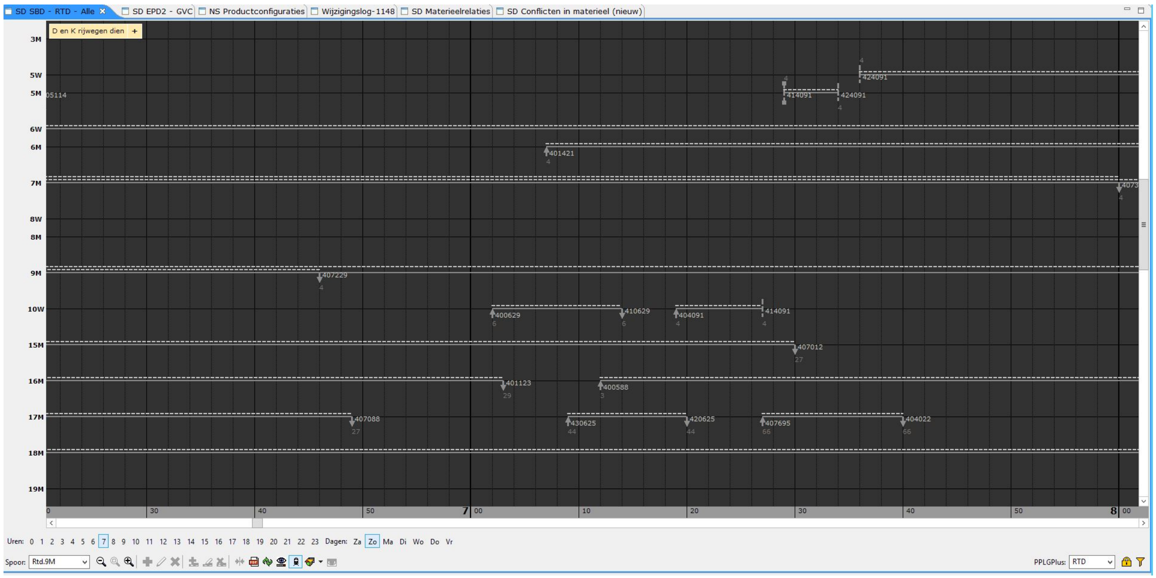

Figure 3.3: Track Occupancy Graph (TOG) of part of the tracks of Rotterdam. An ongoing white line corresponds to the fact that the track is occupied, the white dotted line corresponds to the different trains that occupy the track. The arrows resemble a movement on the track.

it only covers a small amount of the nodes. This tool provides the planner with a proposition where to park which train at a SL. The method behind this tool can be found in Section 2.3.

Node planners are responsible for the schedules of train drivers that drive empty trains between the station and the Service Location. The Personnel Scheduler (PS) provides solutions for this planning problem. This tool is connected to Donna and needs planned activities in order to create a schedule for the personnel. The planner enters the number of train drivers available, at which the PS first generates tasks based on the planned activities and assigns these tasks to personnel. The tool is used by planners to generate an initial crew plan, but since the tool might come up with infeasible solutions because it does not take into account certain factors (e.g. equal workload), the initial planning has to be checked and improved by hand.

3.1.3. Plan environment per department

This Section describes the planning environment of the four departments depicted in Figure 3.2. The planning environment can be described by the plan assignment, the input needed to solve the assignment, the restrictions that are applied and the possible action a planner can take. The goal of this Section is to describe the solution spaces of the four departments. Interviews with employees of different departments gather the information described in this Section. The employees have approved the interviews in order to check if the information was accurate. A summary of the information described in the Section below can be found in Table 3.1.

First, extra information about terms used in the structure is provided:

- **The planning assignment** describes the goal(s) of the planners and the conditions under which the planners must solve the plan problem.
- **Input** describes the (new) input available that needs to be used when solving the planning problem.
- **Restrictions** describes rules, norms, regulations and other important factors that have to be taken into account when solving the planning problem.
- **Possible actions** describes the list of actions a planner can take in order to solve the planning problem. The actions described can be taken by the planner to enlarge the solution space.
- **Solution space** summarizes the information described above in order to visualize the area in which planners need to find a solution.

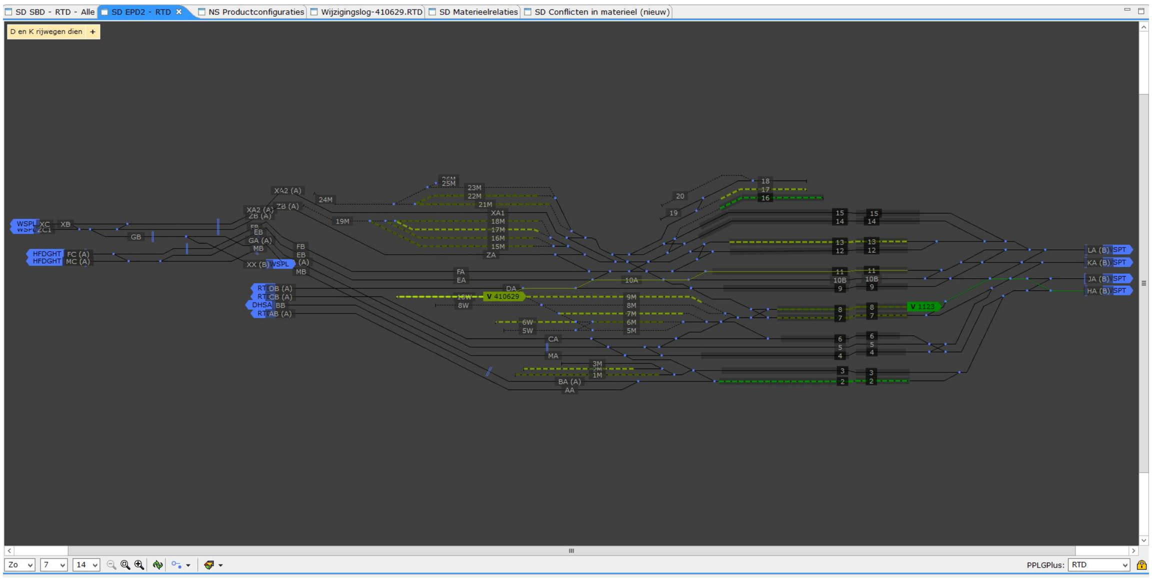

Figure 3.4: Overview of an action of a train over the emplacement of Rotterdam. The small green lines represent the movements over the tracks. In this case, two trains are moving on the emplacement (one on the left side of the station, and the other one on the right side). The green blocks represent the tracks that are claimed by NS, meaning that only NS is allowed to use these tracks for parking and servicing.

Netwerk Ontwerp (NO)

Plan assignment

The planning assignment for the NO department is to create an executable plan for the node within the BD and BDu phase of the planning process. There should be an executable plan to move rolling-stock from the platform to the SL and back at the requested times. Next to that, an executable planning for the train drivers (personnel) is also expected. The goal is to make the plan as efficient as possible. For NO, an efficient plan is to minimize the amount of personnel necessary, while making sure every train is being transferred to the SL in order not to disturb the ongoing services. If services of the SL are integrated into the planning process (see Section 3.1.5 for more information) it is also crucial that all the trains are offered for service on predetermined tracks.

Input for planning assignment

The input of the planning process of the NO planners is the timetable and material planning. Also, a list of long-term or periodic maintenance of tracks is provided as well.

Restrictions

First of all, the hardest restrictions are infrastructure restrictions and safety restrictions. It is not possible to drive somewhere if there are no tracks. Safety is always essential. Within the hierarchical system, it is not possible to change deliberate choices made by departments at an earlier phase, except when nothing else is possible. This means that the timetable also provides restrictions, like routes blocking the accessibility of the SL and departure times of trains from the station. Next to that, the type and amount of rolling-stock and compositions of the trains are set by the rolling stock department. There is also information from third parties, like maintenance of tracks minimizing the infrastructure availability. The last restrictions are restrictions that are set by the departments on a lower level than the NO level. These restrictions are described as shunting norms (for example, the amount of time between the arrival of different times at a SL). In general, these restrictions are more flexible, as long as clear explanations are provided.

Possible actions

A NO planner can perform many actions to solve the problem he or she encounters. First of all, the NO planner can decide the routes of ongoing trains through nodes, but also the routes of shunting trains (if multiple routes are possible). Next to that, the NO planner can determine the arriving platform of trains that end a service at a station and (less preferred) the departing platform of trains that start a service from the station.

Table 3.1: Characteristics like tool usage, possible actions and restrictions per planning department: Netwerk Ontwerp (NO, Network Design), Voorbereiden & Bijstellen (V&B, Prepare & Adjust department), Medewerker Rangeren (MR, shunt traffic manager) and Werkvoorbereider (WvB, Dispatcher). For a better understanding of (x) see *planning responsibility* in Section 3.1.5.

Section	What	NO	V&B	MR	WvB
Possible Actions	Change of route through a node of ongoing trains of the NS	x	x	x	
Possible Actions	Change of arrival or departure platform	x	x	x	
Possible Actions	Change / specify location of train on platform / shunting track	x	x	x	
Possible Actions	Change of start times of shunting movement / actions	x	x	x	
Possible Actions	Split and combine train units	x	x		
Possible Actions	Change of arrival time of trains (within certain norm)	x			
Possible Actions	Request of train units from SL			x	
Possible Actions	Remove actions			x	
Possible Actions	Determine to not accept trains in SL			x	
Possible Actions	Allocate shunting track	(x)	(x)		x
Possible Actions	Allocate services				x
Possible Actions	Recompose trains for departure	(x)	(x)		x
Possible Actions	Park and service extra trains				x
Restrictions	Infrastructural restrictions	x	x	x	x
Restrictions	Safety restrictions	x	x	x	x
Restrictions	Walking time norms	x	x	x	x
Restrictions	Norms of ProRail	x	x	x	x
Restrictions	Movements of external rail operators	x	x	x	x
Restrictions	Out of service restrictions	x	x	x	x
Restrictions	Station bound norms (Stationsboek)	x	x	x	
Restrictions	Services determined by the timetable	x	x	x	
Restrictions	Shunting norms (Catalogus)	x	x		
Restrictions	Rolling stock determined by the rolling stock plan	x	x		
Restrictions	Departure times of trains	x	x		
Restrictions	Change of composition and type of trains	x	x		
Restrictions	Employees			x	x
Restrictions	Request of train units from SL				x
Tools	Opstel Plan Generator (OPG)	x	x		x
Tools	Donna	x	x		
Tools	Personel Scheduler (PS)	x	x		

This is sometimes necessary because the SL is not accessible from all the platforms or routing problems. Next to that, the NO planner determines when a shunting movement takes place. A NO planner can decide to split and combine trains at platform tracks (or at the SL) in order to minimize movements over the emplacement. These actions are all easy to implement and do not need extra explanation. For some more specific actions, planners need to provide an explanation or add approval codes, but these are all node-specific. Examples of these codes are ROBERTO calculations, which provide calculations of travel times between two points on the node. There will be no focus on these points since they can be seen as exceptions and are strongly node-specific.

Solution space

The NO department starts with the creation of the node planning and therefore starts with an 'empty sheet'. Although design possibilities are pretty free, there are restrictions. These restrictions are mostly focused on norms and physical restrictions. There are also resource restrictions, but normally there should not be an overstretch of capacity since this is being thought of by the rolling-stock department. With the conflict checker, a planner can check if there are problems, but unfortunately, the solution space is not fully covered by Donna. Manual checks are necessary in order to find out if the solution is 100% feasible. If exceptions are needed to solve the planning, there is an option to by providing information or a validation code.

Voorbereiden & Bijstellen (V&B)

Plan assignment

The planning assignment for the V&B department is to create an executable plan that is adapted because of the addition of new information from the SD phase. This information is implemented in Donna, and therefore the goal of the planner is to remove conflicts that are detected in Donna. However, during the period of the SD phase, new events or necessary maintenance to tracks can be added. In case this happens, the planner will always be involved if the changes are going to be big. The planner will check if the rolling-stock planning is all right, but also whether the planning for the personnel is feasible. If the plan is conflict-free and feasible, the planner will improve the plan. First part will be to create an efficient plan, meaning to reduce the number of movements and to reduce the number of tasks for personnel in order to decrease the amount of crew necessary. Next to that, the planner will focus on an equal division of the tasks over the personnel (since PS does not take care of this) and to add extra information (e.g. exact parking spots for personnel) in order to improve the operation. If services of the SL are integrated into the planning process (see Section 3.1.5 for more information) it is also crucial that all the trains are offered for service on predetermined tracks.

Input for plan assignment
The input for the V&B planners is the plan created by the NO department, SD information and extra information like new maintenance updates and changes from third party rail users.

Restrictions
The restrictions of the V&B planners resembles the restrictions of the NO planners, as can be seen in Table 3.1. However, there are some differences, since the planning is a little bit more dynamic. ProRail provides better data about the maintenance works, which means that some maintenance work will be removed or different rules apply at the maintenance site (sometimes it is possible to use the tracks while there is maintenance). Third parties are shifting their services. Certain freight paths might be removed because they will not be used. However, all these changes might happen within these four weeks. It is therefore of great importance to notice these changes all the time.

Possible actions
The actions of the V&B planner are the same as the NO planner. For the V&B planner, there is an option to handle individual exceptions when affirmed by calculations. These exceptions are not preferred, so they are often the last option to pick from (if no other action is possible to execute).

Solution space
The solution space of the V&B department is quite the same as the solution space of the NO department except the fact that the deadline is closer and more information is available. The higher amount of information can open new possibilities on one side, though on the other side it might also increase the capacity on the node because of significant events that were not expected in the long term. Next to that, there is the interference of ProRail warning the planner in case conflicts are still in the planning when the deadline approaches.

Medewerker Rangeren (MR)

Plan assignment
The goal of the MR is to execute a provided plan. If no disturbances occur, this should be the case. However, chances that disturbance occur are high in a network that is highly utilized. Therefore the second assignment is to solve last-minute disturbances. These disturbances might cause further problems that harm the complete planning. In that case, it is for the MR and the rest of the staff (for the ongoing lines etc.) to make sure that everything gets back to plan as soon as possible. The norms used in Donna are higher compared to the norms used in the real world. Because of this, many disturbances can be solved easily because extra time is available in the real world. If everything goes according to plan, the dispatcher has time to look forward in the planning to find mistakes or dangerous situations where, because of real-life information, improvements are possible. The MR department is the first department that has a clear KPI, which is the fact that all the trains that should start a service at the node are in the right composition (and of the right material).

Input for the assignment
The input for the MR are parts of the plan of V&B. The track occupancy graph, together with the activity plan are the only things the MR needs. The MR receives real-time information from ProRail, who is managing the network. Based on this information, the MR can predict the future a little bit and can adjust the planning in order to cope with this expected future.

Restrictions
The standard restrictions of the MR are the same as NO and V&B such as restrictions because of infrastructure and norms of ProRail. However, there are not many more restrictions, since the dispatcher is operating the real-time. There might be real-time failures or problems due to personnel that calls in sick or does not call in at all.

Possible actions
The actions that a MR can take up any form, as long as norms of ProRail and NS are taken into account. There are different levels of preferred actions, but when necessary, everything can be done that is allowed in order to continue the service of moving passengers. The actions of changing routes of ongoing trains and changing arriving or departing platforms are not preferred, since this means that information needs to be changed for passengers. Moving start times of shunting is possible, as long as it is permitted by ProRail and personnel is available. The action to request specific train units from the SL (even if they are not clean) is needed if rolling stock disturbances are causing problems to a train service, meaning that trains do not arrive. This action will always come from higher level and is an order for the MR (as well as for the WvB). The removal of actions is allowed if trains do not arrive and if these do not disturb the ongoing lines. It might be possible that specific actions are the only actions that can be executed, even though they harm ongoing services over a node. Next to this, the MR can decide not to accept train units (units that are decoupled), because no personnel or space is available at the SL or no possible route is available, and the unit will block a platform for an ongoing train. The reaction of this action is that a train needs to continue its service with more train units than needed, causing higher operation costs.

Solution space
When there are no disturbances, there is no need for a solution space for the Medewerker Rangeren. Small disturbances can be solved easily because of the high norms that are used by the prior planners, and the lower norms that are used in real life. For bigger disturbances, there is a threat that they will cause other disturbances. In order to not cause a domino effect, the solution space of the Medewerker Rangeren is big since there are not that many additional restrictions next to the standard restrictions. The MR also has to possibility to request new trains from the SL and to park extra trains at the SL (if possible). Only if the capacity of the SL is reached this might become a problem, but chances are small since SL are empty during the day. Next to that, employees are a restriction that might hinder the MR in its actions.

Werkvoorbereider (WvB)

Plan assignment
The assignment of the Werkvoorbereider is to execute the services as planned. A secondary task for the Werkvoorbereider is to plan the next day at the SL. This short amount of time between planning and execution is possible because ProRail does not need to check the feasibility 56 hours before execution. In the chain, the SL is extremely reliable on the Medewerker Rangeren, since he or she decides (and operates) the planned availability of trains at the entrance of the SL, just as he or she expects clean and checked trains at a predetermined time at the exit of the SL. In the time between these two points, the Werkvoorbereider has the time to clean, wash and check the train at the SL. Because of this, the KPI of the Werkvoorbereider is measured by the number of dirty starting trains leaving the SL. The smaller this amount, the better the performance of the SL. Next to this KPI, planners will also try to look for mistakes in the planning to solve these on beforehand, and the fact to improve a planning by minimizing the number of activities.

Input for plan assignment
Since the Werkvoorbereider executes and plans, a division is made to describe the input for the WvB necessary for both tasks. To create a planning, a WvB receives an activity plan from the V&B department that describes which trains arrive at the SL at what time, and which trains need to leave the SL at what time. Next to this, there is also a proposed matching schedule, stating which train unit (coming from one train) is becoming which outgoing train. With this information, the WvB can plan which train will be parked at which track, and when and how services will be applied to the train. For the execution task, the primary input is the plan from the WvB of an earlier shift. However, since SLs are often NCBG, which means that changes can be made very quickly, it is often the case that a Werkvoorbereider only uses part of the planning. Reasons for this choice is because the WvB has its preference or does not understand why specific actions are implemented in the plan or improve the plan according to him. The SL absorbs disturbances in the network because of

its flexibility. This often results in a deviation of the planning to ensure that transport services can continue according to plan.

Restrictions
The standard restriction applies for the WvB like infrastructural restrictions, safety restrictions, walking time norms, norms of ProRail and movements of external rail operators, in case of shared SL. Next to that, there are out of service restrictions (also on SL, tracks need repairs and maintenance). There is a certain amount of employees available to execute tasks, and not all employees are able to perform all the tasks. The last point has to do with unexpected orders to deliver a train immediately. The WvB does always have to respond to this appeal unless no train is present at the SL. The critical restriction of a WvB is the capacity of a SL, which often consist of a service capacity and a parking capacity. These numbers are determined once every year and are based on different factors, like shunting tracks and tracks allocated for services. Exceeding this capacity results in many difficulties for the WvB, since the freedom of movement is reduced, which means that it is hard to recompose trains and to wash trains (extra tracks are needed for these actions). Also, when the capacity is reached, trains are placed on tracks that are reserved for certain services, meaning that these services cannot be executed any more.

Possible Actions
The possible actions of a WvB are the allocation of shunting tracks to trains, allocating services to trains (except cleaning, since a third party performs this), recomposition of trains for departure (this means that shunting space is needed to move trains). The WvB might need to park and service extra trains if needed. This means that the WvB should always know what the usage of the possible capacity is and if it is possible to increase the capacity (if needed) by more efficient parking of trains. Knowing that the SL is the flexible part of of a node (and indirectly, of a complete timetable), its usage might be crucial in order to minimize the impact of disturbances to the complete network. A simple example of this case is if a particular train line is disrupted, it is not wanted that all the trains on that line stop and park at the station since this can result in congestion around the station meaning that the disturbance hinders other lines.

Solution space
The solution space of the WvB is flexible since the planning is done on a short term, and there is a need to have flexibility on the SL in order to solve disturbances. The planning task of a SL alone is quite simple since there are not that many restrictions. Capacity is the key restrictions that a WvB has. Because of this, the WvB can easily replan the planning, even during execution of the plan.

3.1.4. Relations between node planning departments

The Integrated DEFinition for Function model (IDEF0-model) is a tool that is developed to describe a whole system and define its functional scope (Grover and Kettinger, 1999). It is often used for analyzing and designing complex systems. By developing an IDEF0-model, a standard and understandable representation of a system can be utilized as a common language and point of reference. Using an understandable model is especially useful in systems in which information sharing is essential. The IDEF0-tool consists of several system levels, varying from A0 to A4. The A0 chart is the most high level. The IDEF0 diagram consists of relations, tasks, blocks and arrows in which the system behaviour can be described. With an increasing level number (going up from A0 to A4), the system is analyzed at lower levels in more detail. This Section analyses the node planning process using the IDEF0 method. In order to explain the working of the IDEF0 method, an A0 diagram has analyzed the complete planning process. An A1 level diagram is presented afterwards. Digging into deeper layers is not necessary, because an integrated approach is needed, which can be completely explained on the A1 level of the IDEF0.

Figure 3.5 shows the A0 representation of the node planning process. The Figure is used to explain the use of an IDEF0 diagram. The process receives (and needs) information, which is categorized by the arrows entering the block on the left side. For the node planning process, this is the BD, BDu and SD information, feedback information and other information. This information will be specified on A1 level. From the top, the controls are presented: Node information, KPI's and objectives that describe the goals and the constraints of the node planning process. The arrows from underneath represent the tools and machines that are being used during node planning. On the right side, the products of the node planning process are presented, which are the execution of the node process and feedback.

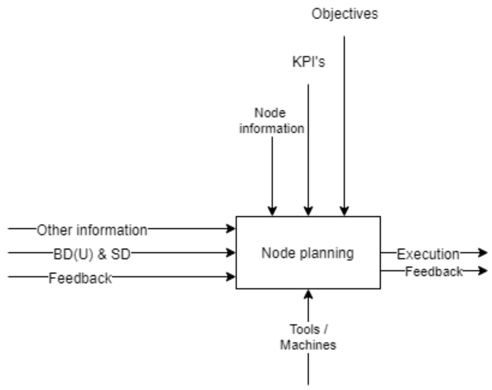

Figure 3.5: Node planning process on IDEF A0 level

The A1 level of the IDEF0 diagram is presented in Figure 3.6. It is crucial to describe the definition of arrows within the IDEF0 diagram. The KPIs and objectives differ per planning department. The KPIs are measurable and set for evaluation reports, while the objectives are less strict and can have different values depending on the planner. They do not have to be measurable. Node information contains all the information and rules of a node, meaning that it includes the infrastructural constraints and safety constraints, but also information about walking times and other essential information. Some departments have as output specified constraints for departments that adapt the planning at a next level. These constraints are the results of decisions made within a precise moment before the execution of a plan. The best example is the amount of personnel that is available on a node. At a certain point, a decision is made about the number of train drivers, mechanics and cleaning staff. This amount is set from that moment on and is hard to change in the proceeding planning phases. From the tools and machines part of a block, Donna appears quite often. For more information about Donna, see the Section above. The last arrows that need to be discussed are the input side of a block. the BD(u) and SD stand for all the information that is related to the different planning phases from the timetable planners. The term other information relates to information about the infrastructure (decommissioning of individual tracks for maintenance) which is more important for the long term planning, and real-time information that is important for short term planning.

Now that the arrows are explained, a description of the process can be explained that is presented in Figure 3.6. Based on the BD and BDu information from the timetable planner, the NO department creates a node plan in Donna. This plan (which can still contain mistakes and conflicts) is forwarded to V&B department. The V&B planners add the SD information to the planning and start solving the conflicts in the plan in order to have a feasible planning. This feasible planning is forwarded 56 hours before execution to the Medewerker Rangeren, who executes the planning and, if time is available, checks the planning on real-time information. If necessary, the MR adapts the planning and executes the new plan.

For the V&B planner, the plan from the NO planner and the information gathered from the BD(U) and other information is used to create a plan that describes the matching aspect. Minimizing the amount of splitting and combining trains is its objective. This plan is forwarded to the Werkvoorbereider. Together real-time information, the WvB to executes the planning.

3.1.5. Differences between nodes

Node plannings are highly dependent on the characteristics of the nodes. The characteristics are based on the layout of the node, but also on other factors. These characteristics have a significant influence on the planning assignment as they determine essential parameters.

The following categories are set in order to categorize different nodes:

- **Shuffleboard versus carousel:** these are the two general layouts types of SL. The shuffleboard layout can be recognized by the fact that trains are parked on a track where almost all the services can be executed (mechanical checks, interior cleaning and minor repairs). Limited train movements are nec-

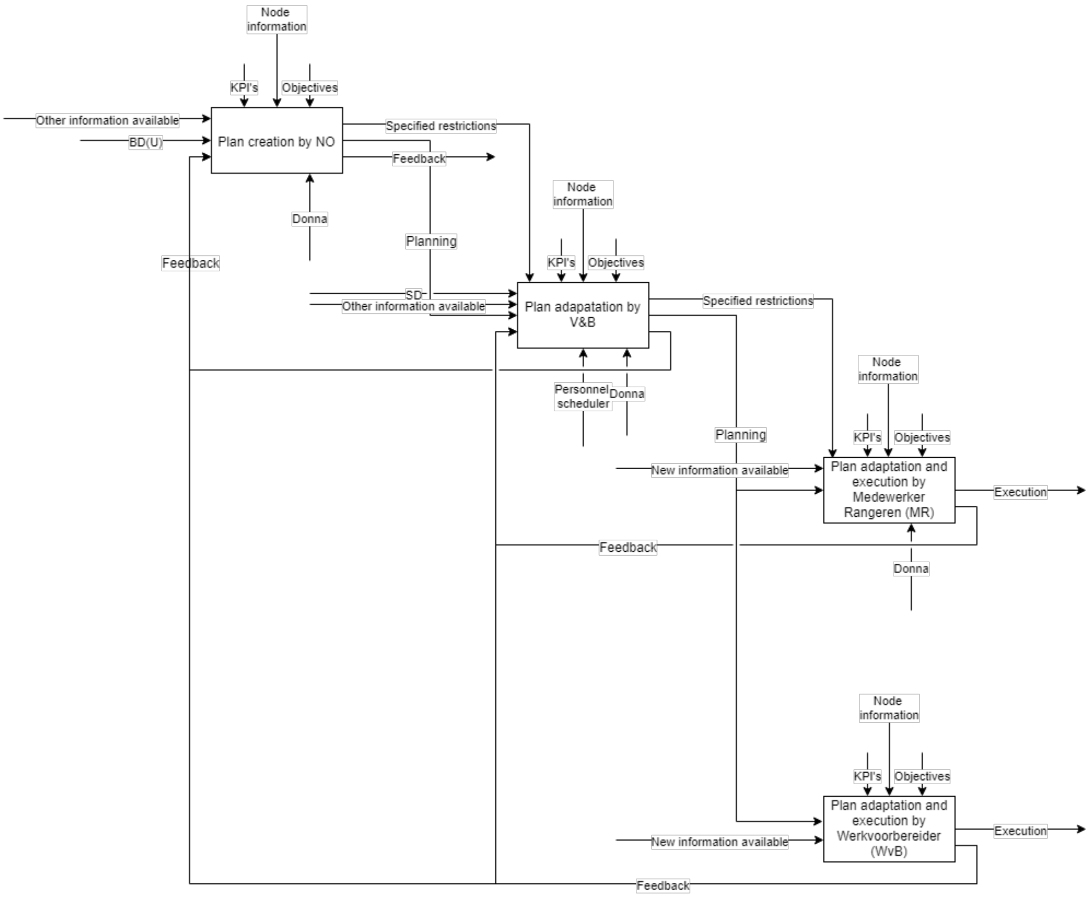

Figure 3.6: Node planning process on IDEF A1 level.

essary. The idea of this layout is that the personnel comes to the train to perform a service (job-shop). A carousel SL can be characterized by the fact that the train moves around the yard to be serviced. The idea behind this layout is that trains move to the personnel in order to be serviced (flow-shop). Clear layouts of a shuffleboard and a carousel SL can be found in Figures 3.7 & 3.8 respectively. In practice, most locations are a mix of both layouts.

- **Concentrated versus scattered SL:** the term Service Location does not mean that all the services (including parking) have to be applied at the same geographical location at the node. Some SLs are scattered over a yard, while others are neatly concentrated. A good example is the SLs of the nodes of Lelystad compared to Rotterdam. Figure 3.9 shows a SL that is scattered over the emplacement, while Figure 3.10 shows a concentrated SL.

- **Accessibility of SL:** the degree of accessibility of a SL differs. Some SLs are accessible by multiple tracks and from multiple directions, while others are only accessible by one single track. Also, the presence of planned movements through the node can influence the accessibility of the SL, because crossing these planned movements is not possible. At last, it also depends on infrastructure if it is possible to reach the SL from every platform. Figure 3.9 shows that it is not possible to reach every platform from every track in the SL and vice versa in one movement. An example: from the tracks selected by the red box in the bottom of the Figure, it is impossible to reach the platforms that are presented in the top of the Figure within one movement. This problem is something that planners need to think about. Departure and arrival platforms can be changed, but this is not wished because passengers are used to the fact that a train always departs or arrives from a certain track. Extra information to passengers is necessary if this action is chosen.

- **Planning responsibility:** There are two types of planning environment that are applicable within NS. There is a segregated planning environment, in which a planning is made for the node and a seperate planning is created for the SL. This is often applied at places where the SL is concentrated and of NCBG (Niet Centraal Bediend Gebied, Not Centrally Operated Area in English), like Utrecht Carthesiusweg (Figure 3.7). The demarcation line between the two responsibilities is at the S-sign, which is the entry or exit of a SL and the transition between CBG (Centraal Bediend Gebied, Centrally Operated Area in English) and NCBG. The NO, V&B and MR are responsible for the planning of the supply and removal of trains at the SL, and the WvB is responsible for the service allocation and transport of trains at the SL. However, there are also integrated planning environments, like Rotterdam (Figure 3.9). At these locations, an integral planning is created that includes the transportation of trains to the SL and the assignment of tracks to trains. In other words: the NO, V&B and MR are responsible for the transport of trains towards the SL, the track allocation and (depending on the SL layout) the need of extra movements because of specific service allocations. The NO, V&B and MR planner must offer the train at certain service tracks (like cleaning platforms) for a predetermined time so that the WvB is able to assign services to that train. These periods are called takt times and should resemble the amount of time that is necessary to execute a service with the addition of some buffer time. In an integrated environment, the NO, V&B and MR planner focuses on the complete transportation of trains over the emplacement, and the WvB assigns teams to execute services during the night. This means that the planning assignment of the NO, V&B and MR planner is bigger compared to the segregated environment, and the planning assignment of the WvB is decreased since it does not include the transportation problem of trains. Within Table 3.1, the *(x)* present the extra tasks for the NO and V&B planners.

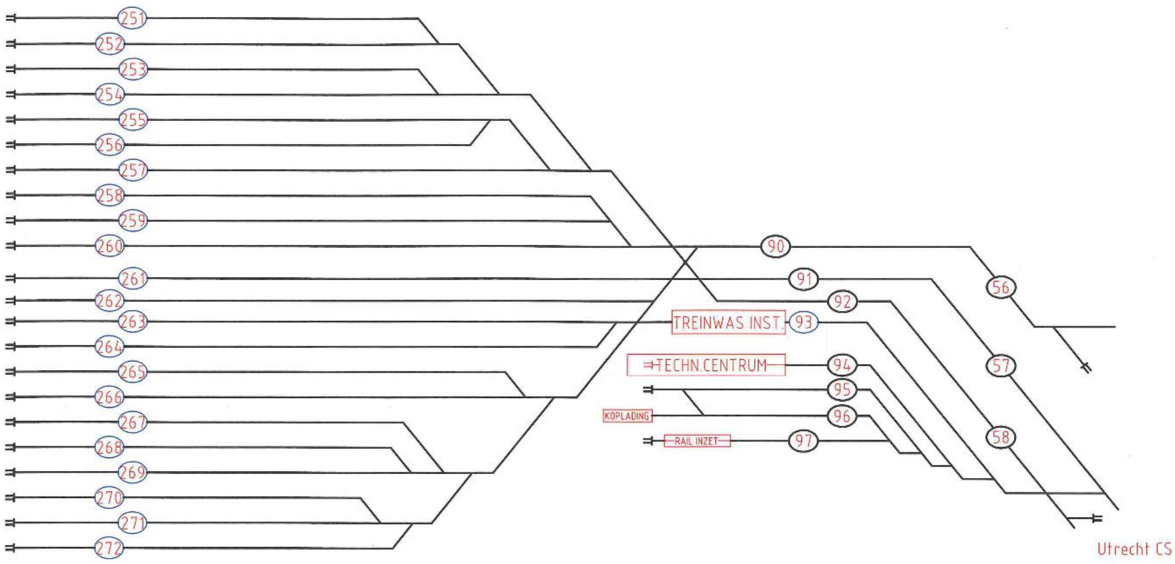

Figure 3.7: Example of a shuffleboard layout of a service location (Carthesiusweg (Ctw) in Utrecht)

Figure 3.11 is a complexity matrix based by van Uden et al. (2017) of all the NS nodes in the Netherlands. The nodes are ranked a 0 or a 1 based on different characteristics: layout and rolling-stock. The higher a node scores, the more complex the node is to plan. The characteristics that are scored are the following:

- Servicing trains is not allowed on all parking tracks
- A TWI is present at the SL
- Many movements are necessary to reach the repairment stations
- The SL is also used for large maintenance of trains, hindering the standard SL services.
- Restrictions of infrastructure at the SL. Think about long term maintenance to tracks, safety issues (use only allowed during daylight) and restriction due to the neighbourhood (not available during certain hours in the night)
- Many movements because of the layout of SL
- It is possible to enter and exit every track from two sides (decreasing the complexity)
- Capacity problems at the SL (both service as parking)
- A large mix of different rolling-stock types

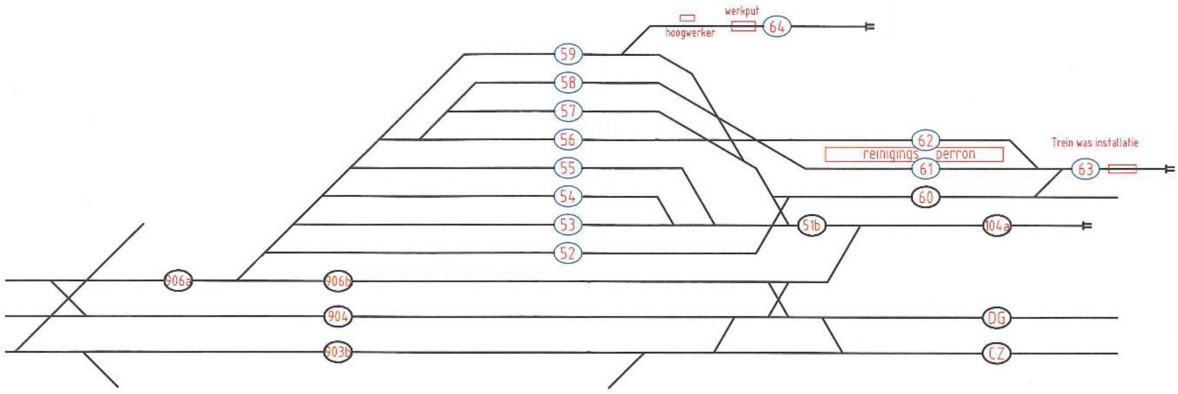

Figure 3.8: Example of a carousel layout of a service location (Kleine Binckhors (Bkh) in The Hague)

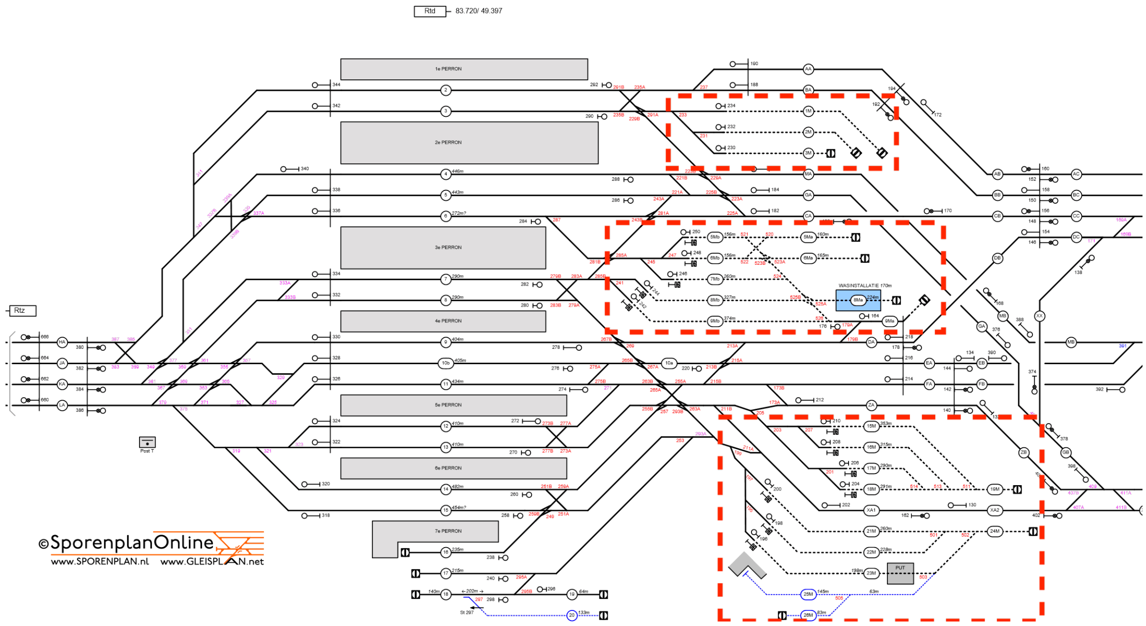

Figure 3.9: Node overview of Rotterdam, where the service location (marked by the red boxes) is scattered over the emplacement (Sporenplan Online, 2019)

- A large percentage of trains that need to be split and combined for their next transporting service
- A large percentage of adjustment necessary during operation

The author of this matrix concluded that the more complex the node is, the better it is to have an integrated planning system instead of a segregated planning system (van Uden et al., 2017). From the nodes with high values Zwolle, Rotterdam and Eindhoven have changed to an integrated planning system. For the other nodes, the planning is segregated or partially segregated (there is no clear line). Even between the three nodes that apply integrated planning, the terms of integration differ. For Rotterdam and Eindhoven, the TWI (washing) is not included in the process, while for the node Zwolle, it is. The level of integration is also node-specific.

3.2. Description of the Hybrid Integral Planning method

HIP is a PAT that is being created by the team of R&D Node Logistics at NS. HIP is used for the experiment of this study because the chances are big that HIP will be implemented in a later phase (Den Ouden, 2019). In order to fully understand the characteristics of HIP, it is necessary to describe this tool. First of all, the general set-up of HIP will be explained (Section 3.2.1) by introducing the sub-tasks of the TUSP that HIP solves and presenting the general operation of HIP. Sections 3.2.2 and 3.2.3 contain information about the input and

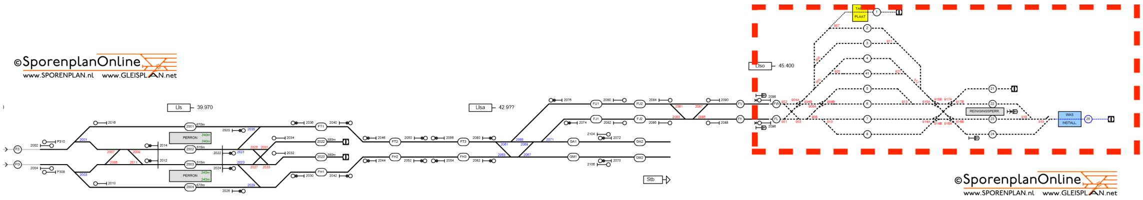

Figure 3.10: Node overview of Lelystad, with a concentrated SL (red box) (Sporenplan Online, 2019)

LOCATIES			Complexiteit door lay-out emplacement								Complexiteit door materieelaanbod					
			Niet alle opstelsporen zijn behandelsporen	TWI aanwezigop het emplacement	Veel bewegingen voor OB/TC?	Aanwezigheid onderhoudsroosters	Aanwezigheid tuin met slot	Emplacement beperkt toegankelijk	Veel rangeren (gevolg van type emplacement)	Vrijwel alle sporen zijn doorrijsporen	Hoge huidige benuttingsgraad	Grote mix aangeboden materieeltypen	Hoog percentage opnieuw samenstellen	Meerdere vervoerders op het emplacement	Hoog % bijsturing tov geplande omloop	
Regio	Naam	Wegingsfactor >>>	1	1	1	1	1	1	1	-1	1	1	1	1	1	Score
RRN	Alkmaar	Amr	1	0	0	1	0	0	1	0	1	1	1	0	0	6
RRN	Den Helder	Hdr	1	0	0	1	0	0	1	0	1	0	0	0	0	4
RRN	Enkhuizen	Ekz	1	0	0	1	0	0	1	0	1	1	1	0	1	7
RRN	Haarlem	Hlm	1	0	0	1	0	0	1	0	1	1	0	0	0	5
RRN	Hoorn	Hn	1	0	0	1	0	0	1	0	1	1	1	0	1	7
RRN	Lelystad Opstel	Llso	1	1	0	1	0	1	1	0	1	0	1	0	1	8
RRN	Amsterdam Zaanstraat	Zstr / Aswp	1	0	1	1	0	1	1	0	1	1	0	1	0	8
RRN	Amsterdam W'meer	Wgm	1	1	1	1	0	1	1	0	1	1	1	1	1	11
RRN	Hoofddorp Opstel	Hfdo	1	0	0	1	0	1	1	0	1	1	1	0	1	8
RRZ	Grote Bkh		1	1	1	1	0	1	1	0	0	1	1	0	1	9
RRZ	Kleine Bkh		1	1	0	0	0	0	1	0	0	1	1	0	1	6
RRZ	Dordrecht	Ddr	1	0	0	1	0	0	1	0	1	0	0	1	1	6
RRZ	Rotterdam		1	1	0	1	1	0	1	0	0	1	0	1	1	8
RRZ	Utrecht OZ	Ut OZ	1	0	0	1	0	1	1	0	1	1	1	0	1	8
RRZ	Utrecht Cartesiusweg	Ut CTW	0	1	1	1	0	1	1	0	1	1	1	0	1	9
RRZ	Utrecht Landstraat	UT LS	0	0	0	1	0	0	1	0	1	0	0	0	0	3
RRZ	Leidschendam	Ldd	0	1	1	0	0	0	1	0	0	1	1	1	1	7
RRZ	Amersfoort Bokkeduinen	Bkd	1	0	0	1	0	0	1	0	0	1	0	1	0	5
NO	Arnhem	Ah	1	1	1	1	0	1	1	1	0	1	0	1	0	7
NO	Arnhem-Berg		1	0	0	1	0	1	1	0	0	0	0	1	0	5
NO	Enschede	Es	1	1	0	1	0	0	0	0	1	0	0	1	0	5
NO	Hengelo	Hgl	1	0	1	1	0	0	0	1	0	1	0	1	0	4
NO	Nijmegen Rep		0	1	0	1	0	0	0	0	0	1	1	1	1	6
NO	Nijmegen Station	Nm	1	0	0	1	0	1	1	1	0	1	1	1	1	7
NO	Zutphen	Zp	1	0	0	1	0	0	0	0	0	1	0	1	0	4
NO	Zwolle	Zl	1	0	0	1	0	0	1	1	1	1	1	1	1	7
NO	Zwolle Oosterhaven		1	1	0	1	0	0	1	0	1	1	1	1	1	9
NO	Zwolle Rgs	ZL Rgs	1	0	0	1	0	0	0	0	1	1	1	1	1	7
NO	Zwolle GE	ZL GE	1	0	0	1	0	1	1	0	1	1	1	1	1	9
NO	Leeuwarden	Lw	1	0	0	1	0	0	0	0	0	0	0	1	0	3
NO	Groningen	Gn	1	0	0	1	0	1	1	0	1	0	1	1	1	8
Zuid	Den Bosch	Ht	0	0	0	1	0	0	1	0	1	1	0	0	0	4
Zuid	Eindhoven	Ehv	1	1	0	1	1	0	1	1	1	1	1	1	1	9
Zuid	Roosendaal	Rsd	1	0	1	1	0	0	1	1	0	1	1	0	1	6
Zuid	Venlo	Vl	0	0	0	1	0	0	0	0	0	0	0	1	0	2
Zuid	Vlissingen	Vs	0	1	0	1	0	0	0	0	1	0	0	0	0	3
Zuid	Maastricht	Mt	1	1	1	1	0	0	1	0	0	1	0	1	0	7
Zuid	Heerlen	Hrl	0	0	0	1	0	1	1	1	0	0	0	1	0	3

Figure 3.11: Complexity matrix showing all the nodes of the NS and ranked on different criteria in order to score complexity (van Uden et al., 2017).

output data of HIP. The last Section 3.2.4 provides a technical detail on the optimization techniques and the solvers used by HIP.

3.2.1. General set-up of HIP

HIP solves the following sub-problems described in Section 2.3: matching, job-shop scheduling for non-human resources, parking, routing and splitting and combining (BMI Projectteam, 2018). Although the plan is to integrate crew assignment also to the tool, at this moment, this sub-problem is neglected. HIP is unique in the way that it solves all the sub-problems parallel, while the tools discussed in Section 2.3 all solved the sub-problems sequential. This has the advantage that sub-solutions are not dependent on earlier solutions, Which opens the solution space and provides more room to optimize the solution. The last advantage of HIP is that it includes the planning of the complete node, instead of only the SL. Trains from and to the SL are also planned since these movements might be the critical factor of a plan.

The solving and optimizing methodology of HIP is presented in order to understand the process that HIP uses to find an optimal solution for the planning problems. First of all, HIP generates an initial solution of the planning problem by solving the sub-problems simultaneously. An elaborated explanation of this part will be presented in Section 3.2.4. Once the initial solution is found, the optimization part starts. This initial solution is adapted using a local-search method. This means that a lot of small adaptations are being made to the solution, in order to optimize the plan. If the adaptations do not optimize the plan, these adaptations are removed, and new small adaptations are executed in order to optimize the plan. HIP is set up in such a way that a balance can be made between the available calculation time and the quality of the plan. If there is little time available, HIP will come up with an executable plan. If there is more time available, HIP will improve the plan and the quality of the final plan will be better based on the different optimization aspects.

HIP uses a cost optimization method to improve the initial solution. However, in order to find the initial solution, some constraints are relaxed Den Ouden (2019). This allows an initial solution with conflicts. These conflicts can be: routes crossing each other, track length violations or delayed departures. There are also hard constraints: resource feasibility and validity of matching. An initial solution is always found as long as the hard constraints apply. The objective of the optimization is the weighted sum of the conflicts. The weights are chosen to be proportional to the difficulty of solving the conflict. By optimizing the initial solution, the amount of conflicts is decreased, resulting in a better plan.

3.2.2. Input

In order to run HIP, data is needed to feed the model. The input for HIP consist of four components: timetable data, data about the trains that start and end their passenger moving services at the node, rolling stock data and node-specific data. From the timetable, the following information is needed:

- arrival and departure times;
- arrival and departure platforms;
- train number;
- the route of trains through the node.

The following information is needed of trains that end or start their passenger moving service at the node:

- train number;
- the arrival or departure time(depending on ending or starting their service);
- the arrival or departure platform;
- the composition of trains;
- train unit numbers (only for incoming trains, trains ending at the node);
- the services needed to be applied for each train unit.

Next to this information, there is also general information about rolling stock types that is needed to solve the planning problem. This general information (per rolling stock type) consists of:

- the amount of carriages;
- the length of train units;
- the standard time to combine train units;
- the standard time to split train units;
- the standard time to change the direction of a train (in Dutch: kopmaken);
- the standard time for a train driver to walk through a carriage (needed when changing the direction of the train);
- all train unit numbers that belong to a certain rolling stock type;
- the standard times of services of train units.

At last, also information about the node is necessary to solve the problem. This information consist of:

- the topology of the node;
- the length of tracks;
- the length of platforms;
- the assignment of special service tracks to distinct tracks (e.g. TWI);
- tracks that cannot be used due to maintenance (including start and end times);
- the amount of different personnel (train drivers, mechanics and cleaning teams) available at the node.

This data can be implemented directly to the tool, but a GUI (Graphical User Interface) can also be used to implement the data. The rolling stock and all the fixed aspects of the location data have to be implemented directly, while the GUI can provide scenario-based information. This information consists of timetable data, information of trains that end or start a service at the node and the tracks that cannot be used due to maintenance works.

3.2.3. Output

The output of HIP is a complete shunting plan of the complete node, including timetable trains passing though (picking up passengers) at the node. This plan contains the following information:

- the assignment of services to train units with a specific place and specific time;
- when which movements are made;
- how train units are positioned at all times;
- when and where train combinations are split or combined;
- the matching between arriving and departing train units

This data can be provided by an activity log, which logs all the activities that occur at the node. Another option is to present the results graphically by the use of the GUI (Figure 3.12). The GUI presents the planning graphically. The location of the train units are presented, just as services that are being applied to train units. The movement of timetable trains and shunting movements (and routes) are also integrated into the graphical presentation of the solution. Next to the planning, the GUI also presents the conflicts that are still present in the solution. The GUI can be used easily to present the conflicts in a graphical manner. Since the GUI allows the user to adapt changes to the solution, it is possible to solve conflicts by hand in order to make the planning feasible.

Figure 3.12: The Graphical User Interface of HIP. The black lines are the track of the node and the trains are represented in different colors (a color represents a rolling stock type). The time is represented underneath and can be changed by the buttons left and right of the time.

3.2.4. Algorithms for initial solution and optimization

HIP first calculates an initial plan that is optimized during the second stage. This Section describes first the algorithms used to create an initial solution for a planning problem. The second part will be focused on the optimization method.

Initial solution

The initial solution exists of a combination of sub-solutions of five sub-problems: Matching, job-shop scheduling for non-human resources, parking, routing and splitting and combining. In order to solve these sub-problems, two methods are used: Mixed-Integer Programming (MIP) and a greedy algorithm. Table 3.2 shows which methods are used to solve which sub-problem. First, the MIP cases are described, after which the greedy algorithm cases are described.

For the **matching** problem, the objective is to minimize the number of splits and combines possible. If train compositions do not have to be changed in order to prepare trains for their next service, this reduces the

Table 3.2: Overview of algorithm used to solve sub-problems of TUSP.

	Algorithm	
Sub-problem	**MIP**	**Greedy**
Matching	x	
Job-shop scheduling for non-human resources		x
Parking	x	
Routing		x
Splitting and combining		

complexity of the planning and the number of tasks that needs to be performed by personnel and increases the infrastructure availability. Notice that a swap of train units also belongs to the matching problem: a train that enters a SL with a train unit with four carriages in the front and a train unit with six carriages in the back exits the SL the other way around. Constraints are added to ensure that each arriving and departing train unit is in precisely one matching and that the train units are in the correct order. Time constraints are added to make sure there is enough time between arrival and departure to execute service task (including possible combining and splitting) and necessary travel time. Extra buffer time is added as well.

For the **parking** sub-problem, the goal of the MIP is to assign one parking track to every train unit. If all necessary service tasks can be performed on the parking track, the train unit will stay at the track. If not all the services can be performed on the track, a post-process assigns a random track later where these services can be executed to the train unit. The objective is to maximize the outcome of the number of trains assigned to a parking track multiplied by a variable that has a value depending on the selected parking track. The value is dependent on the following conditions:

- +0.1 if the route from the arrival track to the parking track is direct (no saw movements necessary);
- =0.1 if the route from the parking track to the departure track is direct (no saw movement necessary);
- -1.0 if the assigned parking track is the arrival or departure track of the train, but parking is not allowed on this track;
- -1.0 if the assigned parking track is the arrival or departure track of the train, but this track is used by at least 2 other arrivals or departures by other trains;
- +10.0 if the assigned parking track is the arrival and departure track of the train and the track is used by 2 or less arrivals or departure by other trains;
- +1.0 for every service task of a train unit that can be executed on the parking track.

The constraints make sure that the trains parked are not longer than the parking track and that train units that have to be split or combined are on the same parking track. The splitting and combining of train units is stated, meaning that no variables can occur. A train is split when it is parked the first time in the SL and a train is combined during the last parking instance at the SL. These activities are added to the plan.

With the solutions of the three sub-problems, a graph can be created that presents the sequence of actions (arrival, parking, servicing and departure) of different train units (Figure 3.13. Between every activity a movement is placed. However, sometimes these movements describe the movement from track X to track X, which is not a movement. In order to fully understand the graph, one needs to know that the splitting and combining action can be part of a movement (if two arrows come out of a movement, a train is split, and if two arrows enter a movement, two trains (units) are combined. Figure 3.13 shows the activities of two arriving trains. The bottom train arrives, is moved to the parking track, is parked (at parking track 1), is combined (two arrows going into one movement) and moved to the departure track and departs. The upper train arrives, is moved and parked at parking track 3. Since two arrows exit the movement, it is clear that the train is split at P3. The most upper part is moved, serviced, moved to the station before it departs. The middle train undergoes a series of movements, parking and scheduling after which it is combined with the bottom train. The combining happens on P1, since both trains are parked on this track. The information of this graph, together with the greedy algorithm, is used to solve the job-shop scheduling for non-human resources and routing. The overall goal of these two sub-solutions is to add time marks to this graph other than the provided arrival and departure time.

For the **routing** and **job-shop scheduling for non-human resources**, the best option is picked that is present. The job-shop scheduling for non-human resources is executed using the greedy algorithm which resource is

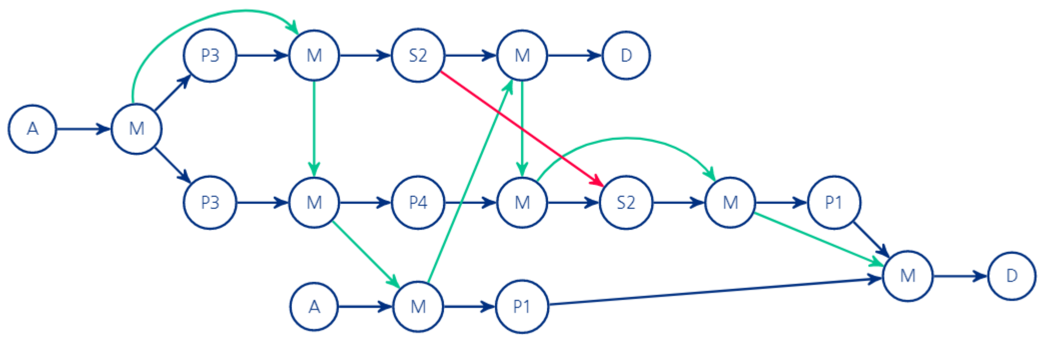

Figure 3.13: Activity graph (with movement between every activity) of two arriving trains (Den Ouden, 2019). The turquoise arrows represent the infrastructure constraints, the red arrows the resource constraints.

used when for what train. The red arrow shows the relation of the assignment of S2 to two different train units. The constraint that is being used is the fact that a resource can only be used to apply services to one train unit at the time. Assigning resources to train units result in begin and end times of services that are applied to these train units. This information is taken into account concerning the routing problem. For the routing problem, infrastructure constraints are needed: train routes cannot cross each other at the same time, and infrastructure needs to be available. To make sure not accidents are possible HIP reserves routes for the entire time the movement takes place. The choice for routes is based on the Dijkstra (shortest path) algorithm. Next to the choice of routes, there is also the choice when the movement is going to take place. An option of a choice can be: a longer route at time X or the shortest route on time Y.

Optimization

When an initial solution is provided, it is time to optimize the solution. The initial solution might contain conflicts since soft constraints are used to find the initial plan. If these constraints are exceeded, they might result in a conflict. The objective of the optimization is to minimize the number of soft constraints that are exceeded. In simpler terms: to minimize the number of conflicts that are present in the final solution.

If the initial solution is found, it presents an initial plan. This plan can be visualized as the graph presented in Figure 3.13 with start and end times of the activities and routes over the node. The time between two activities is based on the travel time of the route that is chosen. In order to optimize the complete plan, local search functions are used. Local search functions make small changes to the initial solution in order to find an improvement (Den Ouden, 2019). The neighbourhood of a solution is the set of solutions that can be reached by making a specific type of small change. HIP uses seven types of the neighbourhood. These are presented in Figure 3.14. In order to understand all the differences that are being made, small information is necessary. The shift movement and shift service neighbourhood mean that a shift is being made in the times a particular movement or shift take place. For example: instead of first cleaning train A, train B is cleaned first. The seven neighbourhoods create several possibilities (exact changes to the initial solution). The number of possibilities, from now on called candidates, changes from neighbourhood to neighbourhood. The number of candidates per neighbourhood for a test scenario with 22 train units is presented in Figure 3.15. Since some neighbourhoods create much more candidates than others, it might be possible that these neighbourhoods will dominate the search.

In order to find out if the changes have a positive effect on the total solution, the First Hill Climber method is used. This method uses all the candidates and shuffles them to get a random sequence of the possibilities. For every candidate, it is determined if the cost of the solutions is lower compared to the costs of the current solution. If this is the case, the current solution is changed to the candidate solution. Overall, this results in the fact that small changes are made to the initial solution. The advantage of using the First Hill Climber method is that it is a fast way to optimize a solution because the first improvement is selected. This might also be a problem of this method since it accepts small improvements made by only one candidate. In contrast,

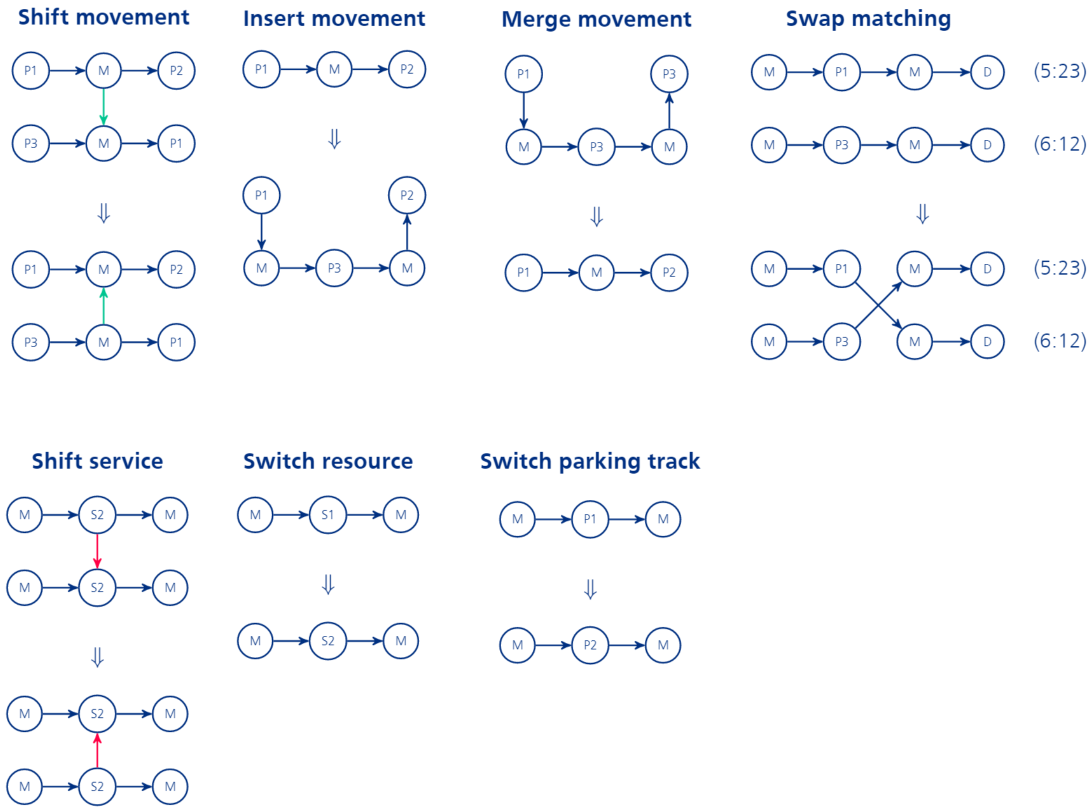

Figure 3.14: The seven types of neighborhood (changes to the solution) proposed by the local search function of HIP. Figure recomposed from (Den Ouden, 2019).

more significant improvements may be possible if multiple candidates are combined. Next to that, the First Hill Climber might get stuck in local optima. This means that no changes are made to the initial problem because the optimal solution is already selected.

Nevertheless, next to the problem concerning the local optima, there are two other problems. The first is the possible domination of neighbourhoods because of the big difference in candidates. The second problem is the fact that the search is pretty unfocused. Since the goal is to minimize the number of conflicts, it is more interesting to change plans where the conflict takes place instead of places where there are no conflicts. The local optima problem and the domination of neighbourhoods problem are solved in the algorithm by using a variable neighbourhood search in cooperation perturbations. In order to decrease the dominance of a neighbourhood, only a limited amount of candidates is selected per neighbourhood. In order to increase the focus of the search, the neighbourhoods are ordered according to success rate with an initial ordering of: switch parking track, merge movement, shift movement, swap matching, shift service, switch resource and insert

Neighbourhood	Number of candidates
Shift movement	812
Insert movement	29917
Merge movement	52
Shift service	21
Switch resource	268
Switch parking track	1130
Swap matching	4

Figure 3.15: Number of candidates per neighborhood of a test with 22 train units at the Kleine Binckhorst, a SL in The Hague, performed by Den Ouden (2019).

Table 3.3: Challenges of the HIP algorithm of which part is already solved. The R&D node logistics department is working to find and implement solutions of not yet solved challenges.

Challenge	Solved	Implemented solution	Possible solution
Local optima	Yes	Perturbation	
Domination of operators	Yes	Variable neighborhood search	
Improve initial solution	No		-
Unfocused search	No		Expanding neighborhood search
Plan evaluation	No		-
Relation neighborhood and planning problem	No		Customizable neighborhoods

movement. To solve the local optimum problem a perturbation action is added to the First Hill Climber that mutates the current solution, if a local optimum is found, by taking a random walk through the neighbourhood in order to optimize another local point. By implementing these three solutions in the First Hill Climber algorithm, the three problems are partly solved.

The unfocused search problem is not yet tackled. Next to this, there are some extra challenges to improve the algorithm. The team of R&D Node Logistics have defined the following remaining challenges: Create a method to generate a better initial solution, the plan evaluation method (defining the times to the graph) is computationally intensive and algorithm customization. This last challenge origins in the fact that some neighbourhoods cause improvements of the capacity on one type of planning problem, but decrease the capacity on another type of planning problem. By customizing the usage of neighbourhoods, the best settings can be determined for different planning problems. Table 3.3 lists the challenges, the already implemented solutions and the not yet implemented solutions in order to solve the challenges.

3.3. Answer to sub-questions 2 and 3

By the analysis of the current state of the planning process of NS provided in this Chapter, it is possible to answer the second sub-question: **How is the integrated node planning process of NS set up?** As many complex production processes, the hierarchical production planning method is used to solve the node planning problem. The availability of information is structured by three planning phases, with each subsequent phase increasing the amount of information available. Coupled to these phases are four departments of which two are responsible for the creation of the plan (NO and V&B), and the other two are responsible for the execution of the plan (MR and WvB). The level of detail increases if more information is available: from a plan for a general (basic-) day to a plan that is adapted to a single specific day. The combination of the amount of information available and the level of detail of the planning needed define the solution space for the departments that create the plan. The departments that execute the plan have an enlarged solution space since the standards of the planning tool used for the preparation are higher compared to the real-life standards. The dependency and hierarchy between the departments are high and can be defined as a real chain. The product (the planning) is transferred on strict deadlines from one department to the other. Although generalization of this production and execution system is applied as much as possible at the NS, node characteristics make it hard to copy the same process. Because of this difference, small adaptations of these processes are applied for different nodes. For complex nodes that are operated from a central location, an integrated approach has been found that includes extra planning actions like assigning tracks to train units.

The second part of this Chapter is used to answer the subquestion: **What are the characteristics of the developed Planning Assistance Tool?** The Planning Assistance Tool is the Hybrid Integral Planning method (HIP). At the moment of writing, the HIP is under development. However, it can be used to solve the following sub-problems: matching, job-shop scheduling for non-human resources, parking, routing, and splitting and combining. In contrast to other PATs, HIP includes the complete node. This means that movements between the station and the SL are planned as well. HIP solves and optimizes a planning problem in two parts. First, an initial solution is generated. Because of the use of soft constraints, this initial solution can contain conflicts. The optimization part has the objective to minimize the number of conflicts. By using an adapted First Hill Climber method, the optimization process is solved by a local search function. The method adds small changes to the initial plan and determines if these changes have a positive impact on the cost function. If this is true, the changes are made to the initial plan. The First Hill Climber method optimizes the plan, but in order to have better results, adaptations are made to the method. These adaptations solve the local optima problem and the fact that certain kind of changes to the initial problem might dominate the optimization. The input of the method consists of information of four different sources: timetable, trains that start or end at a node, rolling stock characteristics and node characteristics. The simple output is an activity log that describes all the actions that need to occur at the node. A Graphical User Interface (GUI) is created that can be used to insert input, and that visualizes the output of HIP. The GUI presents a clear list of conflicts that present in the optimal plan.

4

Experiment set-up

In order to find out how human-generated plans differ from computer generated plans qualitatively, it is necessary to perform comparison experiments. Where quantitative differences tell us something about the numeric differences between plans, it does not provide information about the quality of the content of the plan. For example, two plans can have the same amount of activities, but the sequence of these activities can be completely different. From a quantitative perspective, the two plans are similar, while the plans differ from a qualitative perspective. In order to perform a qualitative comparison, a mixture between a field study and a case study is proposed. This results in a comparison of plans that are based on real-life information: a real node and real expected train traffic. If a hypothetical (non-existing) plan assignment is used, planners might not be in their natural work environment, which might affect the results of the comparison. This Chapter presents the framework of the experiment.

Section 4.1 describes the objectives of the experiment to collect the necessary information. Section 4.2 presents the conceptual model that explains the creation of the human and computer-generated plan and the comparison. In order to make sure that the computer-generated plan can be compared next to the real human plan, it is necessary to implement a real-life planning assignment to the PAT. The data transfer from the current planning tool to the PAT has some assumptions and limitations. Section 4.3 describes the data transfer, the assumptions and limitations. Section 4.4 presents characteristics of the interview protocol that is used for the comparison by the experience planner. The conclusions (Section 4.6) is a summary of this Chapter and answers the sub-question: **How is it possible to identify lacking aspects of a Planning Assistance Tool using an empirical experiment?**

4.1. Objectives of experiment

Before the conceptual model is presented of the empirical experiment, it is first necessary to state the objectives of the experiment. The goal of this part is to gather those aspects that planners think are important to include in the evaluation of different planning options that the current algorithmic planning is not including. The results of the experiment will provide a list of aspects that could be considered to include in the algorithmic approach. Next to the list, the motivation why these aspects are essential for human planners is captured as well. This information is gathered by a comparison between a human plan and a computer-generated plan by an experienced planner. This research can only be conducted in the right way if the planning assignment and the solution space of the PAT resemble the planing assignment and solutions space of the real-life planner. In other words, the input and the rules of that the PAT needs to work with have to be the same as the input and the rules that a planner works typically with.

Next to qualitative aspects, other important differences or similarities are captured as well by the comparison performed. Since the PAT is not yet tested under real-life conditions, the information captured next to the main information can be used to improve the tool. The information will be notified and presented to the NS in order to show the points of attention for further development. The additional information captured by the comparison consist of:

- Differences between plan assignment

- Differences between solution space
- Technical infeasible options provided by the PAT. In other words, options that the PAT provides that are not possible in real-life.
- Similarities between humans and the PAT
- Strong and weak points of the PAT
- Opportunity of the usage of PAT in real-life by planners
- Opinion of planners on Graphical User Interface (GUI) of PAT

Next to that, it is important to notice that the experiment must be performed multiple times and with different planners in order to generate a general list of aspects. Focusing only on one planner or one single planning assignment might provide subjective data, which is undesirable. Repeating the experiment with multiple planners for different planning assignments minimizes the presence of subjective opinions and aspects.

4.2. Conceptual model of the experiment

The conceptual model proposed is given in Figure 4.1. The yellow boxes represent data, the blue boxes represent actions, and the final green box represents the result. A real-life planning assignment is used to create a plan in order to keep the working environment as normal as possible for planners. The final plan can be executed, which is not the case if hypothetical plan assignments would be used.

The real-life plan assignment is a plan that planners have recently worked on or that they are working on at the moment. The real-life assignment describes 24 hours, from 08:00 to 7:59 the next day. Since planners usually plan multiple days in a row, the day is picked where planners needed to make many adjustments because of problems that were in the plan. By picking a more difficult day, chances are higher that planners actively had to look for solutions to solve problems instead of looking through the planning on the automatic pilot.

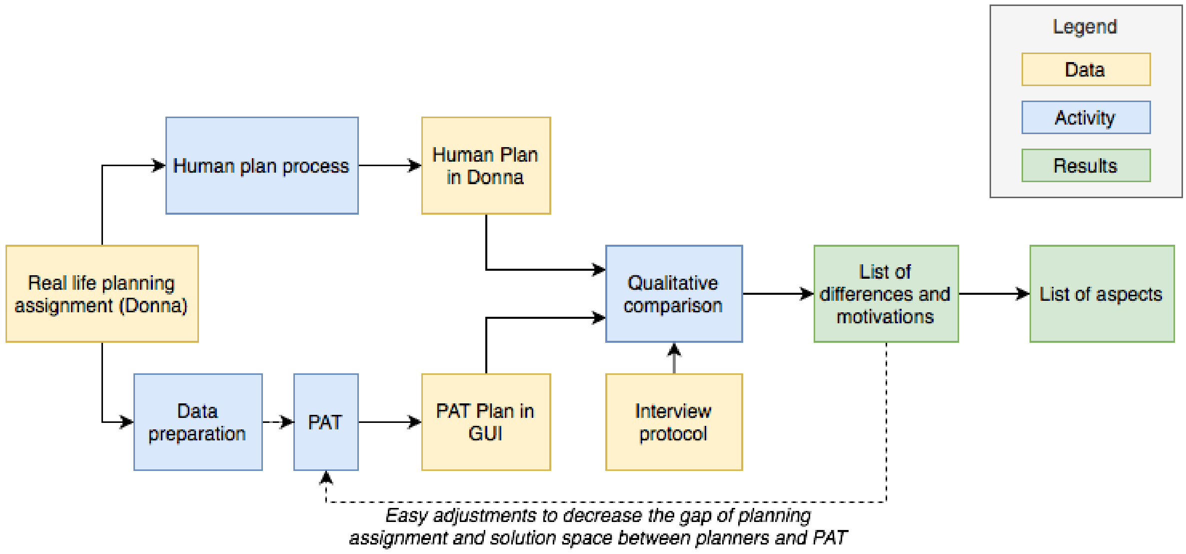

Figure 4.1: Conceptual model of the empirical experiment performed with human planners in order to compare human plannings with computer generated ones

For the Planning Assistance Tool data preparation and transferring is necessary. Data is collected in from the database of Donna (the current planning system), and this data needs to be transferred and adapted so that it can be used as input for the PAT since the PAT is not connected to Donna. This data preparation contains some assumptions and limitations, that is further explained in Section 4.3. Because of this, the planning assignment (input) and the solution space (input) differ compared to the planning assignment and solution space of the human plan process. However, the differences are small, so that the plan produced by the PAT can be compared to human planning.

Based on the experience of the first couple of experiment, easy to fix differences that arise between the PAT and the human planners can be changed for the last set of experiments. This change is advantageous because of the increase in the quality of information captured by the experiments increases. If these simple changes were not made, the same differences would occur for every new experiment made, while no time and attention are given to less apparent points. The dotted line in Figure 4.1 represents these small changes. These points have to be supported by multiple planners, which is why these changes are only applied after the first set of experiments. Next to this, the reasons for this change are of course written down in the list of results.

Once the two plans are created (by planner and PAT), the (qualitative) comparison can be executed. However, the presentation of the two plans differ. The human planning is presented in Donna (the same tool that the planner uses to create the plan). This tool provides much information on different displays. The displays used are the ones that are presented in Section 3.1.2 (Figure 3.3 and 3.4). These displays provide an overview of which train is where at every moment and provide information about movements on the node. The PAT has its own GUI (Graphical User Interface) that integrates the two displays of Donna. An example of this display can be found in Figure 3.12. This display provides the same information as Donna but on one screen instead of multiple. In order to structure and collect the output of the comparison, a clear interview protocol (Section 4.4) is created. Based on the output, a list can be created that describes the differences and their motivations. This list shows all the information captured from the experiment and therefore might cover certain points twice or might give multiple motivations for the same difference. A restructuring and grouping of the list containing the differences and motivations construct the list of aspects, which is the goal of this experiment.

4.3. Data preparation for PAT

Real-life data must be prepared and transferred for the PAT in order to create a plan based on the real planning assignment and with the same solution space. The data is collected from the database of Donna. The input data for the PAT consist out of three different documents: the scenario (which is the same as the planning assignment), the location file and the initial information. Next to these basic documents, it is possible to change, add or remove data manually in the GUI of the PAT. These four modes of input and their content will be discussed in this Section, as also their assumptions and limitations.

The initial data input contains the following interesting information about train types: length, carriages, time to combine and to split two units, time to change driving direction and time for a train driver to walk through a carriage. This information is all standard and can be implemented directly. The addition that has been made to this document is the amount of hypothetical train unit numbers. Long term planners (NO and V&B) do not work with train unit numbers yet since these are coupled to train units in the last phase of the planning (56 hours before operation). The PAT plans in the long term phase already with train unit numbers. However, every train unit number can only be serviced once in the PAT, while in real-life, a train unit can stay multiple times during one day at the Service Location. Therefore, every train unit that enters the Service Location is given a new (hypothetical) train unit number. These (in real-life) not existing train unit numbers must be added to the initial document in order to be recognized by the PAT. Another assumption has to do with a special train composition, which is used for the High-Speed Line (Inter-City Direct). This composition consists of two locomotives with a set of carriages in between. Because this train composition is never changed, this train is changed in the initial data to one train unit that consists of the amount of carriages + 2 (locomotives). The length of the train is changed accordingly.

The location file contains all the data of the infrastructure and the rules applicable to the infrastructure. The infrastructure is described by the tracks (length and the availability of a catenary overhead line (electricity)), switches, buffers and service resources. The service resources are connected to tracks, i.e. service platforms are connected to the tracks where there are service platforms, and the TWI is connected to the track where the TWI is placed. The location data also includes the number of employees (and the different types of employees present at the location). Since the PAT does not yet include the crew assignment sub-problem, the amount of employees present at the service location is high compared to the real amount of employees. The rules that are applicable at the location are mentioned underneath the tracks. There are mainly two rules that are mentioned: is it allowed to change the driving direction of trains and is it allowed to park trains at specific tracks. These rules are set by ProRail and are dependable on track characteristics and safety rules. Specific and short term changes to the infrastructure (i.e. out of service of certain tracks due to maintenance) are added to the PAT by the GUI, which will be explained further down.

For the PAT, the information from the planning assignment will be captured within a scenario. The scenario exists out of three sections: incoming trains (trains that need to be parked/serviced at the node), outgoing trains (trains that are parked that need to leave the node) and ongoing trains (trains passing through the node that might call at the station). The incoming trains need a train number, rolling stock type, (multiple) train unit number(s) (depending on the train composition), arrival time and arrival platform. Tasks can be added to the different train units: internal cleaning and technical check-up. From Donna, almost all this information can be gathered, except the train unit numbers (they are not yet assigned to train numbers) and tasks. This means that all information can be coupled directly to the scenario file if fictional train unit numbers are added. For every train unit, a new fictional unit number is chosen in order not to disturb the process. Tasks are added manually in the PAT. Adding tasks is complex since not all trains need to be internally cleaned or need to have a check-up. During the day trains are parked at a node, but no tasks are executed since there is no crew (cleaning teams and mechanics) present at the Service Location. Tasks are only added to the incoming trains that arrive after the evening peak, and that stay the night at the service location and train units that have a DCR claim (Daily Control and Interior Cleaning). Because these trains have to be identified manually, the tasks are added manually to the scenario as good as possible. For the outgoing trains, the same information except for the train unit number and the tasks is available and used for the scenario file. The train unit number is not given since matching is part of the sub-problems of the TUSP. However, in real-life it might be possible that some matches are fixed.

For ongoing trains, the train number, start and end times are needed, and a route needs to be defined. In the original database, real routes are available of ongoing trains, but unfortunately, these routes cannot be exported correctly. Luckily, the entrance point of the node, the platform and the exit point of the node can be exported with related times. This information is enough for the PAT to create a route that can be reserved for this train. An example of such a data set is the entrance of node by track AA at time 00:02, a waiting activity at platform 1 from 00:04 - 00:08 and exit of node at track GE at 00:10. The PAT will create a route through these points, but this route might be different compared to the route in the planning of Donna. Unfortunately, it is not possible to export ongoing trains of third parties (i.e. freight routes) to the PAT. This is a significant limitation because at specific nodes the number of ongoing trains of third parties is significant and an essential factor for node planners.

The PAT has a limitation when train units need to be added or removed from an ongoing train. At this moment, the PAT does not mention the fact that the train numbers of the ongoing and the stopping or starting train units are the same. Therefore, the PAT automatically identifies a conflict: the platform is reserved for the ongoing train, while the incoming or outgoing train also needs to be at that platform at that time. In order to minimize the number of conflicts, the train units that need to be added to an ongoing train are placed before the train unit is arriving, and the train units that are split-off are appearing after the ongoing train has left the platform track. This means that these train units have been brought forward or have been delayed in the planning assignment of the PAT, which is not the case in real-life since this will not be recognized as a conflict.

The GUI of the PAT can be used to add, delete or change data manually. Data that is added at the end is:

- Tracks that are out of service within a certain time frame (i.e. platform 2 from 02:00 - 05:00).
- Trains that are already at the SL. Since the PAT has a strict plan window (from 8:00 on day 1 to 7:59 on day 2), it might be possible that there are trains available before 8:00 on day 1 and after 7:59 on day 2. The number of trains is low (because of the morning peak), but they have to be added manually because otherwise a rolling stock imbalance or no matching might occur.
- Trains that stay at the SL at the end of the plan window (see explanation above).
- Trains that are not loaded correctly to the PAT. These trains are ongoing trains that have movements during the start time of the scenario (8:00) or during the end time of the scenario (7:59 +1) or midnight (23:59 -> 00:00 +1). The number of these ongoing trains is limited.

Table 4.1 summarizes the assumptions and limitations of the data transfer in order to insert a real-life planning assignment in the PAT. The table represents whether an assumption or limitation has an impact on the planning assignment or the solution space.

Table 4.1: List of the assumptions and limitations that had to be taken in order to transfer a real-life planning assignment to the PAT. The table also shows whether the assumption or limitation causes differences in the planning assignment or solution space.

Item	Assumption	Limitation	Planning assignment	Solution space
Addition of hypothetical train unit numbers in PAT	x		x	
Change of composition of one train for PAT	x		x	
There are more employees in the PAT scenario compared to original scenario	x			x
Tasks are added manually to train unit numbers as good as possible for PAT scenario	x		x	
Fixed matches are not taken into account in PAT scenario		x	x	
Routes of ongoing trains are different in PAT scenario compared to original scenario		x		x
Movements of third parties (other train operators) are not taken into account for PAT scenario		x		x
Some train units are delayed or brought forward in order to minimize conflicts in PAT scenario		x	x	
Some ongoing trains are removed because data is causing conflicts in PAT scenario		x		x

4.4. Interview protocol

For the comparison phase of the experiment, interviews are conducted. In order to keep the interviews effective, a list of sub-objectives is presented:

1. Characteristics of planning assignment
2. Characteristics of human planning
3. Personal opinion of own planning or planning of HIP
4. Identification of differences between plannings
5. Why is the solution of HIP not technical feasible
6. Why is the solution of HIP preferable?
7. Why was human solution picked?
8. What is causing non-preferred solutions of HIP
9. Improvement of the visualization and other characteristics of HIP
10. Grouping of differences and aspects

An interview protocol is used to structure the interview. The interview protocol is developed by the Interview Protocol Refinement (IPR) framework, created by Castillo-Montoya (2016). The method consists of four phases: ensuring interview questions align with the study's research questions, organizing an interview protocol to create an inquiry-based conversation, having the protocol reviewed by others and piloting it. By using the IPR framework, researchers can increase the quality of data they obtain from research interviews. The framework can be applied to structured and semi-structured interviews, but can also be useful for non-structured interviews. The exact interview protocol (in Dutch) used for the experiment is added in Appendix B. The appendix also contains the translated protocol in English. Before the start of these four phases, it is crucial to determine what type of interview will be conducted.

The interview will not be a structured interview since the topic (the comparison between a human plan and a plan created by the PAT) will differ from interview to interview. Because of this, a fully structured interview is not favourable. However, some structure can be added to the interview like an explicit begin and ending of the interview in the form of introductory and conclusive questions. Therefore, the semi-structured interview will be chosen to focus on.

Because people have complex experiences that do not unravel neatly before the researchers, the researcher wants to have intentional and necessary interview questions (Castillo-Montoya, 2016). In order to structure the questions according to the depth of the topic, the questions are divided into a begin, middle and end part. The beginning part is focused on understanding the planning assignment (identifying the hard and the easy parts of the assignment) and will provide general information about the state of the human solution.

The middle part is where researchers ask the questions most connected to the study's purpose (Rubin and Rubin, 2012). Within this middle part, the differences between the plans are provided in two different ways. First, the plans are compared chronologically, where the focus is on the difference in individual actions. The second part is focused on the state of the complete SL at it busiest moments (during the afternoon and the middle of the night), where a more general comparison is made between the two plans.

The end part of the interview summarizes and concludes the information provided above. An overall opinion about the PAT is asked for, that provides information about the priority of improvements. Next to this, the differences and aspects are grouped by the planner. The last question does not focus on the experiment objectives, but gives the respondent an option to provide information about the complete planning process.

In order to check the alignment of the interview questions with the experiment objectives, a matrix is created (Table 4.2) that maps interview questions to experiment objectives. As can be seen, the questions are ordered to the begin, middle or end part of the interview to give the interview some structure. As can be seen in Table 4.2, the interview questions are aligned to defining the differences, presenting which option is preferable and why this option is preferable. Based on these questions, a clear list can be created that describes the differences and the motivation behind these differences.

Table 4.2: Link between interview questions and interview objectives

		Sub-objectives interview									
Interview Questions		1	2	3	4	5	6	7	8	9	10
Evaluation planning assignment	1	x									
Evaluation planning assignment	2	x	x								
Evaluation planning assignment	3	x									
Evaluation planning assignment	4		x								
Evaluation planning assignment	5			x							
Evaluation planning assignment	6	x		x							
Solution comparison	1a				x						
Solution comparison	1b					x					
Solution comparison	1c						x				
Solution comparison	1d							x			
Solution comparison	1e			x							
Solution comparison	2a			x	x			x			
Solution comparison	2b				x				x		
Solution comparison	3						x				
Conclusion	1			x		x			x		
Conclusion	2			x						x	
Conclusion	3										x
Conclusion	4										x
Conclusion	5			x						x	x
Conclusion	6										

4.5. Case study

Specifics of the case study used for the experiment are described in this Section. The case study will exist out of a comparison between a human solution of a real-life planning assignment and the PAT solution. The comparison will be performed by a research group that exists out of experienced planners. The goal of the case study is to identify differences between the PAT solution and the human solution. In order to find out if these differences are node-specific or generic, a second comparison will be performed using a different node. Information about the PAT, the locations and the research group are presented within this Section.

4.5.1. Planning Assistance Tool

As stated already earlier in the report, the Planning Assistance Tool that is going to be used is called the Hybrid Integral Planning method (HIP). From this moment, HIP will be used instead of PAT in the report, but it should be clear that these terms mean the same. HIP is an integrated planning tool that is able to produce a shunting plan that includes the complete node: the train station and shunting yards. Since a real-life planning assignment is needed for the experiment, this data needs to be prepared so that it can be implemented into HIP. The data preparation is performed by a python script that uses a CSV file as input (export of Donna) and transforms all the necessary data to a YAML file that can be implemented in HIP. The script makes it possible to prepare and implement a lot of data easily.

4.5.2. Locations

The first test will take place at Eindhoven. This choice is made because this node is planned using integrated plannings and is dealing with evident capacity stress. Referring back to Chapter 3, a node with integrated planning means that planners plan the shunt movements towards and on the shunting yard and the service activities that need to take place. A non-integrally planned node is a node where only the shunt movements towards and from the shunting yard are planned. The skill to plan integral is needed for the experienced planners that are needed to compare the two solutions. Eindhoven is a suitable node where many planners are able to plan integral. Next to this, a certain level of capacity stress and change is necessary in order to differentiate the NO planning with the V&B planning. Next to these two points, Eindhoven will be used for a more significant project of the NS in which HIP will be tested. From the NS perspective, the introduction of HIP to the planners of Eindhoven by this experiment eases the start of the more significant project.

In order to understand the location a little bit better, a layout of the node Eindhoven is provided in Figure 4.2. Visible in the Figure is the station (1), and the two shunting yards (2 and 3). Both of these shunting yards are Service Locations; however, not on every track all services can be executed. Shunting yard 2 (Figure 4.3) is called the garden by planners and is accessible from the platform tracks with one movement. The yard is only accessible from the station side. Shunting yard 3 (Figure 4.4) is called the East-side (in real life, this shunting yard is placed east of the station), and this shunting yard is only accessible with a saw movement. The yard is accessible from both sides. This shunting yard does have three tracks that are accessible without a saw movement; however, interior cleaning services are not possible at these tracks. In the layout, the ongoing lines represent the tracks that are Centrally Operated Area (CBG). The dashed lines represent that Not Centrally Operated Area's (NCBG).

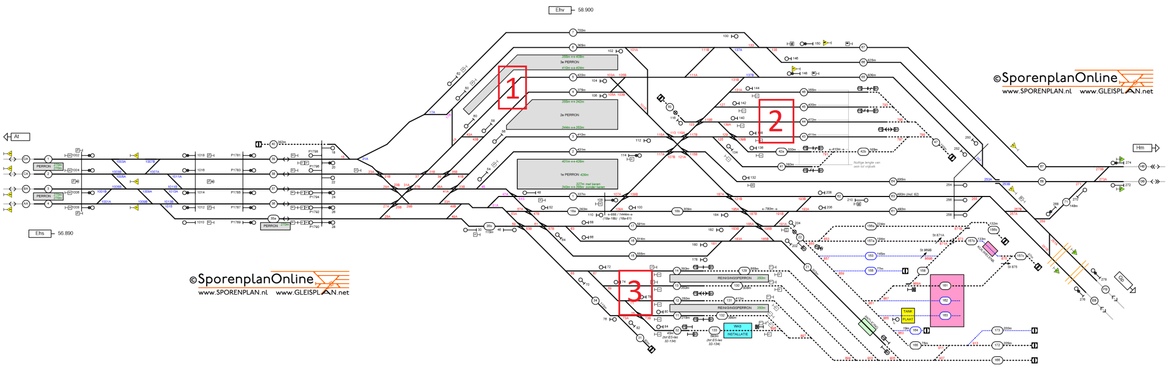

Figure 4.2: Layout of Eindhoven, with the station (1) and two shunting yards: the Garden (2) and the East-side (3) (Sporenplan Online, 2019).

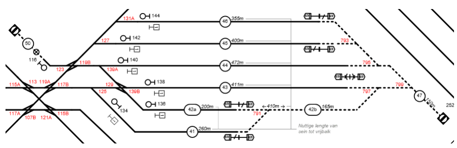

Figure 4.3: Layout the Garden (Sporenplan Online, 2019).

Eindhoven can be characterized as a hub in the national network of the NS. Intercity trains arrive here from the central part of the Netherlands in order to continue to Limburg and vice versa. The Hub aspect can be found back in the time table of Eindhoven since intercity trains are planned in such a way that passengers are able to change trains. Next to this, Eindhoven is also the regional hub. Sprinter trains enter the station

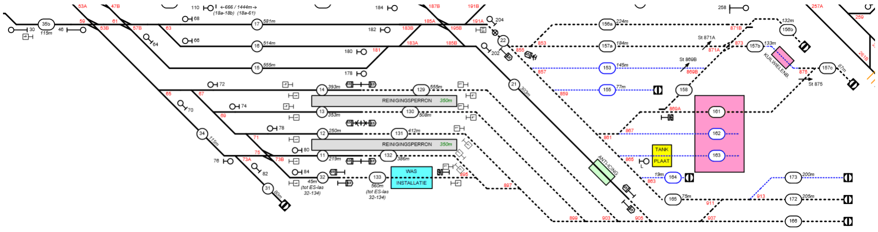

Figure 4.4: Layout of the East-side (Sporenplan Online, 2019).

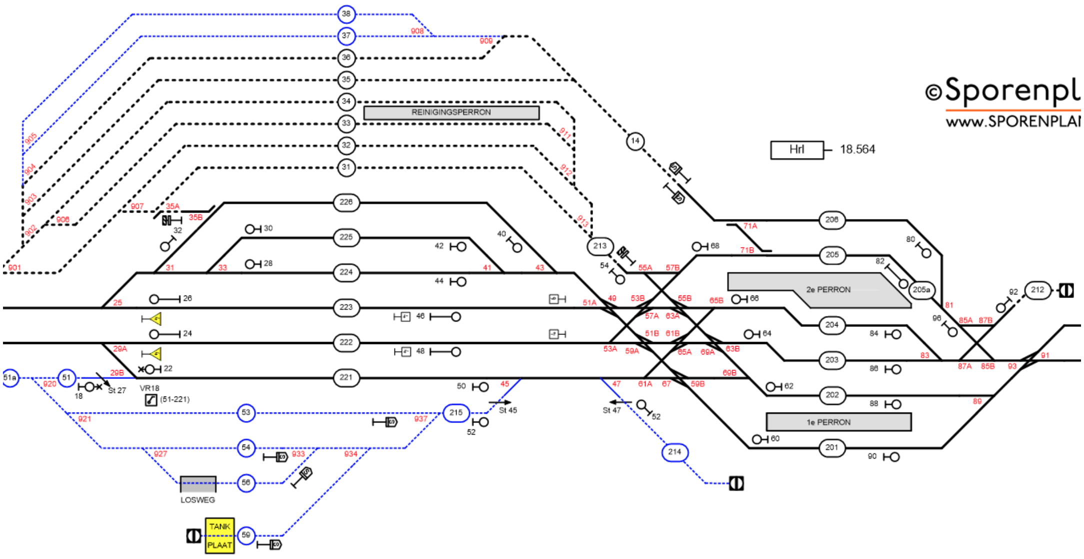

Figure 4.5: Layout of Heerlen (Sporenplan Online, 2019). On the right the station is visible with two platforms.

before the intercity train arrives and depart after the intercity has left. These connections allow the regional passenger to transfer to the national network and vice versa. Because the passenger demand to Limburg is lower compared to the passenger demand toward the central part of the Netherlands, trains towards Limburg are sometimes shorter compared to trains going towards the central part of the Netherlands. For the node, this means that many trains need to bee split-off our coupled during the day. For the regional traffic, this is only the case after the evening peak.

As stated in the introduction of the case study, a verification node is used in order to find out if differences are node-specific or general. The verification node is Heerlen: another node in the South of the Netherlands but smaller compared to Eindhoven. Since NS only has one service that ends at Heerlen, only one type of rolling stock is parked and serviced at this node. Next to the singularity of rolling stock, Heerlen differs since it is a terminal station for the NS, meaning that the NS only has one platform track. There are no ongoing trains of the NS calling at Heerlen. The layout of Heerlen is presented in Figure 4.5.

Heerlen is picked as a verification node because many planners that plan Heerlen are also certified to plan Eindhoven. Next to this, the node is also planned integral, meaning that services to train units are planned as well. Since the goal is to verify the generic and node-specific differences, new differences will not be added to the list of differences found for Eindhoven.

4.5.3. Research group

The node planning process consists of four different types of planners (for more information, see Section 3.1.1). For this experiment, the Voorbereiden en Bijstellen (V&B) planners are used. The V&B planner has as input a Basic Day update (BDu) planning and specific day information in order to create the Specific Day (SD) planning. According to the number of changes (from the specific day input), the plan assignment increases for the V&B planner. However, the period for the V&B planner to solve the planning assignment is from x-4 weeks to x-56 hours (with x being the day of operation). Since Eindhoven is an integrated node, the V&B planner plans all shunt movements and services: from the station towards the shunting yards as well as the movements in the shunting yards. Another advantage of picking the V&B planners is the fact that they are with a more significant number compared to the NO planners.

The research group of Eindhoven consists of 5 planners. With these 5 planners, 6 different real-life planning assignments are used to compare the solutions of the human planners and HIP. The planner that has created the human solution is used for the comparison. A simple calculation shows that one planner has compared two planning assignments. This planner is planner A, and the two planning assignments are called A1 and A2. Characteristics of the different real-life planning assignments are presented in Table 4.3.

Table 4.3: Quantitative characteristics of the different real-life planning assignments used.

Indicators	**A1**	**B**	**A2**	**C**	**D**	**E**
Date	2-3 Jan	5-6 Dec	30-31 Dec.	26-27 Feb	24-15 Feb	6-7 Feb
VIRM4	15	19	23	38	36	17
VIRM6	12	17	12	17	20	16
FLIRT3	80	57	79	58	69	73
FLIRT4	43	75	47	81	80	67
SLT4	1	0	0	2	1	1
SLT6	1	7	3	2	4	2
ICR Gvc-Ehv 9	36	36	36	36	36	36
ICM3	3	3	3	3	3	3
Sum of train units	191	214	203	237	249	215
Arriving trains	148	148	138	152	154	144
Train units per arriving train	1,29	1,45	1,47	1,56	1,62	1,49
Departing trains	154	156	147	161	163	148
Train units per departing train	1,24	1,37	1,38	1,47	1,53	1,45
Tracks out of service	0	0	0	13	0	0
Ongoing trains	536	534	512	547	556	511
Level of difficulty	Easy	Medium	Medium	Hard	Hard	Medium

4.6. Answer to sub-question 4

In order to conclude this Chapter an answer is provided to the sub-question: **How is it possible to identify lacking aspects of a Planning Assistance Tool using an empirical experiment?** An experiment set up is presented where the preparations are explained of the comparison between a human plan and the HIP plan. First of all, a real-life planning assignment is picked that a planner has worked on or is working on. If a choice is possible, the planning assignment is picked that was difficult for the planner to solve, since it contained many problems. By picking such a real-life planning assignment, it is ensured that the planner is comfortable with the plan and can easily remember why he or she made specific decisions.

For the HIP plan, the real-life planning assignment had to be prepared for the input requirements of HIP. A script is written that quickly translates the data from the real-life planning tool Donna to the used Planning Assistant Tool developed by the NS (HIP). Unfortunately, a 100% correct translation between the information of Donna and HIP cannot be created due to missing data. The problems that this partially incorrect or lacking data transfer causes are notified. For some missing information, assumptions are made. The rest of the missing information is characterized by limitations of HIP for real-life plan assignments.

Based on the input generated, HIP can create a solution to the real-life planning assignment. The next step is to compare the human plan with the HIP plan in a qualitative way. Since experience is necessary, the planner that has created the human plan is asked to compare the two plans. In order to structure the comparison process and the flow of information, an interview protocol is created. The results of the comparison are captured in a list of differences between the human plan and the HIP plan and motivations behind certain decisions of planners. In order to improve the qualitative comparison, easy and small adjustments are done to HIP to improve the information gathered by the comparison. These easy and small adjustments had to be notified by multiple planners before they were changed to HIP in order only to adjust things that are supported by multiple planners. By restructuring of the list of differences and motivations, the list of aspects that are not included in HIP is created.

5 Results

This Chapter describes the results of the empirical experiments of the case-study described in Chapter 4. These results consist of the list of differences and the list of aspects, as can be seen in Figure 4.1. The focus is first placed on the list of differences identified during the experiment. If the main information about the differences is presented, the list of aspects is explained using clear links towards the differences found.

This Chapter first presents the main differences that were identified for almost all experiments (Section 5.1). The second Section (5.2) describes the differences there are between planners. Section 5.3 presents the results of the verification node in order to find out whether differences are generic or node-specific. The last Section (5.4) describes the transformation from the list of differences to the list of aspects by categorization of the differences.

5.1. Main differences

From the empirical experiments, a list of 59 differences is identified between the human plan and the HIP plan. The motivation why humans plan different compared to HIP was also collected. A complete overview of the results can be found in Appendix C. This list contains all the differences highlighted during the experiments. If differences were identified during multiple experiments, the differences were grouped. The last column presents the number of experiments where the difference was identified. In this section, all the differences are mentioned that are identified in 4 or more experiments. The exact difference number is placed between brackets at the end of the paragraph. From the ten main differences, two differences (difference 5 and 9) were expected due to technical or data limitations as described in Section 4.3.

First of all, there was the general difference that trains were not parked at the correct service tracks, because of the misassignment of services to train units. Human planners assign services to trains based on intuition and information that cannot be exported from the database. The assignment of services is a manual job because it is highly dependable on a lot of different factors (e.g. the time the train stays at the SL or if a train is already clean and checked). Because of this, it is hard to define a correct set of characteristics that can be used to assign services to train units. Because of this, for all experiments there were train units that did not have any services assigned or had wrong services assigned, resulting in the parking of these train units on incorrect tracks (difference 2).

Next to this, HIP parked the same type of rolling stock on different tracks, whereas the planners prefer to park the same type of rolling stock at the same track since this eases the possibility to exchange train units without extra movements (difference 3).

Train units of third parties were not included for HIP, resulting in specific movements of the trains of the NS that were not possible for human planners. This caused differences for all experiments (difference 5).

Another crucial point is that human planners have an evident preference to park certain rolling stock types at certain tracks. This preference is based on specific characteristics of this rolling stock, for example, the length of rolling stock (difference 7).

Another main difference that was identified in many experiments is the fact that for a sprinter service that calls at Eindhoven and changes train number, HIP made certain irrational matches. Where usually the incoming train will continue after a stop of some minutes as the outgoing train, HIP suggested changing the train with a train from the SL, most often causing a departure delay. Within human plans, such irrational matches were never found. It has to be stated that this problem only occurred if a specific part of the incoming train was split off the train combination in order to be placed at the SL (difference 9).

Another generally found difference is that human planners try first to fill the parking tracks where services can be executed before filling the tracks where trains can only be parked. HIP does not consider this in its choice for a parking track (difference 11).

Based on the new configuration of trains and their future usage, human plan differently compared to HIP. Humans park specific train units at specific tracks if these need to be paired with later incoming trains or if trains leave the Service Location as an empty rolling stock train (no need to pass a platform track). HIP does not take this into regard, meaning that these specific tracks are less used in the plan of HIP (difference 15).

Within the human plan, parking tracks are filled with more trains compared to the the HIP plan. The utilization of parking space if more important for humans compared to HIP (difference 20).

the HIP plan contains a lot more movements compared to the human plan. Parked trains are moved on the SL in order to be serviced by the HIP plan compared to the human plan, where the train is already parked at the service track. HIP chooses routes that contain more saw movements compared to the humans considering the difference in plans (difference 21).

There are some tracks that are highly utilized by HIP and rarely in the human plans. This is specifically the case when looking to one track in the back of the "garden". Due to regulations between planners, this track is seldom used by human planners, whereas HIP uses this track a lot (difference 26).

5.2. Differences between planners

Next to the main differences, some differences are less widespread. Some of these differences only occurred in one experiment. The motivations behind certain differences were sometimes contradictory between planners. In other words, the solution of HIP was identified as wrong, but the explanation was largely different between planners. These differences between planners were notified and are explained in this Section.

The first difference between planners is the amount of time for passengers to get off a train before this train is shunted towards the SL (difference 4 and 6). There are some agreements for certain train services, however, no official rules are present.

The second difference that was remarked is how planners handle new (not already parked) rolling stock types at the SL (difference 16). Two options are possible. The first option is to park the train of this not yet present rolling stock type on a new track. By doing this, the planner limits the chance that this new rolling stock type interlocks an certain rolling stock type. The first option optimizes the robustness of the planning in a way that every rolling stock type is accessible at every time without moving another train unit. The second option is that the train unit of the new rolling stock type is parked at the same track, interlocking the train unit(s) of other rolling stock type(s). This option optimizes the utilization factor of the parking tracks.

The last difference between planners is the possibility to park trains (when serviced) at the platform during the night (differences 3 and 14). There are different opinions about the possibility to store trains at the station (on a platform track). In general, it is wanted to keep the train as long as possible at the SL to perform services. However if the services are planned correctly, certain planners agreed that it is possible to park trains at platform tracks if capacity of the SL is exceeding. The fact that the plan of HIP sometimes stores multiple trains at the platform track (in the order of departure) also caused different reactions of the planners. Some planners found it a feasible (and executable) solution, while other did not found it a feasible solution since it causes uncertainty for passengers which train is departing.

5.3. Node specific or generic differences

With the verification process executed on Heerlen, it was possible to verify whether differences are node-specific or generic. In Appendix C, the node-specific and generic differences are identified. As can be identified, there are not that many node-specific differences. However, the generic differences are categorized into different types. First of all, there is just generic, meaning that the problem is fully generic. The problem does occur at Eindhoven as well as Heerlen.

The second group are the generic differences that include data that is dependent on train services. Different types of train services call at different stations, which all have their own rules. An example of Eindhoven is the intercity to The Hague Central Station. This train uses a small part of the High-Speed Line, and because of this particular rolling stock is used that require specially trained train drivers. Every train service comes with different rules that have to be taken into account at the node.

The third group consist of the generic differences that include data-dependent on the layout of the node. For these differences, Section 3.1.5 can be seen for the differences. Different node layouts require different techniques and rules. However, a node often consist of a mixture of layouts, meaning that these techniques and rules have to be mixed as well.

The fourth group of generic differences are based on safety rules. These are simple rules that apply everywhere, but these rules have been unknown or not specified to the developers of HIP.

The last group of generic differences are dependent on the different rolling stock that needs to be parked and serviced at a node. Every type of rolling stock requires different services and comes with different rules. These rules are generic, but differ from node to node based on the presences of the rolling stock.

The verification performed at Heerlen showed that many differences are generic. However, two things need to be notified: First of all, Heerlen is a small node that has a less complex planning assignment compared to Eindhoven. The characteristics of Heerlen differ compared to the characteristics of Eindhoven. The second point is that there might be differences that apply both on Eindhoven and Heerlen, but not at other nodes. This might cause the problem that we assume certain generality, but that this might not be true.

5.4. Aspects

Based on the difference and the additional information like motivation and condition, the results of the empirical experiments were translated into aspects using categorization. This Section first describes these aspects in general (Section 5.4.1). Subsection 5.4.2 represent links between the differences and aspects in order to understand the aspect better. An overview of the relation between the experienced differences and the aspects is presented in Table 5.1.

5.4.1. Aspects in general

Six different aspects were recognized being:

- simplification of planning assignment;
- capacity optimization;
- increasing flexibility;
- optimizing plan for KPI's;
- norms and agreements;
- safety.

These aspects are all shortly explained underneath. The linkage between the aspects and the differences is provided in the Subsections.

Simplification of the planning assignment is one of the major aspects that are typically in the human plan process, but is lacking in the HIP process. For humans, the complete shunting plan process is enormous and therefore, they tend to divide the big problem into smaller problems. Even the TUSP consist of multiple subproblems in order to make the problem more comprehensible. However, planners think already ahead when solving the plan assignment. By making rational decisions, they simplify the planning assignment because the number of needed actions is decreased. These rational decisions can be translated into a smaller solution space for human planners. Instead of taking into account every possibility, the amount of possibilities is limited to the preferable. If these possibilities are all not possible, the set of possibilities will be enlarged by less preferred options. Because of these preferred options, planners often talk about preferences (e.g. the preference to park a specific rolling stock type at a particular track, or the preference to use a specific route). However, there is always a rational thought behind these preferences. It can be stated that for almost all the differences, the preferences are widely supported by all planners.

The second aspect is **capacity optimization**. Capacity can best be described as the infrastructure capacity and resources capacity. Infrastructure capacity can be translated to parking track usage and the usage of certain tracks for saw movements. The resource capacity can be related to the availability of train drivers. The fewer tasks for train drivers, the higher the availability of these train drivers, the higher the capacity. Service resources like the Train Washing Installation are also taken into account. Next to this, it can be stated that the minimization of movements leads to a higher capacity since new movements are possible.

Planners also try to **increase flexibility** in the operational phase. In other words, the planners try to create a plan that takes into account the effects of disruptions. An example of this can be the fact that planners try to keep one parking track free, in case a train cannot continue its service and needs to be parked at the SL. Disruptions can have a big impact on the shunting yard and indirectly the planning within the yard. By thinking rational about several disruptions scenarios and real-life experience from dispatchers, a clear list of preferences is known by planners to think about when planning. However, these preferences are subordinate to preferences that ease the planning assignment because planners first need to create a feasible planning before optimizing it. The addition of flexibility is therefore based on the possibilities within the plan.

Another aspect is the **optimization of the plan for KPI's**. The V&B team itself does not use KPI's to measure their planning, but their planning can have an impact on the KPI's that are used for other teams (like dispatchers). Optimizing for KPI's means that a human planner makes certain choices because they might be better for service in general. These KPI's can be train-related (amount of dirty trains starting their service) but also passenger related (amount of passengers that is able to make their connection). Planners can adjust this last KPI since they can change the arrival and departure platform of trains. If connecting trains are arriving at platforms that are located close to each other, the connection time (the time a passenger needs to go from one train to another) is lower than if the trains arrive at two platforms that are located far from each other.

Norms and agreements are also an aspect that is not yet fully integrated into HIP. Some norms are not implemented at all in HIP. There is also the challenge that norms and agreements can be implemented as hard or soft constraints. The level of hardness of a norm or agreement is dependable on the planner.

The last aspect that planners take into account is the **safety** aspect. Safety contains traffic safety, but also the safety of the crew. The general traffic safety rules are applied in HIP, but certain safety rules are based on the specific characteristics of the train or location. Next to this, the safety rules of the crew lack in HIP. The safety rules that apply narrow down the solution space, limiting the number of possible options. The last thing that is unwanted is that crew is not able to perform planned actions or need to stop specific actions, because safety rules are (going to be) violated.

Most of the differences could be related to aspects, however, some differences are caused by technical or data limitations. These limitations are partly expected, because of existing limitations of the algorithm, the lack of data and the way conflicts are identified. However, there are also some limitations that were unknown to developers. From another perspective, not all data is known by planners, and sometimes hard norms are not yet official.

5.4.2. Relation between differences and aspects

This Section describes the relations between the differences and the aspects. First, Table 5.1 is presented showing the links between the differences and the aspects. The links are discussed per aspect afterwards. At the end of each paragraph, the number(s) between brackets resembles the difference(s). Since differences might have multiple relations to different aspects (see Table 5.1), the most important relationship is represented.

Overview of relations between differences and aspects

Table 5.1 is an overview of the relation between the experienced differences and aspects. A double plus (++) represents the most important relation. A single plus (+) represent smaller relations or indirect relations.

Table 5.1: Relations between the experienced differences and the aspects. Some differences can be related to multiple aspects. The most important or direct relations between a difference and an aspect is marked by a double plus (++). Smaller or indirect relations are marked by a single plus (+).

#	Difference	Increasing flexibility	Optimizing plan for KPI's	Technical limitation	Simplification of planning assignment	Capacity optimization	Norms and agreements	data limitations	Safety
1	HIP brings trains earlier to the station compared to planner.	++	+						
2	Rolling stock is not placed at the correct service tracks, because these services have to be noticed by a planner and have to be assigned by the planner		+	++					
3	Planners like to park rolling stock types and forms (amount of carriages) on the same track. HIP parks these randomly over the shunting yard				++	+			
4	HIP does not take into account time for passengers to get out of trains, resulting in trains that are immediately moved towards the SL after arrival at the platform.						++	+	
5	HIP does not take into account train movement of third parties. Human planners do have to take this into account limiting the movements over the node.							++	
6	After arrival, trains are moved later to the SL in human plans compared to the plan of HIP						++	+	
7	Human planners prefer certain tracks for certain rolling stock combinations because of optimal usage of the tracks				+	++			
8	Human planners take into account the fact that cleaning teams and mechanics are present at the SL during certain time periods. In HIP, the trains can be serviced 24/7		+				++		
9	HIP contains a lot of difficult matches between train units, while in real life these matches are set.			++					
10	HIP does not assign services to trains that only need to be serviced before moving on to another node		++					+	
11	HIP parks trains at parking tracks that are not servicing track, while service tracks are still available. Human planners will always first fill the service tracks, before using only parking tracks	+	+		++				
12	the HIP plans to change directions (performs saw movements) at tracks where this is not allowed because of (traffic) safety.								++

Table 5.1: Relations between the experienced differences and the aspects. Some differences can be related to multiple aspects. The most important or direct relations between a difference and an aspect is marked by a double plus (++). Smaller or indirect relations are marked by a single plus (+).

#	Difference	Increasing flexibility	Optimizing plan for KPI's	Technical limitation	Simplification of planning assignment	Capacity optimization	Norms and agreements	data limitations	Safety
13	It is possible to change driving direction at certain tracks, only if this is according to the traffic safety rules								++
14	HIP brings multiple trains to the platform in the early morning, exceeding the platform length. In human plans, this is never identified		+						++
15	For certain rolling stock types or train series, planners prefer certain tracks to place these train units		+		++				
16	Some planners park train units of a different rolling stock type at a track that is not yet occupied	++							
17	For certain rolling stock types or train series, planners prefer certain tracks to place these train units		+		++				
18	The time to change directions for certain times are different between HIP and the human planner							++	
19	Some planners prefer to park trains at already occupied tracks in order to keep other tracks empty	++				+			
20	Human planners try to optimize the utilization rate of parking tracks					++			
21	The human plan has less movements compared to the HIP plan	+			++	+			
22	Planners combine trains, then move them and then split them, whereas HIP moves these trains separately					++			
23	HIP brings trains earlier to the station compared to human planners, which is not preferable.	++	+						
24	HIP brings trains earlier to the station compared to human planners, which is not preferable.	++	+						
25	the HIP plans empty material first to a platform track, after which it is moved the the SL. A human planners uses these parking tracks or plans the train immediately to the SL	+		+		+		++	
26	Some tracks are highly used by HIP and almost never by the human planner.				++				+
27	Human planners park a long stay train at a track where it does not hinder other trains and therefore also does not have to move				++		+		
28	HIP used tracks that are not assigned to the NS or that cannot be used for planners						++		
29	the HIP plans a long parking break between certain saw movements. Human planners always try to perform a saw movement with an as small break as possible.	+				++			
30	This platform change cannot be done by HIP, while the planner is able to change arrival (and departure) platforms			++	+				

Table 5.1: Relations between the experienced differences and the aspects. Some differences can be related to multiple aspects. The most important or direct relations between a difference and an aspect is marked by a double plus (++). Smaller or indirect relations are marked by a single plus (+).

#	Difference	Increasing flexibility	Optimizing plan for KPI's	Technical limitation	Simplification of planning assignment	Capacity optimization	Norms and agreements	data limitations	Safety
31	Within HIP, trains are allowed to have a delayed departure			++					
32	In real life, trains are allowed to enter service dirty, whereas in HIP, this is not allowed			++					
33	HIP used tracks that are not assigned to the NS or that cannot be used for planners						++		
34	the HIP plans empty material first to a platform track, after which it is moved the the SL. A human planners uses these parking tracks or plans the train immediately to the SL	+				+		++	
35	Not all trains are serviced over night in HIP, while the human planner offers a service for every train unit that stays at the SL overnight.		+	++					
36	the HIP plans to change directions (performs saw movements) at tracks where this is not allowed because of (traffic) safety.								++
37	The human plan has less movements compared to the HIP plan	+			++	+			
38	Routes in the human plan are different compared to the plan of HIP							++	
39	Human planners park a long stay train at a track where it does not hinder other trains and therefore also does not have to move				++				
40	In order to optimise the use of the tracks, planners prefer to place certain rolling stock types at certain tracks				+	++			
41	In order to optimise the use of the tracks, planners prefer to place certain rolling stock types at certain tracks				+	++			
42	Planners prefer to fill first all the service tracks, before the parking tracks without service resources are used		+		++				
43	HIP reserves complete routes for ongoing trains, meaning that these tracks are all not possible to be used, while the human planner has more freedom because the exact location of a train is used and not the complete route.			+				++	+
44	HIP only moves split-of train units of ongoing trains after the ongoing train has left the station. Human planners plan the movement of the split of train unit sometimes earlier, by moving it in the opposite direction of the ongoing train to the SL.					++			
45	Rolling stock is not placed at the correct service tracks, because these services have to be noticed by a planner and have to be assigned by the planner		+	++					

Table 5.1: Relations between the experienced differences and the aspects. Some differences can be related to multiple aspects. The most important or direct relations between a difference and an aspect is marked by a double plus (++). Smaller or indirect relations are marked by a single plus (+).

#	Difference	Increasing flexibility	Optimizing plan for KPI's	Technical limitation	Simplification of planning assignment	Capacity optimization	Norms and agreements	data limitations	Safety
46	HIP reserves complete routes for ongoing trains, meaning that these tracks are all not possible to be used, while the human planner has more freedom because the exact location of a train is used and not the complete route.			+				++	+
47	HIP identifies a conflict if a train is parked at a track where this is not possible, even if this duration is short. For human planners, this is not a conflict, which means that no further actions are necessary to make the plan feasible.			++					
48	Track 166 can be used for short parking in order to combine and split trains or to wait before the next movement is possible.			++					
49	HIP only moves split of train units of ongoing trains after the ongoing train has left the station. Human planners plan the movement of the split of train unit sometimes earlier, by moving it in the opposite direction of the ongoing train to the SL.					++			
50	In the garden area of the shunting yard, HIP uses track 46 a lot in order to bring trains from the right side of the yard to the station. These movement include a saw movement. Human planners never apply these kinds of movements, because their set-up is based on the FILO (First In Last Out) system.				++				+
51	Because of the high utilization rate of track 46, simple rematching of rolling stock is not executed. These simple rematches reduces the amount movements necessary.				++	+			+
52	Empty rolling stock is planned by HIP over platform tracks, which is not favored by human planners	+			++				
53	The human planner uses the parking tracks that can be entered from the main tracks immediately (no saw movement) and from both directions.				++				
54	the HIP plans to change directions (performs saw movements) at tracks where this is not allowed because of (traffic) safety.								++
55	HIP parks this rolling stock in the garden, whereas the human planner prefers the south side of the shunting yard.	+			++				
56	HIP allows services at tracks where service resources are not present		+	++					

Table 5.1: Relations between the experienced differences and the aspects. Some differences can be related to multiple aspects. The most important or direct relations between a difference and an aspect is marked by a double plus (++). Smaller or indirect relations are marked by a single plus (+).

#	Difference	Increasing flexibility	Optimizing plan for KPI's	Technical limitation	Simplification of planning assignment	Capacity optimization	Norms and agreements	data limitations	Safety
57	HIP parks trains at tracks, where long term parking is not permitted	++			+				
58	It is possible to change driving direction at certain tracks, only if this is according to the traffic safety rules								++
59	Some tracks can only be used if other (comparable) tracks are out of service and specific rules are made with the original user of these tracks (third parties)						++		

Increasing flexibility

Planners also took certain actions to add flexibility to the planning and indirectly to the operational phase. The first difference that led to this reason is the fact that HIP brings trains to the station when their service tasks are executed, while a human planner will maximize the parking time of the train at a service location. By doing this, the planner gives Operation and Service (responsible for the service execution at the service location) more flexibility. They have a bigger period to execute services to the train. This time is maximized, taking into regard the fact that in the morning many trains have to be brought to the station in a small period. Since the amount of train drivers is restricted, it might be needed to bring trains earlier to the station. Indirectly, this aspect has a direct influence on the KPI's, since there is more time to execute services. (1, 23, 25)

Another form of increasing flexibility in the operational phase is that human planners prefer to park different types of rolling stock on different tracks. By doing this, every type can be brought to the station in one movement. The inclusion of one rolling stock type is not wanted since disruptions might demand the immediate need for specific rolling stock. This aspect conflicts with the optimal usage of parking tracks aspect. (16)

Within the HIP plan, trains are parked on tracks that are used for saw movements, which means they are crucial for the accessibility of different parts of the emplacement. Parking a train at these locations will block these parts, which indirectly decreases the flexibility during the operational phase. Therefore, human planners do not park trains on these tracks. (57)

During the day, some planners prefer to park trains at tracks where other trains are already parked in order to keep other tracks completely free. By doing so, they create fallback options if a disruption occurs, while the initial planning can still be executed, because the initial planning is based on only a certain amount of available tracks.

Optimizing plan for KPI's

Next to normal trains that are serviced and parked at Eindhoven, this node is also used to service trains that are parked at another node. In other words, certain trains are only cleaned in Eindhoven and are parked at Weert since cleaning is not possible at Weert. These trains are called DCR-trains (Daily Check-up and Interior Cleaning). Since this characteristic of the train is not included in the data, human planners have to notice these by themselves. HIP is not able to notice DCR-trains and therefore does not allocate services to these trains resulting in dirty and unchecked trains leaving the node and decreasing the KPI's of the node. (10)

Technical limitations

Unfortunately, the experiments also provided a lot of technical aspects: differences not caused by human aspects, but caused because of technical limitations of HIP. Some of these technical aspects are already mentioned in the limitations of the experiment (Section 4.3). The limitations that caused differences are:

- Another technical aspect that caused many differences has to do with the service allocation part of the complete TUSP model. The model is not specified to the exact planning assignment of the VB planners, which means that the service allocation is too specific compared to the human planning. VB planners offer trains for service at a SL, but the specific services are not planned. It is for the last-minute dispatchers to execute the exact service allocation. For V&B planners, HIP is too detailed with its exact service allocation. (2, 35, 45)
- HIP is not able to change the arriving and departure platform for arriving or departing trains. Planners have the option to do so, taking into account strict rules. (30)
- A departure time is a hard constraint for human planners. Within HIP, this constraint is soft, meaning that it is possible for a train to departure late. (31)
- The services that need to be executed to a train unit are a soft constraint for human planners. If a train needs to depart but is still being serviced, this service is ended earlier (or completely skipped) to make sure the train is ready to depart. Within HIP, this constraint is hard, meaning that the train needs to be completely serviced before departure. (32)
- There is a continuing sprinter service in the timetable that changes train numbers at Eindhoven. Within HIP, this train is seen as an arriving train and a departing train, which means that HIP needs to match the arriving train with the departing train. In most cases, this is performed all right if the train consists of the same rolling stock types. However, if the incoming train is of a different combination compared to the outgoing train (because one train unit is left at Eindhoven), HIP does not match the train units correctly. (9)
- Some trains need to park for a short time on tracks where this is not allowed. Within HIP, this parking function will be notified as a conflict, while in real life this is necessary, and no conflicts will be identified for short parking. (47,48)
- Within the plan of HIP, some trains are serviced at tracks where this service cannot be applied because there are no resources at these tracks. These conflicts only occurred with one type of service that is needed for a new kind of sprinter train (Flirt). These conflicts are not noticed in human plans. (56)

Simplification of planning assignment

Human planners prefer to place trains of the same rolling stock type (same rolling stock and amount of carriages) on the same track. By doing this, it is easier for them to swap trains later during the planning process. An example: if train unit A is locked-in by train unit B, and A and B are the same rolling stock type and length, it is easy to swap the train numbers on these trains. If this is not the case and train unit B is another type of rolling stock, this is not possible and moving train unit B is necessary to free A. Next to this simplification, parking trains by rolling stock type is also beneficial during disruptions, if a train unit of a certain rolling stock type is unexpected necessary. (3)

Different types of rolling stock need different types of services. Since not all services can be applied at all the tracks, planners park trains based on their rolling stock characteristics. In other words, cleaning an ICM-3 (rolling stock type) can only be performed at track X. These limitations help planners to simplify the plan assignment since they are only focused on the tracks where a rolling stock type can be serviced instead of all the tracks. HIP does not know the exact rolling stock limitations and therefore places rolling stock on tracks where not all service tasks can be executed to the relevant rolling stock type. (15, 17)

All the human plans contained fewer movements compared to the HIP plans. Human planners minimize the number of movements because the fewer movements there are in a plan, the simpler the plan is, which is indirectly a simplification of the planning assignment. On the longer term, this also increases the flexibility of the plan during the operational phase. Besides this, it also reduces the tasks for train drivers, meaning that it has a positive effect on the capacity: there is more "train driver" capacity available. (21, 37)

If trains are matched in a way that they are parked for a long time at the shunting yard, human planners prefer to park these trains on certain tracks to minimize problems for other trains or movements. In the operational phase, this minimizes the number of movements of these long-stay trains. Within the planning of HIP, these trains are moved a lot over the emplacement, increasing the number of movements and the tasks for train drivers. (27, 39)

Human planners also prefer first to fill the tracks where service resources are available, before parking trains at tracks where no services can be executed. First of all, this eases the solving of the service allocation, since trains are already on service tracks. Next to this, fewer movements are necessary because trains need to be moved to a service location, which increases the availability of train drivers. (11, 42)

Since empty rolling stock does not need to stop at a platform for boarding or deboarding passengers, planners prefer not to plan these trains to platform tracks. Platform tracks are normally highly utilized, and every train that does not need to be routed over these tracks decreases the pressure on these parts of the rail network. In regards to this, planners will also not plan empty train towards parking or servicing tracks that are easily accessible from or to platform tracks. By not performing these two tasks, the planning assignment is easier to solve. In the operational phase, there will be more flexibility since there is less pressure on highly utilized parts of the network. (52, 55)

If parking or service tracks are accessible immediately from the main track (no saw movement needed) and can be reached from both sides, these tracks are preferred for empty rolling stock that arrives in one direction and leaves in another. (53)

A big difference between the plans from HIP and the human plans is the fact that HIP uses one track a lot, that is (almost) never used by planners (track 46). This track is used by HIP to move trains from the back of the "garden" either to another "garden" track or via a free "garden" track to the station. Human planners only use this track for the creation of new train combinations or if a train needs to be parked for a long time (longer than the time the planner needs to plan). However, this is never the case, which means that this track is often forgotten (and not used) by human planners. Not using this track changes the method that needs to be used for parking trains at the yard (from carousel to shuffleboard). This means that the First In Last Out (FILO) method needs to be used for parking. (26, 50, 51)

Capacity optimization

Planners try to optimize the usage of parking tracks. All planners know what the optimal combination of rolling stock is for every track in order to use the maximum amount of available length. This knowledge is functional for planners with their decision where they want to park rolling stock. If a track is exactly long enough to park two VIRM VI trains (190 meters of the 200 meter available used), this track is preferred to park VIRM VI trains. (7, 20, 40, 41)

Another good example of capacity optimization is the fact that human planners try to minimize the number of movements in order to minimize tasks of train drivers (increasing the availability of train drivers). Examples of these differences are the high number of movements in the HIP plan and the fact that human planners reduce movements by combining trains, moving them and splitting them. HIP sometimes does the opposite and split trains, after which the two train units are moved to the same location. Another good example is the long parking breaks between saw movement that HIP plans. Human planners try to keep this break as short as possible, regarding the possibilities for the second movement. (22, 29)

The human plan showed multiple times the option to move two train units in opposite directions from the same track at the same time. According to the traffic safety rules, this is possible, but the planning created by HIP does not use this function at instances where this is possible. This movement empties a track in a faster time, making it available for new trains. Especially for platform tracks, this feature is advantageous. The most notable instance when this is happening is if the last train unit of an ongoing train is shunted to the other direction of the ongoing train. (44, 49)

Norms and agreements

Differences also occurred because of soft norms and agreements with other departments in the chain. An example is that trains that finish their service need some time at the platform in order to let all the passengers get off and to check the train on sleeping passengers by the train conductor. This time is not incorporated in

HIP, which results in trains being moved to the Service Location immediately after arrival at the platform. (4, 6)

Another norm that is not included in HIP is the fact that employees of the SL (that execute the task) are active (or present) during specific periods. For example, the cleaning times are present at the Service Location from 8 PM to 6 AM. This means that interior cleaning can only be executed within these periods. (8)

Real-life planners do not use any tracks that are not predetermined for their use. At Eindhoven, there is also a technical centre, but V&B planners are not allowed to use any of these tracks for their plan. the HIP plan used from time to time these tracks to park trains and to execute services. If essential tracks are out of order, some of the tracks that are normally predetermined for third operators can be used (by agreement) by the NS. (28, 33, 59, 61)

Data limitations

Some data was also wrong in HIP, which caused differences between human plans and the plans created by HIP. Some differences were expected (see Section 4.3), but new differences were identified as well. These limitations are related to the following data:

- The missing of third party movements like freight trains etc. The lack of this information results in a larger solution space for HIP, while in real life these were not possible. (5)
- Time to change directions for certain train types. This is the time that is set for train drivers to walk to the other side of the train and to set the driving direction to the opposite direction. (18)
- In real life, all train drivers are allowed to drive on CBG. A specific amount of train drivers is only available to drive on NCBG. For all the shunt movements, train drivers are used that are able to drive on both. However, it is possible for planners to plan empty trains directly to shunting tracks if they are CBG. This has advantages because these trains can be parked with one direct movement to the shunting yard (fewer movements needed) and not additional train driver is needed (increased capacity of train drivers). HIP does not contain information about the CBG and NCBG and neither about the authorisations of train drivers. (25, 34)
- The routes of ongoing trains differ between the human plan and the HIP plan. Some of the routes of HIP make use of the shunting yard, causing conflicts. Since the information of the normal routes cannot be extracted from the Donna Database, the data that is inserted in HIP consist of the begin and the endpoint of the route. HIP determines the route itself, which is sometimes using shunting tracks. In real life, this would never happen, since shunting tracks are reserved for shunting movements and not for movements of ongoing trains. (38)
- The correct crossing times and follow up times are not yet implemented into HIP. There are general rules, but these do not apply everywhere, providing more freedom for human planners than for HIP. This means that in the human plan, routes can cross each other earlier compared to the plan created by HIP. Next to this, tracks are reserved for (ongoing) movements for a long time in HIP, resulting in fewer options for movements in HIP. (43, 46)

Safety

In real life, it is possible to change directions for a saw movement on certain tracks if the train consists of only one train unit because train drivers can walk interior from one cabin to the other one. Since train drivers do not need to exit the train, no specified walking path is needed. This specific freedom of movement is not implemented in HIP. (58)

Some tracks are used in the HIP plan to park or change directions of trains where this is not allowed due to safety constraints. There are no paths next to the tracks, so the train driver is not able to reach or leave the train in a safe manner. In the human plan, a parked train never occurred at these tracks. (12, 13, 35, 54)

HIP might bring multiple trains to the same platform if services are starting up in the early morning. As long as the traffic safety conditions are reached, this is allowed. However, this option is not preferred by planners because they cannot plan the exact location where a train needs to be parked at the platform. This means that it might be possible that either train drivers have to exit trains at locations where this is not safe, or in a worse case, passengers have to enter trains that have not fully entered the platform. This might result in dangerous situations, which is unwanted. Next to this, passengers have to walk longer and might miss their connection because their train is at the end of a platform. Indirectly, not using this option increases the KPI's for passenger movements. (14)

6
Extending HIP

This section describes the process of creating extensions for the differences and aspects that were identified during the experiment with planners and categorized into aspect groups. Together with the developing team of HIP, extensions are proposed. These propositions are described in Section 6.1. Due to time constraints, a selection had to be made regarding the implementation of the proposed extensions. The determination process which aspects are implemented is described in Section 6.2. This selection of extensions is implemented to the (algorithm of the) HIP. This implementation process is described in Section 6.3. Every extension affects the performance of HIP, and these differences are mentioned, and reasons are presented why these differences occur. By proposing extensions, choosing extensions to implement and implementing them, HIP is changed based on the aspects identified by the experiments. Testing the adjusted HIP is described in the next Chapter.

6.1. Defining extensions for aspects

The list of differences and the list of aspects that are missing or misinterpreted in HIP provides a good base for the extensions of HIP. The process of defining extensions and the implementation is made visible in Figure 6.1. Visible is the list of differences and the list with aspects as input data to start this process. Based on these two input sources, a selection of differences is identified that can be solved using extensions. The selection was made based on the number of appearances of differences over all experiments and the level of importance motivated by the planners. The list of differences is visible in Table 6.1. In cooperation with the development team of HIP, sessions were held in order to propose extensions that can be implemented to adjust HIP based on the differences. Differences with a low priority were of less interest compared to differences and aspects that had a high priority. Because of this, certain differences and aspects did not have extensions proposed.

The extensions that were proposed could be of different forms. These forms can be categorized into three different groups: extensions that adjust the constraints of the problem, extensions that adjust the objective function of the problem and extensions that adjust the input of the algorithm. A choice has been made to add as many extensions as possible concerning the time available, meaning that extensions do not need to be too hard to implement or need to take up much time to implement.

Within this section, the proposed extensions are explained per aspect category. Table 6.1 presents the differences and aspects, the level of prioritization and the proposed extension. The number before the next paragraphs indicate the difference presented in Table 6.1. The relation between the extensions and the aspect categories can be found in the list above the table. The table also presents extra information that is explained in Section 6.2.

6.1.1. Simplification of the future plan assignment

Difference 1: Planners preferred place trade the same rolling stock type on the same track. By implementing parking weights, the algorithm can optimize the parking spots for different kind of rotating stock types as preferred by human planners. Based on the results of the six experiments performed, different costs can

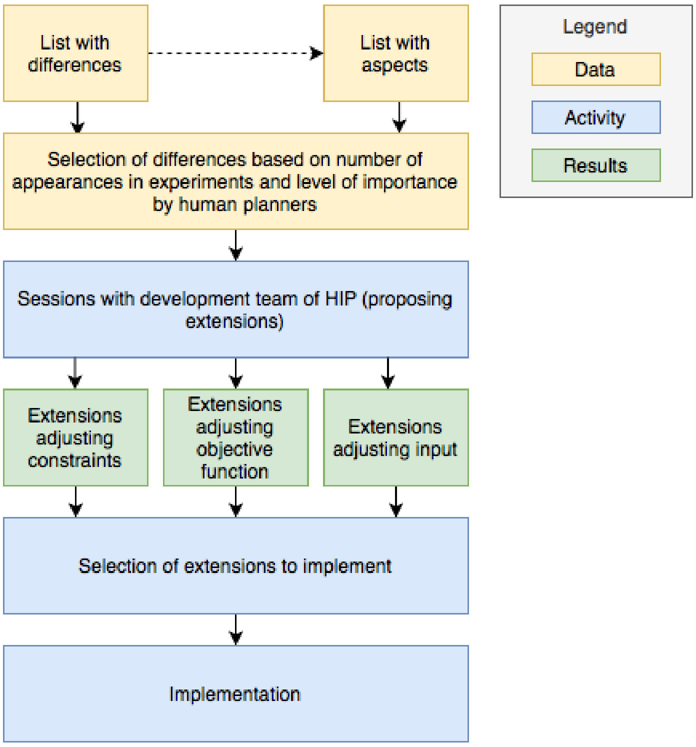

Figure 6.1: The process of defining, selection and implementing extensions

be assigned for different tracks per rolling stock type. Including the parking weights positively in the penalty function will steer the algorithm to place certain types of rolling stock on preferred tracks.

Difference 2: Since different types of rolling stock need different types of services, planners also prefer to place different types of rolling stock at tracks where the services can be executed. Just as above, the addition of parking weights to the optimization function will increase the correct placement off rolling stock at tracks where services can be executed.

Difference 3: The large number of movements in the plan of HIP can be reduced by shifting the cost function of the number of movements from the penalty cost function towards the primary optimization function. However, the amount of movement can also be reduced by allowing the extra coupling moving and splitting, which is sometimes used in real planning solutions. This is, however, not possible withing HIP. Since the amount of movements is highly dependable on a lot of other factors, there is no straight forward solution to tackle this problem.

Difference 4: The fact that long-staying trains need to be placed at specific tracks in the shunting yards is location-specific as well as planner-specific. Next to this, these trains have to be identified by the planning assistance tool. Since the experiments focused only on one single night, long-staying trains could not be identified. Because of the missing of long-staying trains, no further extension is proposed to remove this problem.

Difference 5: In order to fill tracks where services can be applied before tracks where no services can be applied, the parking weight function can also be used. By preferring tracks where services can be applied above tracks where no services can be applied HIP will automatically fill tracks with service resources earlier then tracks without service resources.

Differences 6 & 7: In order to minimize the use of platform tracks by empty rolling stock, the planning assistance tool is already on the right track. In real life, these trains are planned to tracks that are directly accessible from the free tracks but that are no platform tracks. However, in HIP, the trains cannot be planned to these tracks, since “entering and exit points” are not present at these tracks. Therefore, the choice is made to let the trains enter the node at the edges, so that HIP can directly allocate these trains to parking tracks. However, the fact that not all train drivers are allowed to drive on the shunting tracks (also see the data limitations) is not added to HIP. Therefore, some of these movements are not feasible in real life. Next to that, for the tracks that are normally used for empty rolling stock parking is not allowed in HIP, while in real life this is. By removing this constraint, these tracks might be used to park empty rolling stock.

Difference 8: Identifying empty rolling stock is hard for HIP, since this information is missing and is in real life identified by the human planner. Therefore, parking empty rolling stock at preferred parking places (or repelling it), is not possible. However, at Eindhoven, the empty rolling stock partly consists of a rolling stock type that is unique at this location. Adding preferred parking tracks for these trains by using parking weights can be used to solve this problem partially.

Difference 9: The last problem is the fact that the garden part of the service location is used differently by HIP compared to human planners. This is partly due to the high utilization of one specific track by HIP compared to humans. Figure 6.2 represents the garden and indicates the track with a red arrow. By removing this track, HIP is forced to use another method to fill the garden area, since this area is being changed from a carousel type of shunting yard to a shuffleboard type of shunting yard. This last type requires the First In Last Out method (FILO) to efficiently fill this yard, which is also the method used by human planners.

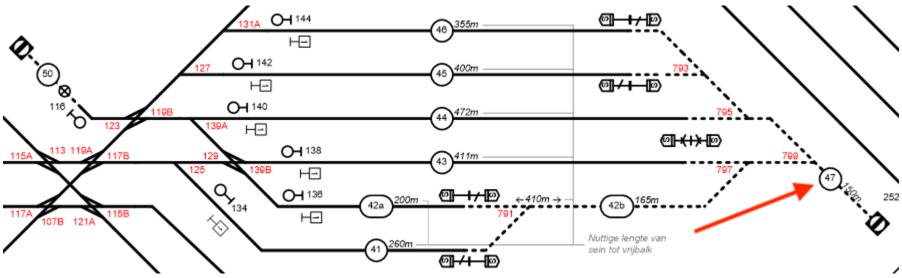

Figure 6.2: The garden yard, where the specific track (47) that is removed is indicated by the red arrow.

6.1.2. Capacity optimization

Difference 10: In order to maximize the usage of parking tracks, certain combinations of rolling stock types are preferred by human planners. These combinations were missing at the solution of HIP, because the demand never exceeded the parking capacity. However, planners prefer to always optimize the usage of tracks in a way that the length of the tracks is optimally used considering parked trains. If a train of 100 meter needs to be parked, planners prefer to park this train on a track that is a little bit longer, but not extremely long in order to keep the occupancy of the track as high as possible. In order to partially implement this rule to HIP, parking costs are added to certain tracks if the optimal combination of rolling stock is parked at these tracks.

Differences 11 & 12: The wish to minimize the number of tasks and the duration of tasks for the train drivers and other personnel is something that is not yet taken into account since crew assignment is not yet part of HIP. If crew assignment is added to HIP, it will be likely that HIP will decrease the number of tasks and the duration of the task in order to optimize the work schedule for the crew.

Difference 13: It should be possible for HIP to shunt trains in the opposite direction of the ongoing train. However, this option has never occurred in the solutions proposed by HIP. HIP identifies conflicts if a shunting train has been rejected from an ongoing train. It means that the rejected train is in conflict and the standard procedure of HIP is to let the conflicted train wait. The developing team of HIP is trying to neglect these conflicts if the train number of the ongoing train and the rejected train is the same. This adaptation takes quite some time since it is complicated to adapt this in the code, and the prioritization is low.

6.1.3. Increasing flexibility

Difference 14: In the early morning, planners try to optimize the duration a train is offered at a service track before it is moved to a platform track so that it can be used for moving passengers. This optimal value is highly dependent on the situation and the planner, and therefore no standard value can be assigned. However, it is possible to prefer the service track a little bit more with the use of the parking weights. Since these parking weights are time-dependent, it is more likely that the adapted HIP will park trains for a long time at the service tracks compared to the original HIP.

Difference 15: Difference 11 & 12: During the day, planners prefer to park different types of rolling stock on different tracks, so that all types can be brought to the station in one movement if needed (no inclusion of rolling stock types). However, this aspect only applies during the day, and not all planners focused on these points since it is officially out of their scope. Because of this, the prioritization is low.

Difference 16: HIP parks many trains at essential saw tracks, meaning that complete parts of the shunting yard are blocked. By not allowing the parking of trains at these tracks, this problem can be solved.

Difference 17: Keeping tracks completely free if possible is a wish of some planners. Since all planners do not use this method, no adaptations are proposed for HIP. An extensions that could be taken is to reward the algorithm if a train is parked at an already occupied track.

6.1.4. Optimizing plan for KPI's

Difference 18: Since the current data from Donna does not contain information whether a train only needs to be serviced at the SL (or has been serviced already at another SL), it is currently impossible for HIP to perform different actions for DCR trains. Identification by HIP is unwanted because the likelihood of misidentifying is high. The good news is that Donna will soon contain information about the DCR trains, meaning that special labels can be given to these trains in HIP.

6.1.5. Norms and agreements

Difference 19: Adding times for passengers to get off terminating trains and times to get on starting trains within HIP is possible since these two actions are included in HIP. The only problem is that no official value is given for the time needed to board or unload passengers.

Difference 20: Within HIP, the running times for certain services can be set. In other words, the times when certain services can be executed. For the planners, it is unclear what the exact times are for different types of crews (cleaning teams, mechanics, etc.). Therefore, a general indication is used for all the crew from 18:00 to 06:00.

Difference 21: Tracks that are not predetermined for the NS can be removed from HIP in order to make sure that these tracks are needed for the solution presented by HIP. Another advantage is the fact that it makes the plan assignment less complicated, optimizing the solution time of HIP.

6.1.6. Safety

Difference 22: Special rules are hard to implement. An example of such a rule is the fact that changing driving directions is allowed at all tracks if the train driver can switch cabins without leaving the train unit. Since these options are not "standard" for planners (and require much extra documentation within the plan), it is not necessary to focus on these exceptions.

Difference 23: The change of parking and driving directions allowances in HIP are easy and will be implemented.

Difference 24: Although bringing multiple trains to the platform at the start of a new service day is unwanted, it is sometimes needed or beneficial concerning crew capacity. Because of this, the option is not removed from HIP. However, the track length violations is a conflict, which means that HIP will try to solve these conflicts. Within the experiments, the running time was not long enough to exclude these conflicts out of the solution.

6.1.7. Technical limitations

Difference 25: The misallocation of services to train units is a process that is hard to implement. In real life, planners first plan which trains are parked at the service location during the night. Based on this information, they allocate service tasks to the train units. It means that the matching is performed before the tasks are allocated. HIP can swap the matching afterwards to improve the solution. It means that it is unclear which trains are parked during the night at the Service Location. Because of this, the service task allocation to incoming train units is part of the input data. To better resemble the planning process of the planner, this allocation needs to be coupled to outgoing train units. In other words, instead of assigning a cleaning task to a train unit that enters the service location, it might be better to state that an outgoing train unit has to be clean when leaving the service location. This other way of allocating services is hard to implement.

Difference 26: The fact that HIP is not able to change arrival and departure platform and human planners are, is known by the developing team. This feature will be added in a later phase.

Difference 27: The reason why the departure time is no hard constraint is because it relaxes the solution space, making it easier for HIP to find an initial solution. Making the departure time a hard constraint negatively affects the performance of HIP. Therefore it is not possible to make the departure time a hard constraint. However, it is possible to increase the weight of the departure time, increasing the necessity to solve these conflicts related to other conflicts.

Difference 28: Servicing is currently a hard constraint. It is necessary, because HIP focuses on the shunt plan, including the services. Making these constraints soft can lead to plans that do not contain services at all, meaning that the extra value of HIP compared to other planning algorithms is not significant.

Difference 29: Certain train matches are incorrect (no rational choices are made by HIP). It has to do with the fact that the train number of a train changes when calling Eindhoven. This change of train number causes problems. The fixed matching between these trains can be easily adjusted in the input of HIP.

Difference 30: Short parking is identified as a conflict. Within algorithms, clear rules need to apply, meaning that parking is parking, whether it is for 1 minute of for an entire day. Weather in real life, trains sometimes have to wait before they can continue their shunt movement, within HIP, this will be identified as conflict. Adding a short parking activity that is not penalized can solve this problem, but on the other hand, there are no real values what the maximum time of this short term parking needs to be.

Difference 31: Some services are executed at tracks where no service resources are available. This wrong combination is caused by the fact that for certain services, mobile resources can be applied, increasing the number of tracks where these services can be executed. In real life, these mobile resources are never used. Planners have not yet reached capacity problems which required the use of these devices. These devices can be removed from HIP.

6.1.8. Data limitations

Difference 32: Adding third party movements are hard since this data cannot be extracted in bulk from Donna. Since these third party movements are irregular (both in the timetable as well as the route), it is not possible to add hypothetical movements. If the third party movement data would be available, the implementation of the third party movements can be executed quickly.

Difference 33: For some train services, different times apply for the change of driving direction. These differences in times can be changed easily in HIP, since it is only a change of value.

Difference 34: The difference between centrally operated area (CBG, accessible for all train drivers) and non-centrally operated area (NCBG, accessible only for train drivers that had a specific training) is not implemented in HIP. This information needs to be added since certain trains are shunted directly to a shunting yard using train drivers that are only allowed to drive at the centrally operated area. Adding these new constraints will be difficult since it is location-specific as well as train driver-specific. However, because the crew assignment is not yet part of HIP, adding this feature does not have a high priority.

Difference 35: The exact information of the movements of ongoing trains cannot be extracted in bulk from Donna. The only way to add the specific routes is by hand. This addition is not doable within a reasonable time. The fact that many trains are using a vital shunting track can be solved, by removing the access point from the open track to this shunting track. In real life, this switch is never used, only for specific football trains, since the platform for PSV Eindhoven (soccer club) Stadium is next to this station.

Difference 36: The developing team knows the fact that crossing times are incorrect within HIP. The problem is that it is pretty unclear which norms should be handled since a lot of different rules, regulations and norms exist. There are the official norms of ProRail; there are the different sheets that planner use; but there are also the current norms of Donna (to check conflicts). However, these norms are sometimes outdated. This problem is known in the industry, but no agreement has yet been made which norm is leading. Therefore, the change of norms is postponed until the industry has made a decision regarding the different norms.

Difference 37: The fact that tracks are reserved for the entire time a train moves over this route is something that is also known by the development team of HIP. However, it is challenging to change this system to a reservation of tracks that trains need to pass, because it would include the estimated speed of trains, which are hard to predict around a node. The developing team is researching the method that Donna is using to identify and calculate the position of trains at certain times in order to implement it to HIP.

6.2. Selection of extensions to implement in HIP

The list of proposed extensions is long, so specific extensions are chosen to be implemented. This process was executed after the extensions were proposed in cooperation with the development team of HIP. First of all, based on the prioritization, the extensions are proposed. If the prioritization is low, there was no real need to spend much time proposing extensions, resulting in some differences and aspects that have no proposed extension. If the priority is of a medium level, it would be nice to have the difference solved. The high priority differences and aspects resemble the need to have differences and aspects.

Based on the proposed extension, the level of difficulty of implementation was decided within the development group. Extensions could be easy to implement, medium-hard to implement or hard to implement. The level was based on different aspects like the difficulty of coding, the complexity of the extension, the availability of data and the difficulty of adapting the overall structure of the algorithm.

With the priority and the level of difficulty known, the team was able to determine which extensions could be implemented by the researcher. The last factor that played a role was the fact that the extension was already planned to be implemented soon. If this was the case, implementation of the researcher was not necessary in order to minimize double work.

The proposals of extensions might not be the best option to solve the difference. It might be possible that other extensions are better compared to the chosen ones, but they are more challenging to implement or not yet possible because of the current structure of the algorithm or the lack of data.

Based on the proposed extension, which might not yet be the best solution to the difference, the team identified if further improvement is necessary to the extension. It might be the case that the proposed extension is simple and does not cover the whole problem. However, since the other options are more challenging to implement or at that moment not possible, this extension is proposed as the best option for that problem within this circumstance. Other improvements might be the fact that values used need some verification, or that more research is necessary to data.

Table 6.1 summarizes the process described above and the results. The level of priority is given using three scales: high, medium or low. The type of the proposed extension is provided as well. An *I* represent an extension that changes the input of the model, a *C* stands for extensions that changes constraint(s) of the model. The *OF* represents extensions that change the objective function of the model. The letter *T* stands for a technical change of the model, meaning that it affects the existing structure of the model. The last three columns show if an extension is chosen to be implemented and if an improvement of the extension is possible in order to capture the difference better. The last column presents whether the R&D team is already planning the implementation of the extension. In that case, the extension is not implemented in this research.

Table 6.1: Table containing the difference, the priority (H = high, M = medium, L = low), the proposed extension(s), the type of extension (*I* = input, *C* = constraint changes, *OF* = objective function changes and *T* = technical), the difficulty of applying (same as priority, using high, medium and low), the fact if the extension is chosen to be implemented, whether improvement is possible of the implementation and if the implementation is not yet added, if the R&D team is planning on doing so. The colors represent yes (green) or no (red).

#	Difference	Priority	Proposed extension(s)	Type	Difficulty	Chosen	Improvement possible	On planning
1	Place rolling stock types and forms on the same track	M	Parking weights	OF	M	Yes	Yes	No
2	Park trains at specific service locations (based on rolling stock types)	H	Parking weights	OF	M	Yes	No	No
3	Decrease the amount of movements	H	Change of cost and penalty function	OF	L	No	-	No
4	Park long stay trains at preferred tracks	L	-	-	-	No	-	No
5	Fill service tracks first, before filling normal parking tracks	H	Parking weights	OF	M	Yes	No	No
6	Decrease routes empty rolling stock that are using platform tracks	H	Change of parking constraints	C	L	Yes	Yes	No
7	Decrease parking of empty rolling stock at certain shunting tracks	M	Parking weights ‘	OF	M	Yes	Yes	No
8	Increase parking of empty rolling stock at certain tracks, if arriving and departure directions differ	M	Change of parking constraints	C	L	Yes	Yes	No
9	Change parking method for garden shunting area	H	Change of driving directions constraints	C	L	Yes	No	No
10	Fill tracks with regard to optimum usage	M	Parking weights	OF	M	Yes	Yes	No
11	Decrease the amount of tasks for train drivers	L	-	-	-	No	-	Yes
12	Decrease the length of tasks of train drivers	L	-	-	-	No	-	Yes
13	Move trains in opposite directions at platform tracks	H	Neglect conflicts with same train number	T	M	No	-	Yes
14	Maximize parking on service tracks at service location	L	Parking weights	OF	M	Yes	Yes	No
15	Park different types of rolling stock at different tracks	L	(Parking weights)	OF	M	Yes	-	No
16	Do not park trains at tracks that are used for saw movements	H	Change of parking constraints	C	L	Yes	No	No
17	Keep tracks free if possible	L	Reward parking of trains at already occupied tracks	OF	M	No	-	No
18	Identify DCR trains and assign tasks to them	L	-	-	-	No	-	No
19	Add times to let passengers leave the train before shunting towards yard	H	Provide duration times for boarding and deboarding	I	L	No	-	No
20	Include opening times for SL (when crew is present at SL)	H	Add opening times for SL	C	L	Yes	No	No
21	Only use tracks that are predetermined for NS	M	Remove tracks in layout	I	M	Yes	No	No

Table 6.1: Table containing the difference, the priority (H = high, M = medium, L = low), the proposed extension(s), the type of extension (*I* = input, *C* = constraint changes, *OF* = objective function changes and *T* = technical), the difficulty of applying (same as priority, using high, medium and low), the fact if the extension is chosen to be implemented, whether improvement is possible of the implementation and if the implementation is not yet added, if the R&D team is planning on doing so. The colors represent yes (green) or no (red).

#	Difference	Priority	Proposed extension(s)	Type	Difficulty	Chosen	Improvement possible	On planning
22	Add option to change driving direction at tracks if train exists of only one train unit	L	-	-	-	red	-	red
23	Parking and changing directions not possible at every track	H	Change of parking constraints; change of driving directions constraints	C	L	green	red	red
24	Exceeding of platform length by multiple trains in the beginning of the morning	L	This conflict will be removed as HIP runs longer	T	L	green	red	red
25	Misallocating of services to train units	L	Change assignment of services; add now manually	I	H	green	green	red
26	No changes possible in arrival or departure platform	M	Add option to HIP	T	H	red	-	green
27	Departure time is currently a soft constraint in HIP	H	Change into hard constraint (unwanted)	T	H	red	-	red
28	Servicing is currently a hard constraint in HIP	H	Change into soft constraint (unwanted)	T	M	red	-	red
29	Some train matches need to be set	H	Add fixed matches to input	I	L	green	-	red
30	Short parking is identified as conflict	L	Add short parking activity	T	M	red	-	red
31	Some services are executed at tracks where service resources are not present	H	Remove mobile resources	I	L	green	red	red
32	Third party movements are not present in HIP	H	Add if data is available	I	L	red	-	red
33	Time to change driving directions for rolling stock are incorrect	M	Change driving direction times	I	L	green	red	red
34	Information about permissions of train drivers and CBG and NCBG are missing or incorrect	L	Add constraints to tracks and crew member	C	H	red	-	red
35	Routes of ongoing trains are incorrect in HIP	M	Remove switch in layout	I	M	green	green	red
36	Crossing times and follow-up times are incorrect	M	Add correct norms if known	I	M	red	-	green
37	Reservation of tracks for movements decreases the amount of possibilities related to human plans	L	Use predicted train position instead of reservation	T	H	red	-	green

In general, the table presents the different levels of priority and the difficulty of the proposed extensions. These insights are presented in Figure 6.3. From the 37 differences and aspects, 13 were identified with a low priority level, 9 with a medium priority level and 15 with a high priority level. Because of certain low priority levels, 32 extensions are proposed. From these 32 extensions, 13 were identified with a low difficulty level, 14 with a medium difficulty level and 5 with a high difficulty level.

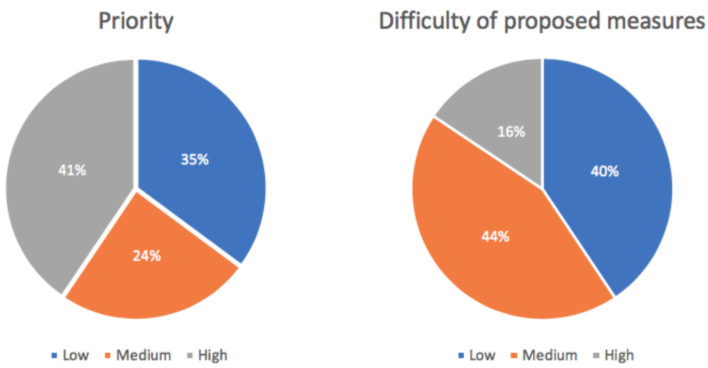

Figure 6.3: General information about the priority and the difficulty of the proposed extensions

Next to the general information gathered, it is also interesting what are the characteristics of the extensions chosen to be implemented. Figure 6.4 presents these characteristics. First of all, 54% of the extensions was chosen to be implemented (20 out of the 37). From the chosen group, 50% of the extensions are identified as perfect for the difference that it needs to solve. Eight proposed extensions need further improvement because the extension will probably not cover the complete difference. Of two extensions it is unclear if improvement of the extensions is necessary or not. For the first one, number 15 of the table (the need to park different types of rolling stock at different tracks), parking weights can be used as an indirect extension. However, this extension does not steer the algorithm to actively park same types of rolling stock at the same track. The second extension (29 in the table) where no insights can be provided about the need for improvement is the fact that certain train matches need to be set. The reason behind this uncertainty is the fact that this difference is node-specific, and therefore no general information can be provided.

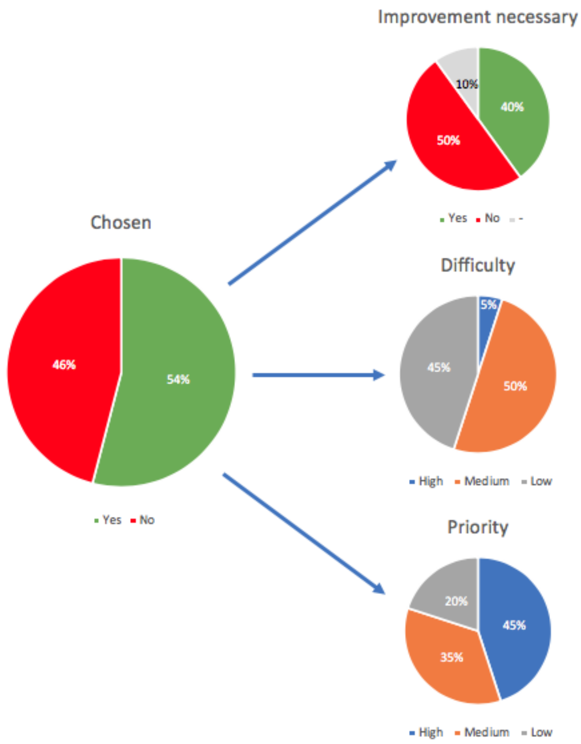

Figure 6.4: Information about the chosen extensions to be implemented.

From the chosen extensions, the number of difficult extensions is 1, the number of medium-difficult extensions is 10, and the level of easy extensions is 9. Relating the chosen extension to the priority level of the differences and aspects, it can be stated that there is a clear focus on the high priority differences and aspects with 9 of the 20 extensions. The group of medium priority differences and aspects that is chosen is 7, and the amount of low priority differences and aspects is relatively low with 4. Based on these outcomes, it is clear that the choice procedure succeeded in two ways: focus on the high priority differences and aspects and a large amount of simple or medium-difficult implementations, that however still do need some improvement. Based on this choice process, a clear list of extensions can be implemented. A description of the implementation of these extensions is provided in the next Section.

6.3. Implementation of extensions in HIP

The choice process results in a clear list of extensions that need to be implemented. This section will describe how the extensions are added or changed in HIP. The extensions are explained using the type categories: extensions adjusting the constraints of HIP; extensions adjusting the objective function of HIP, extensions adjusting the input of HIP. Since none of the technical differences to HIP are chosen, this category has no content and therefore is not discussed in this section. In order to understand the exact technical changes to the algorithm, certain links are made to terms explained in Section 3.2.

6.3.1. Extensions adjusting the constraints of HIP

The points that will be adjusted to the constraint part of HIP are the following: the change of parking constraints; the change of driving directions constraints and the addition of opening times for the Service Location. These three adjustments have a direct effect of the solution space since they reduce or increase the space with the constraints. These adjustments are necessary in order to make sure the solution space of HIP is the same as the solution space of human planners.

The adjustments to the parking and possibility to change driving direction at certain tracks is presented together since these implementations are almost the same. The constraints are stated in the location file, where for every track it is stated if parking is allowed and if it is possible to change driving direction at the track. An example will be: for track 44, it is possible to park, but it is not allowed to change driving direction. This information can easily be adjusted in the algorithm by changing a "Yes" to a "No" or vice-versa. By doing this, it is possible to adjust the rules per track. According to the differences, multiple changes are made regarding the parking rules and change of driving direction rules.

It is essential to state that the rules are based on human planners. The rules that planners are applying exist of the real rules (from ProRail, the infrastructure manager) rules between planners and other departments or employees, and certain rules that simplify the planning assignment. An example of this is the fact that although it is possible to park at a certain track on a shunting yard, it is not preferred at all by planners since walking times from and to this track are irrational long. These long waling times result in a fact that train drivers need to walk a lot, which is generally not preferred by train drivers, and that train drivers are occupied with one task for a long time, which affects the productivity. Next to the fact that rules are made between planners that the usage of this specific track is unwanted because of the inefficiency of train drivers, rules between planners and train drivers are made that the usage of this track is minimized as much as possible. Since the usage of these kinds of tracks is not notified in any of the human plans in the experiment, the decision is made to adjust the constraints of these specific tracks.

The removal of parking possibilities and the possibility to change the driving direction decreases the solution space in general since more rules are added to the space meaning that fewer options are possible to park a train or change the driving direction of a train. The decrease of solution space also has an impact on the complexity of the planning assignment, since a decrease of options directly affects the complexity of choices made by the algorithm. Instead of having 40 tracks where trains can be parked, the number of tracks decreased to 35. For the local search approach that is used to optimize the solution, a decrease of options means that fewer candidates have to be checked. Instead of 40 options in the original HIP, only 35 options are possible in the adjusted HIP.

The addition of the opening times of the SL is performed by adjusting the time of the availability of resources needed for service. For example, a service platform is needed for the interior cleaning of trains. For the original HIP, the availability of the service platform was 24/7, meaning that the service platform could be used every moment of the day. By adjusting the availability of these resources to the evening hours and night, it is possible to add real opening times to the Service Locations. It means that trains can already be parked next to the service platform, but that the interior cleaning starts when the service platform is available. This constraint results again in fewer options a service can be executed, decreasing the number of options.

6.3.2. Extension adjusting the objective function of HIP

The only extension that is added that adjusts the objective function of HIP is the addition of parking weights. The addition of parking weights solves a lot of differences and aspects, although some differences and aspects are partially solved. Although this extension affects the optimization function because an extra element is added (the preference track rewards), it is still possible to identify the differences between the original HIP and the adjusted HIP. It is necessary to compare the original HIP with the adjusted HIP in the next Chapter.

The parking weights are determined based on the information of planners provided during the experiments. In Table 6.2, the parking tracks are presented in the first column. The rows present the different types of rolling stock present at the node. The fact that the parking weight can differ per rolling stock increases the effectiveness of this extension since preferences between rolling stock types can be captured as well with the parking weights. The relation between the resources at certain tracks (parking, interior cleaning, technical check-up, etc.) and the rolling stock types are leading when deciding the weights for different tracks and rolling stock types. The higher the number in the Table, the more preferred the type of rolling stock is for the specific track. An example: Flirt trains need to undergo a particular service to empty the wastewater tanks. This service is only possible at Eindhoven at tracks 11, 12, 131 and 132. Table 6.2 shows that the combination between the Flirt types and these specific tracks score high compared to other tracks. In real life, planners only plan Flirt trains at these tracks and therefore reserve these tracks for this particular train type. This is also made visible since for almost all these tracks (11, 12 and 132), no weights are given if a VIRM type or an ICD type is parked at these tracks. For these train types, other tracks are preferred. For the train types SLT and ICM, no real preference tracks are indicated based on the experiments. However, certain combinations of trains are preferred by planners as well. A good example visible in the Table is the combination of an ICD train and a SLT4 train. These unique combinations are also added to the parking weights.

The weights are influencing the penalty cost as follows. First of all, the amount of seconds a train is parked at a specific track is multiplied by the parking weight. The addition of all these values is multiplied by 0,0002 and subtracted from the penalty score. It means that it is possible to have a negative penalty score if the parking weights value is larger than the cumulative values of the number of movements and the parking time violation.

6.3.3. Extensions adjusting the input of HIP

The extensions adjusting the input of HIP are the following:

- manual assignment of services;
- addition of certain fixed matches to input;
- the change of driving direction times;
- the removal of mobile resources;
- the removal of tracks in the layout of the node
- and the removal of a switch in the layout

These extensions can be grouped into two groups concerning the method of implementation. For the scenario-specific input, like the manual assignment of services and the addition of specific fixed matches to the scenario, the Graphical User Interface (GUI) was used. It is the most straightforward and most user-friendly method to change these scenario-specific details. Within the GUI, it is possible to assign services to train units and to add fixed matches to the planning assignment. Since these two extensions are scenario (or planning assignment) specific, it is not possible to implement these changes within the algorithm. Next to this, both the assignment of services as the addition of fixed matches have to be performed manually, since the identification of problems where assignments of services or matches are wrong have to be notified by a human.

Table 6.2: Parking weights of Eindhoven per rolling stock type. The weights are based on the experiments performed.

Tracks	Rolling stock types							
	VIRM IV	VIRM VI	ICD	FLIRT 3	FLIRT 4	SLT 4	SLT 6	ICM 3
11	0	0	0	1	1	1	1	1
12	0	0	0	1	1	1	1	1
13	1	1	0	0	0	1	1	1
14	1	1	0	0	0	2	2	1
129	1	1	0	0	0	1	1	1
130	1	1	0	0	0	1	1	1
131	0	0	0	1	1	1	1	1
132	0	0	0	1	1	1	1	1
41	0	0	0	1	1	0	0	0
42	1	2	0	1	1	1	1	1
43	0	0	2	1	1	1,25	1,25	0
44	1	1	0	1	1	1	1	1
45	1	1	1	1	1	1,25	1,25	0
46	0	1	0	1	1	0	0	0
15	0	0	0	0	0	1	1	1
16	0	0	0	0	0	1	1	1
17	0	0	0	0	0	1	1	1

The second group is consist of the extensions that can be implemented in the primary input of the algorithm since this information is not scenario-specific but general for all planning assignments. The group consist of the last four extensions of the list presented above. For the change of driving directions time, the information about rolling stock types had to be adjusted. The last three extensions (the removal of mobile resources, the removal of tracks and the removal of a switch) were changed in the input information concerning the layout of the node. These adjustments have a direct impact on the initial solution space since this information describes the boundaries of the solution space (the number of total tracks, the number of switches for routes, etc.).

7
Evaluation of adjusted HIP

This Chapter answers the last sub-question: **What is the impact of the integration of missing aspects to the Planning Assistance Tool (HIP in this case)?** In order to answer this question, a comparison is executed using a real-life planning assignment and a theoretical planning assignment. The three comparisons that took place are presented in Figure 7.1. The yellow arrows represent the comparisons that took place using real-life planning assignments. The blue arrow represent the comparison that took place using theoretical planning assignments. A comparison between a human plan and the adjusted HIP solution on real-life planning assignments is used to capture the qualitative differences. With the help of the same planners used for the first experiment, the solution of the adjusted HIP for a real-life planning assignment is compared to a human planning. This experiment provides answers to which extent the implementations are improving HIP and its solution from a planners perspective.

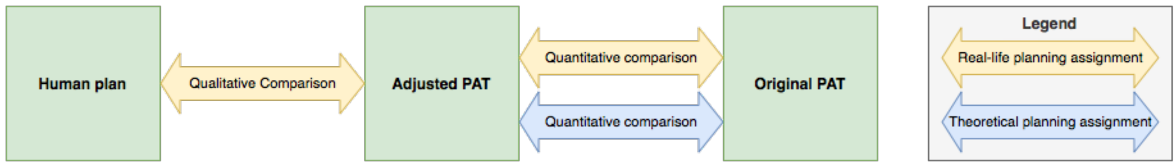

Figure 7.1: Explanation of the three comparisons discussed in this Chapter. The comparison between the human plan and the adjusted PAT is based on real-life planning assignments and has to goal to capture the qualitative differences. The two comparisons between the adjusted HIP and the original HIP have the goal to capture the quantitative differences based on real-life planning assignments and theoretical planning assignments.

With a real-life planning assignment a quantitative comparison is performed between the adjusted HIP and the original HIP. In this comparison, the solution of the adjusted HIP is compared to the solution of the original HIP considering real-life planning assignments. This comparison provides information in what way the implementation affect the quantitative indicators of the solution. The goal is to understand if the adjusted HIP performs better or worse compared to the original HIP.

Next to real-life planning assignments, the last experiment that is executed is based on theoretical planning assignments. These fictional planning assignments differ with each other in a level of complexity. Again, a comparison between the adjusted HIP and the original HIP is be performed in order to find out how the adjusted HIP performs regarding planning assignments with a different level of complexity.

This chapter starts with the introduction of the quantitative indicators in Section 7.1. The explanation of the indicators is followed by comparisons based on real-life planning assignments in section 7.2, containing both the qualitative as the quantitative comparison. Next, the comparisons based on the theoretical planning assignments are described in Section 7.3. The Chapter concludes with Section 7.4, where the answer on the sub-question described above is given.

7.1. Indicators

In order to compare different HIP plans with each other, indicators are necessary. These indicators are all quantitative. The quantitative indicators can be grouped into three different groups: conflict costs, penalty costs and other indicators. The first two groups describe an indicator and certain sub-indicators. This means that the sub indicators have a direct relation to the value of the main indicator (conflict or penalty costs). The three groups of indicators are explained underneath.

7.1.1. Conflict costs

The **conflict costs** indicates the costs of conflicts of the plan. These costs are minimized by the local search optimization. If the new costs are lower compared to the original costs, the new plan is assumed to be better and the original plan is discarded. Therefore, the conflicts costs represent the primary optimization phase of the plan. If these costs are 0, the secondary optimization value is being optimized, which is the penalty costs. The cost score consist of the crossings, the delayed departures, the delayed arrivals, the track length violations, the incorrect combination and the delayed fixed trains.

The number of **crossings** indicates the number of not allowed crossings of shunting trains.

The number of **delayed departures** indicates the number of trains that depart with a delay. It only includes the trains that were parked (and serviced) at the node, no ongoing trains.

The number of **delayed arrivals** indicates the number of trains that arrived with a delay. It only includes the trains that need to park (and have service) at the node, no ongoing trains.

The **track length violations (TLV)** indicate the number of times a track length is exceeded.

The number of **incorrect combination** indicates the number of trains that depart in a wrong composition.

The number of **delayed fixed trains** indicates the amount of ongoing trains that are delayed because of the shunt plan. Ongoing trains (timetable trains) are trains that call at the station of the node, but that continue the service. These trains do not need to be parked or serviced at the node.

7.1.2. Penalty costs

The **penalty costs** represents the penalty score. The lower this score, the less penalties are present in the plan. As stated above, this penalty score is optimized as soon as the the conflict costs is zero. The penalty cost consist of the number of shunt movements and the illegal parking time.

Number of movements indicates the amount of shunt movements that are in the solution. This quantitative indicator can also be used as a qualitative indicator. A lot of planners indicated that the amount of movements of HIP plan felt high compared to their own plan. Next to that, certain movements did not have a direct goal; these movements were unnecessary. Although the real number of movements of the plan of the planners are not counted, it is clear that the amount of movements of HIP exceeds the amount of movements planned by the human planner.

Illegal parking time violation indicates the time in seconds a train is parked at a track where this is not allowed.

7.1.3. Other indicators

Candidates checked indicates the amount of candidates that are checked by HIP in order to optimize the initial plan. The higher this number, the more options HIP has checked. If the same running time is used and the number of candidates checked has increased, this is positive for the solution approach of the algorithm. In other words, the plan assignment might be a little less complex, since more candidates were able to be checked within the same time limit. The number of candidates is not part of one of the two costs functions.

7.2. Comparisons based on real-life planning assignments

Comparisons based on real-life planning assignments are conducted and the results are described in this Section. The Section starts with the qualitative comparison, where the plan of the adjusted HIP is compared to the original plan of the human planner. When the qualitative comparison is presented, it is time for the quantitative comparison. But before this is explained, it is first interesting to have a better look to the real-life planning assignments. The quantitative comparison between the original and adjusted HIP is described at last in this Section.

7.2.1. Qualitative comparison

The solutions of the real-life planning assignments created with the adjusted HIP were presented to the planners that have planned these plan assignments in real-life. They were able to compare the solution with the first solution (created by the original HIP) and with the solution that they have created (the human planner solution).

First of all, the overall opinion of the planners was that the adapted HIP produced a better solution compared to the original HIP. The word better needs to be related to two different ways of looking. First off all, the planners found less mistakes in the solution. Trains were not parked at tracks where this is not possible, there were no movements that were not allowed and all the norms and rules agreed upon with third parties were enforced. The plan contained less conflicts compared to the original solution.

The second meaning of better refers to the fact that the new solution resembles the plan of the planner more compared to the old solution. All planners said that they would agree upon a large amount of the decisions HIP made regarding the solving of the plan assignment. The fact that trains were parked according to general rules between planners and actions were clear, planners were able to understand the solution in general easier, as well as the motivation behind certain actions.

Before going more in depth, the planners all agreed upon the fact that the new solution provided a better plan for the garden area of the Service Location (for a detailed drawing, see Figure 4.2. The usage of the track at the end of the garden areas was minimized, and the trains were correctly parked according to the first in, last out (FILO) method. This method is used by the planners as well. This was a big improvement for the planners and solved a lot of side problems as well. The new solution also produced a better parking plan for the east side of the Service Location (see Figure 4.2 for a better understanding). The different rolling stock types were parked according to the norms and rules that planners have internally and externally, and logical reasoning. These parking tracks were well used for the trains that need to undergo a service, as well as for the trains that are going to be part of an empty rolling stock train, and because of this need to be parked at specific tracks in order to easily exit the Service Location without blocking the platform tracks.

In order to substantiate the points above, the list of differences (Appendix C) between the original HIP and the human plan was presented to the planners as well, and for every difference, the question was asked if the difference still occurred or if the difference was partly or even completely solved. The differences that were partially or completely solved are presented in Table 7.1 including some small remarks.

Table 7.1: This table contains the differences that are partially (status is yellow) or completely (status is green) solved within the adjustment of the PAT. Remarks are added which adjustment solves the difference between human and PAT plan. Differences that were not impacted by the adjusted PAT are not presented in this table.

#	Difference	Status	Remarks
1	HIP bring trains earlier to the station	(yellow)	Due to the fact of parking weights.
2	Rolling stock not placed at correct service track	(yellow)	Due to the fact of parking weights, forcing certain rolling stock to certain tracks.
3	Planner likes to park rolling stock types and forms on the same track	(yellow)	Due to parking weights.
7	Human planner prefer certain tracks for certain rolling stock combination in order to optimize usage of the tracks	(green)	Due to the fact that these rolling stock combinations are added to the parking weights.

Table 7.1: This table contains the differences that are partially (status is yellow) or completely (status is green) solved within the adjustment of the PAT. Remarks are added which adjustment solves the difference between human and PAT plan. Differences that were not impacted by the adjusted PAT are not presented in this table.

#	Difference	Status	Remarks
8	Human planners take into account that cleaning teams are not 24/7 present at the Service Location		Adding opening times of Service Locations.
9	The solution of HIP contains a lot of difficult (irrational) matches between train units		Solved by setting the matches of the specific train service beforehand.
10	HIP does not assign services to all trains that need to be serviced		Due to the manual assignment of these services.
11	HIP parks trains at non-service track while service tracks are available		Solved due to the fact that service tracks have a higher ranking for the parking weights.
12	Change of driving directions not possible at certain tracks		Solved due to the addition of constraints.
14	HIP brings extra trains to platform in the early morning		Due to parking weights.
15	Planner likes to park rolling stock types and forms on the same track		Due to parking weights.
16	Human planners park train units of a different rolling stock type at a track that is not yet occupied		Due to the parking weights.
17	Planner likes to park rolling stock types and forms on the same track		Due to parking weights.
18	The time to change driving directions for certain rolling stock types is incorrect		Solved due to the change of time.
20	Human planner prefer certain tracks for certain rolling stock combination in order to optimize usage of the tracks		Due to the fact that these rolling stock combinations are added to the parking weights.
21	Human plan has a lot less (saw) movements		Less movements because of parking restrictions at certain saw tracks.
23	HIP bring trains earlier to the station		Due to the fact of parking weights.
24	HIP bring trains earlier to the station		Due to the fact of parking weights.
25	Empty rolling stock trains are using the platform tracks while this is not necessary		Due to the better functioning of the empty material tracks available.
26	Some tracks are highly utilized by HIP and not by the human planner		Making these tracks less attractive for HIP.
28	HIP used track that are not assigned to the NS		Solved because these tracks are deleted from HIP.
33	HIP used track that are not assigned to the NS		Solved because these tracks are deleted from HIP.
34	Empty rolling stock trains are using the platform tracks while this is not necessary		Due to the better functioning of the empty material tracks available.
35	HIP does not assign services to all trains that need to be serviced		Due to the manual assignment of these services.
36	Change of driving directions not allowed at certain tracks		Solved due to extra constraints for the change of driving directions
37	Human plan has a lot less (saw) movements		Less movements because of parking restrictions at certain saw tracks.
38	Routes in HIP are different compared to the human plan		Solved by removing a never used switch, making the problematic routes not possible any more.

Table 7.1: This table contains the differences that are partially (status is yellow) or completely (status is green) solved within the adjustment of the PAT. Remarks are added which adjustment solves the difference between human and PAT plan. Differences that were not impacted by the adjusted PAT are not presented in this table.

#	Difference	Status	Remarks
40	Human planner prefer certain tracks for certain rolling stock combination in order to optimize usage of the tracks		Due to the fact that these rolling stock combinations are added to the parking weights.
41	Human planner prefer certain tracks for certain rolling stock combination in order to optimize usage of the tracks		Due to the fact that these rolling stock combinations are added to the parking weights.
42	HIP parks trains at non-service track while service tracks are available		Solved due to the fact that service tracks have a higher ranking for the parking weights.
44	The movement of split-off train units is planned by HIP after the ongoing train has left the station		Due to the set matching of this specific train type.
45	The solution of HIP contains a lot of difficult (irrational) matches between train units		Solved by setting the matches of the specific train service beforehand.
46	HIP does not assign services to all trains that need to be serviced		Due to the manual assignment of these services.
49	Certain tracks can be used for short parking, combining and splitting in real-life		These functions are added to these tracks.
50	The movement of split-off train units is planned by HIP after the ongoing train has left the station		Due to the set matching of this specific train type.
51	HIP does not use FILO method to fill the garden		Solved by adjusting the use of the track in the back of the garden (track 47).
52	Simple rematching is not preferred for HIP		partly solved by adjusting the use of track in the back of the garden.
53	Empty rolling stock trains are using the platform tracks while this is not necessary		Due to the better functioning of the empty material tracks available.
54	Empty rolling stock trains are using the platform tracks while this is not necessary		Due to the better functioning of the empty material tracks available.
55	Change of driving directions not possible at certain tracks		Solved due to the addition of constraints.
56	Empty rolling stock trains are using the platform tracks while this is not necessary		Due to the better functioning of the empty material tracks available.
57	HIP allows services at tracks where no service resources are available		Solved by deleting the mobile resources.
58	HIP parks trains at tracks where this is is not allowed		Solved by addition of parking constraints.

The length of the list is long, meaning that a lot of differences that were present in the original HIP were not or less present in the solution produced by the adapted HIP. However, there were also some points that drew the attention of the planners. The largest point, which is worth mentioning, is the fact that for the East Side of the Service Location, the amount of movements decreased a lot, and not all movements had a clear or direct goal. A lot of movement were performed, while other options were possible and would have been chosen by the planner. This main point, together with some smaller points, is addressed to the development team.

The reason why the amount of movements were noticed by the planners was because this was the new thing to focus on since a lot of problems experienced earlier were not present due to the measures. It might be possible that these movements were also present in the original solution, but were not noticed of not as important to mention compared to the other problems. This shows that more iterations, as described in this report, are necessary to improve the algorithm in order to resemble the human way of planning.

7.2.2. Explanation of real-life planning assignments

In order to understand the differences between the original HIP and the adjusted HIP it is needed to describe the different characteristics of the real life planning assignments used. But before identifying the differences between the planning assignments, first the main characteristics of all the planning assignments are retrieved. The planning assignments all describe 24 hours of a weekday (from Monday to Friday), starting from 08:00 till 7:59 the next day. The planning assignments did not contain trains that stayed for multiple trains: all trains that arrive will leave within the time period of the planning assignment (within the 24 hours). In general, the same types of rolling stocks are present at the node: VIRM4, VIRM6, Flirt3, Flirt4, SLT4, SLT6, ICD 9 Gvc-Ehv and ICM3. The number of train units per rolling stock type differ from planning assignment to planning assignment.

The numerical characteristics of the planning assignments are presented in table 7.2. The different planning assignments are named from A to E, where A exist out of two versions, A1 and A2, since these two planning assignments were planned by the same planner. The planning assignments are presented in the second to last rows. The indicators are presented in the first column and describe the number of train units per rolling stock type that needs to be parked (and possibly serviced) at the node per 24 hours, the amount of arriving and departing trains at the node (so no ongoing trains) and the average number of train units per arriving and departing train. The table also presents information about the number of tracks that are out of service in the planning assignment and the number of ongoing trains of NS. This number only represents the traffic of NS, so no other rail traffic is included.

Table 7.2: Quantitative characteristics of the different real-life planning assignments used.

Indicators	**A1**	**B**	**A2**	**C**	**D**	**E**
Date	2-3 Jan	5-6 Dec	30-31 Dec.	26-27 Feb	24-15 Feb	6-7 Feb
VIRM4	15	19	23	38	36	17
VIRM6	12	17	12	17	20	16
FLIRT3	80	57	79	58	69	73
FLIRT4	43	75	47	81	80	67
SLT4	1	0	0	2	1	1
SLT6	1	7	3	2	4	2
ICR Gvc-Ehv 9	36	36	36	36	36	36
ICM3	3	3	3	3	3	3
Sum of train units	191	214	203	237	249	215
Arriving trains	148	148	138	152	154	144
Train units per arriving train	1,29	1,45	1,47	1,56	1,62	1,49
Departing trains	154	156	147	161	163	148
Train units per departing train	1,24	1,37	1,38	1,47	1,53	1,45
Tracks out of service	0	0	0	13	0	0
Ongoing trains	536	534	512	547	556	511
Level of difficulty	Easy	Medium	Medium	Hard	Hard	Medium

Table 7.2 shows that certain planning assignments involve more trains than others. Especially planning assignments C and D contain more trains units that need to be parked at the node, as well as more ongoing trains compared to the other planning assignments. This higher amount of trains is logical, since these planning assignments describe days of the carnival week, which is a big event in the area around Eindhoven (C is February 26 - 27, and D is February 24 - 25). These planners both indicated that their planning assignment was harder compared to normal days. The higher number of average train units per departing or arriving train indicates that trains were longer compared to normal days. Between these two planning assignments, planning assignment C also contained 13 tracks being out of order during the night, increasing the difficulty level of the plan assignment compared to D and the rest. planning assignment A is the planning assignment with the least train units and the lowest number of train units per departing or arriving train. However, the amount of ongoing trains is relatively high. planning assignment B, A2 and E are comparable concerning the amount of train units per arriving or departing train, and the sum of the train units. From these differences indicated in the table, three groups can be created that present the different difficulty levels of planning assignments: easy (A1), medium (B, A2 and E) and hard (C and D). However, it should be noted that the planning assignments are grouped based on limited quantitative data and not on qualitative data.

7.2.3. Quantitative comparison

The quantitative comparison about the original HIP and the adjusted HIP concerning the real-life planning assignments is presented in Appendix E. The running time of HIP was, just as during the experiments, 10 minutes (600 seconds). The amount of seeds is either two (for A1 and B) or three (for the rest). The seeds resemble the amount of different solutions that could be calculated by one run. The difference in the amount of seeds is based on the fact that the results from the experiments are used for the original HIP. For the first two experiments, only two seeds were used, whereas for the last four experiments three seeds were used. For the qualitative comparison: the Graphical User Interface (GUI) automatically shows the solution that has the lowest conflict costs.

The overview of the comparison between the original HIP and the adjusted HIP on real-life planning assignments is made visible in Table 7.3. This Table presents the differences of the different planning assignments and the average over these 6 planning assignments. The differences between the original HIP and the adjusted HIP differ quite from planning assignment to planning assignment. Only for the illegal parking time violation indicator, it can be stated that a general decrease is visible over all planning assignments. Concerning the other indicators, less hard statements can be made. The conflict costs decreases generally with 19% when the adjusted HIP is used. However, these costs increase for planning assignment C and D. The same is applicable for the delayed arrivals and track length violations. The penalty costs decreases for all planning assignments except for planning assignment A2. The amount of crossings decreased for all planning assignments for the adjusted HIP, except for planning assignment C. The delayed departure decreased, except for A2 and C. The amount of TLV decreases with 11%. The adjusted HIP has a negative influence on the incorrect combinations, since these increase for all planning assignments except E. The difference on the number of movements between the original HIP and the adjusted HIP differs from planning assignment to planning assignment, but overall there is no noticeable difference. The number of candidates checked decreased for the adjusted HIP, which means the adjusted HIP investigates less candidates for the local search optimum.

Table 7.3: Average of comparison between original HIP and adjusted HIP on real-life planning assignments. The first indicator presents the number of seeds these results were based on. The six different planning assignments are presented in the columns, and the average column at the end represents the average of these six planning assignments.

Indicators	**A1**	**B**	**A2**	**C**	**D**	**E**	**Average**
Conflict costs	**-16%**	**-90%**	**-13%**	**35%**	**1%**	**-31%**	**-19%**
Crossings	-81%	-86%	-37%	18%	-6%	-65%	-43%
Delayed departures	-29%	-87%	19%	46%	-6%	-18%	-12%
Delayed arrivals	-100%	-98%	-6%	23%	3%	-55%	-39%
Track length violations (TLV)	0%	-20%	-100%	96%	20%	-60%	-11%
Incorrect combinations	0%	300%	0%	20%	20%	-100%	40%
Delayed fixed train	74%	-93%	-9%	19%	2%	-28%	-6%
Penalty costs	**262%**	**-23%**	**27%**	**-39%**	**-26%**	**-29%**	**29%**
Number of movements	-8%	1%	5%	7%	-2%	0%	0%
Illegal parking time violation	-54%	-78%	-68%	-56%	-47%	-63%	-61%
Candidates checked	-26%	12%	-1%	12%	-4%	4%	-1%
Number of seeds	2	2	3	3	3	3	N/A

No clear relations can be drawn between the level of difficulty of the planning assignments and the performance of the adjusted HIP compared to the original HIP. The group of hard planning assignments (C and D) have in common the increase of the conflict costs, an increase of the delayed arrivals and an increase of the track length violations. The use of the adjusted HIP on the low level of difficulty planning assignment (A1) shows improvements on certain indicators, but not on the penalty costs (extremely large increase) and the delayed fixed trains. Next to that, the amount of candidates checked is also a lot less compared to the results of the original HIP. For the planning assignments that are grouped as medium difficult (B, A2 and E), the amount of candidates checked increased by using the adjusted HIP. A lot of indicators show the same trends within this group. However, this holds not for the penalty costs, delayed departures and incorrect combinations, where planning assignment A2 seems to act different of the adjusted HIP compared to the other two planning assignments. A last remark must be made concerning the amount of movements, which increases or stays stable for all medium planning assignments with the adjusted HIP.

For a more graphical representation of the results, Figure 7.2 is created. The graph clearly envisions the differences between the experiments to each other and to the average line in black. The differences per planning assignment per indicator are clearly marked. This graph helps to indicate the outliers that are present. Clearly are the largest outliers of planning assignment A1 (green) and B (yellow). The graph also shows the fact that for certain indicators the differences between planning assignments are not big, like for the number of movements, the illegal parking time violation and the number of candidates checked. For the penalty costs, the planning assignments B, C, D, and E are also pretty close together. From a more general view, it is visible that the performance of planning assignment E has improved by the adjusted HIP and of planning assignment B if TLV and incorrect combinations are not taken into account.

Figure 7.2: The average performance of the adjusted HIP of the different planning assignments compared to the original HIP. The indicators are shown on the x-axis, the difference in percentage on the y-axis. The two outliers are not presented in this graph: planning assignment A1 has a penalty costs increase of 262%, planning assignment B has an increase of incorrect combinations of 300%.

7.3. Comparisons based on theoretical planning assignments

Within this Section, the adjusted HIP is tested and compared with the original HIP using theoretical planning assignments. The goal is to find out how the adjusted HIP performs compared to the original HIP. Theoretical planning assignments are created, which are described in Section 7.3.1. The theoretical planning assignments function as input for the original and adjusted HIP. The results of these experiments are presented in

Section 7.3.2. At last, Section 7.3.3 contains the conclusions that can be drawn based on the results of the comparisons.

7.3.1. Creation of theoretical planning assignments

In order to create a diverse set of theoretical planning assignments, the base is used of a real-life planning assignment. This planning assignment is stripped by removing all the trains that need to be parked and service from the planning assignment. The stripped real-life planning assignment only contains the ongoing trains, which are trains that are out of the scope of node planners. From this strip base, different planning assignments are created by adding trains of one rolling stock at a time and by testing different rolling stocks. The three rolling stock types that are used are the VIRM rolling stock, the ICR Gvc-Ehv 9 rolling stock type and the FLIRT rolling stock type. The choice for these types of rolling stock is based on the high amount of train units of these rolling stock types that are present at the node.

Based on the stripped real-life planning assignment and the possibility to add (combinations of) different types of rolling stock, six different planning assignments can be produced. Figure 7.3 presents the combination of planning assignments. On the x-axis, the complexity of the different rolling stock types and the attached train services are shown. On the y-axis, the demand of train units that needs to be parked and serviced at the node is presented. By building the planning assignments this way, a set of 6 planning assignments is created that differ on a complexity base and on a demand base. These planning assignments can be seen as representative in order to test the performance of the adjusted HIP.

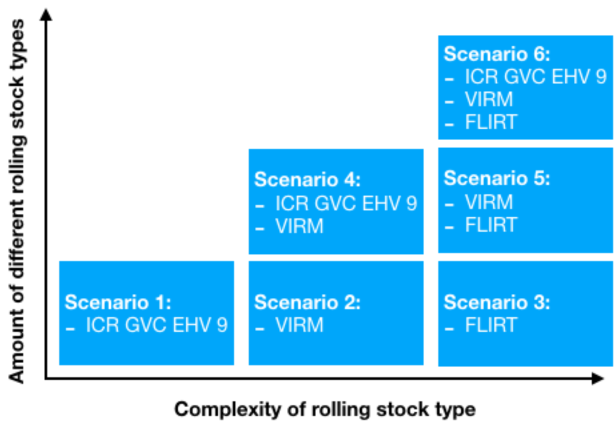

Figure 7.3: Graphical representation of the different planning assignments based on the complexity per rolling stock type and the demand of train units

In a more quantitative way, the characteristics of the planning assignments are presented in table 7.4. Planning assignment 1, 2 and 3 are characterized since only one rolling stock type is added to the ongoing trains. For planning assignment 1, this is the ICR GVC-EHV 9. This train exists out of one train and cannot be combined. This is visible in the train units per arriving and departing train, which is 1 for this planning assignment. For planning assignment 2, this number already increases because more trains will arrive or depart that consist of a combination of VIRM4 or VIRM6. The ratio train units per arriving or departing train is again higher for planning assignment 3. The trains that arrive or depart for the FLIRT service more often consist of two train units than of one. Planning assignment 4, 5 and 6 are combinations of planning assignment 1, 2, and 3.

Based on this information, it can be stated that planning assignment 1 contains the most easy planning assignment, since no trains have to be combined and split. For all the other planning assignments, the difference between the ratio train units per arriving train and train units per departing train is different, which assumes that at least one combination or split is necessary. The largest difference between these two ratios is in planning assignment 3 and 5 (0,07%). Since planning assignment 6 is the combination of planning assignment 5 and planning assignment 1, we know that for the FLIRT and VIRM trains the amount of combinations or splits necessary will remain the same. The table clearly substantiates the graphical representation of the different planning assignment (Figure 7.3). Based on these characteristics, it can be stated that the complexity of the planning assignments increases from 1 to 6, but that it is not clear if 3 or 4 is more complex. Planning assignment 3 has a higher ratio of train units per departing and arriving train, but planning assignment 4 consist of two rolling stock types.

Table 7.4: Quantitative characteristics of the theoretical planning assignments. The top number in the columns is related the the theoretical planning assignment shown in Figure 7.3.

Indicators	1	2	3	4	5	6
VIRM4	0	17	0	17	17	17
VIRM6	0	16	0	16	16	16
FLIRT3	0	0	73	0	73	73
FLIRT4	0	0	67	0	67	67
ICR GVC-EHV 9	36	0	0	36	0	36
Sum of train units	36	33	140	69	173	209
Arriving trains	36	26	76	62	102	138
Train units per arriving train	1,00	1,27	1,84	1,11	1,70	1,51
Departing trains	36	27	79	63	106	142
Train units per departing train	1,00	1,22	1,77	1,10	1,63	1,47
Ongoing trains	511	511	511	511	511	511

7.3.2. Results of comparisons based on theoretical planning assignments

The results presented in this Section are calculated taking the averages of 9 seeds. This means that 9 different solutions were calculated and the average was taken in order to increase the stability of the results. The running time of HIP was 2 minutes (120 seconds).

Table 7.5 presents the results of the original HIP. Planning assignment 1 is, as was expected, easy to solve, since the conflict costs are 0. The penalty costs, which is directly related to the number of parking movements for the original HIP, is 1. This number cannot be lower, since the amount of movements (one towards the SL and one from the SL) is lower compared to the amount of train arriving and departing. This is possible, because one train is arriving at the station and does not have time to go to the SL, so the driving direction is changed at the platform. According to the conflict costs, it can be stated that planning assignment 3 is more complex compared to planning assignment 4. It can be related that complexity is more dependent on the number of train units arriving, number of arriving and departing trains and the ratio of train units per arriving or departing train than an increase of rolling stock types.

Taking the complexity to a next level, Table 7.5 clearly shows that the complexity of planning assignments is not cumulative, but increases exponential if a planning assignment is made more difficult. This can be based on the fact that planning assignment 4 is a combination of the planning assignments 1 and 2, planning assignment 5 of 2 and 3 and 6 of 1 and 5. If the indicators of planning assignments 1 and 2 are added, they differ compared to the indicators of planning assignment 4. Next to the fact that the complexity of two added planning assignments is different compared to the complexity of the single planning assignments combined, the running time also plays a role. HIP must solve the two planning assignments in the same time as it normally would have solved one planning assignment. Since this research is not focused on the link between the planning assignment complexity and the results of HIP, this subject is ended here. However, further research might be interesting within this field.

The data of Table 7.5 is visualized in Figure 7.4. This visualization can be used in order to mark large differences in results for indicators between the planning assignments tested. For the original HIP, the illegal parking time increases rapidly for the more complex planning assignments (5 and 6), just as the penalty costs.

The results of the adjusted HIP is presented in Table 7.6. Also for the Adjusted HIP, the first planning assignment is easy to solve regarding the conflict costs. For the adjusted HIP, the conflict costs and the penalty costs increase when the planning assignments are getting more complex. Concerning the conflict costs, planning assignment 4 seems to be less complex compared to planning assignment 3. The penalty costs however show

Table 7.5: Results of original HIP. The first column represents the indicators, the second to the seventh column show the results per planning assignment.

Indicator	**1**	**2**	**3**	**4**	**5**	**6**
Conflict costs	**0**	**316**	**562**	**419**	**1195**	**1512**
Crossings	0	4	2	6	12	14
Delayed departures	0	3	10	4	17	20
Delayed arrivals	0	1	2	2	7	9
Track length violations	0	1	0	1	3	6
Incorrect combinations	0	0	1	0	1	1
Delayed fixed train	0	16	16	20	44	57
Penalty costs	**1**	**2**	**5**	**4**	**21**	**47**
Number of movements	70	75	94	149	154	233
Illegal parking time violation	0	36	100	63	1541	3930
Candidates checked	20395	14061	9815	10127	6734	5592

something different.

Compared to the results of the original HIP, the penalty costs increase at a faster rate. Next to the parking costs, the illegal parking time violation has higher levels compared to the original HIP results. This is due to the fact that more tracks are marked as no parking tracks. However, it might occur that trains have to wait (park) at these tracks for a short amount of time before a next movement can be executed. Taking into account the fact that the penalty costs is directly related to the illegal parking time violation explains the large increase of penalty costs if complexity is getting larger.

Table 7.6: Results of adjusted HIP. The first column represents the indicators, the second to the seventh column show the results per planning assignment.

Indicator	**1**	**2**	**3**	**4**	**5**	**6**
Conflict costs	**0**	**217**	**336**	**306**	**680**	**846**
Crossings	0	2	0	2	4	8
Delayed departures	0	3	5	5	10	12
Delayed arrivals	0	0	1	0	5	6
Track length violations	0	0	0	0	0	0
Incorrect combinations	0	0	0	0	1	0
Delayed fixed train	0	10	14	15	27	34
Penalty costs	**1**	**80**	**32**	**109**	**365**	**410**
Number of movements	70	93	97	170	170	237
Illegal parking time violation	0	886	507	2515	24693	28640
Candidates checked	18517	13928	9937	10465	6909	5716

The data presented in Table 7.6 is made visual in Figure 7.5. The Figure resembles the Figure of the original HIP, except for the illegal parking time violation and the penalty costs, which are higher for the adjusted HIP. Taking into regard the outliers or differences, the illegal parking time still draws attention. However, the increase of this value for the different planning assignments is more gradually, since planning assignment 4 bridges the gap between planning assignment 2 and 3 and planning assignment 5 and 6. Another remark that have to be noticed is that the scores of planning assignment 5 and 6 for the different indicators are closer to each other. This might suggest that for the adjusted HIP, planning assignment 5 is of the same level of complexity compared to planning assignment 6.

With the two results presented, it is time to present the differences of the indicator values per planning assignment. Table 7.7 presents the differences per planning assignment, and in the last column the average difference is presented. First of all, lets start with the large increases of both penalty costs and illegal parking times. The reason for this difference is already presented, but an interesting thing to mention is the fact that the differences decreases if planning assignments become more complex.

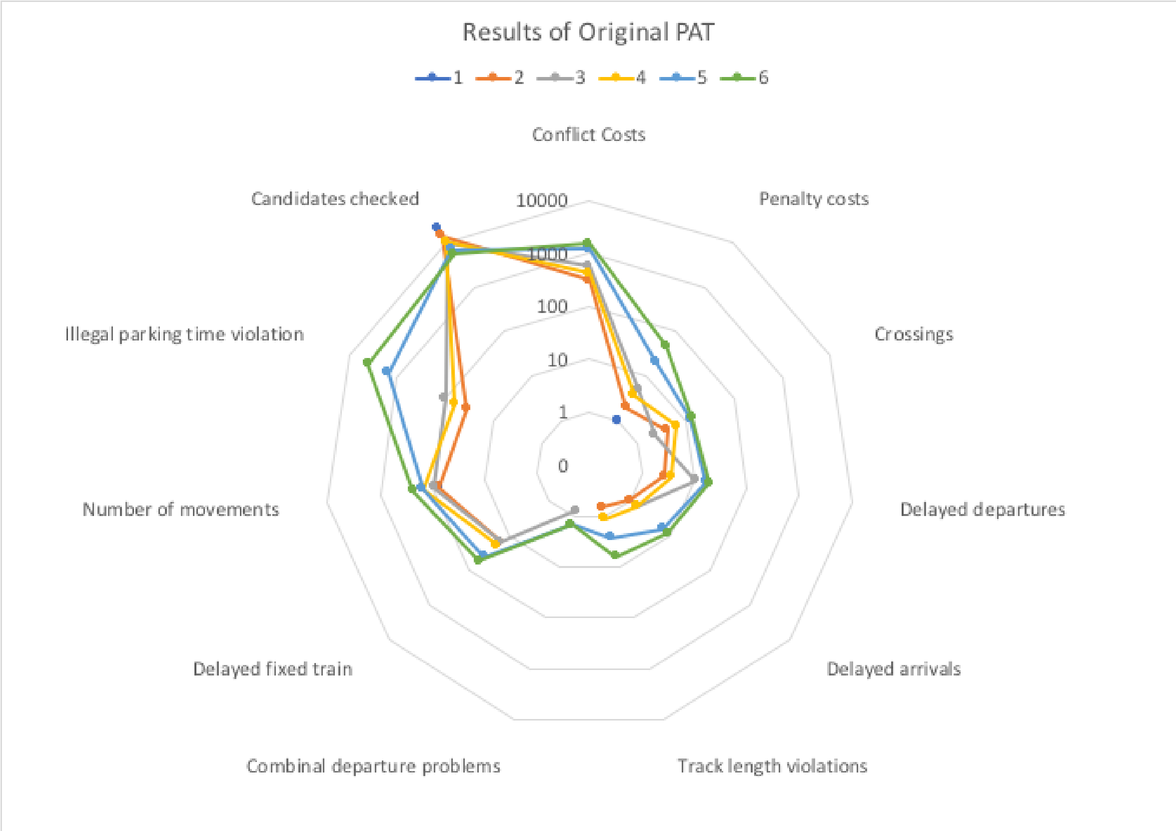

Figure 7.4: Results of original HIP for the different planning assignments. The different planning assignments are marked by the different coloured lines. The axis is logarithmic.

It seems that for the conflict costs and the indicators that are directly related to the conflict costs, the adjusted HIP performs better, since the value of HIP are for almost all indicators and planning assignments negative. The adjusted HIP performs better on al conflicts compared to the original HIP.

Something where the adjusted HIP scores negatively on is the number of movements. a clear increase of the number of movements is recorded for the adjusted HIP. In other words, the solution of the Adjusted HIP contains more movements compared to the solution of the original HIP. The increase of movements is however planning assignment specific.

The last thing that have to be noticed is the fact that the adjusted HIP checks overall the same amount of candidates compared to the original HIP. A very interesting trend is visible: the amount of candidates checked by the adjusted HIP for the less complex planning assignments is less compared to the original HIP. For the planning assignments that have a higher level of complexity, the amount of candidates being checked by the adjusted HIP exceeds the number of the original HIP.

The differences presented in Table 7.7 is made visual in Figure 7.6. The Figure does not include the data set of planning assignment 1, since there are (almost) no differences between the adjusted HIP solution and the original HIP solution. Planning assignment 1 is not representative for this part of the research, since the planning assignment is too simple to discover differences. Because of the removal of planning assignment 1, the average (shown in black diamond shaped points) has changed a little bit with regards to the average values presented in Table 7.7.

The Figure shows the difference between the performance of the different indicators based on the planning assignments. For the conflict costs, the crossings, the delayed fixed train, number of movements and candidates checked, it seems that the adjusted HIP improves or deteriorates these indicators within a certain specified range. For the other indicators it might be harder to forecast the difference of the adjusted HIP. The

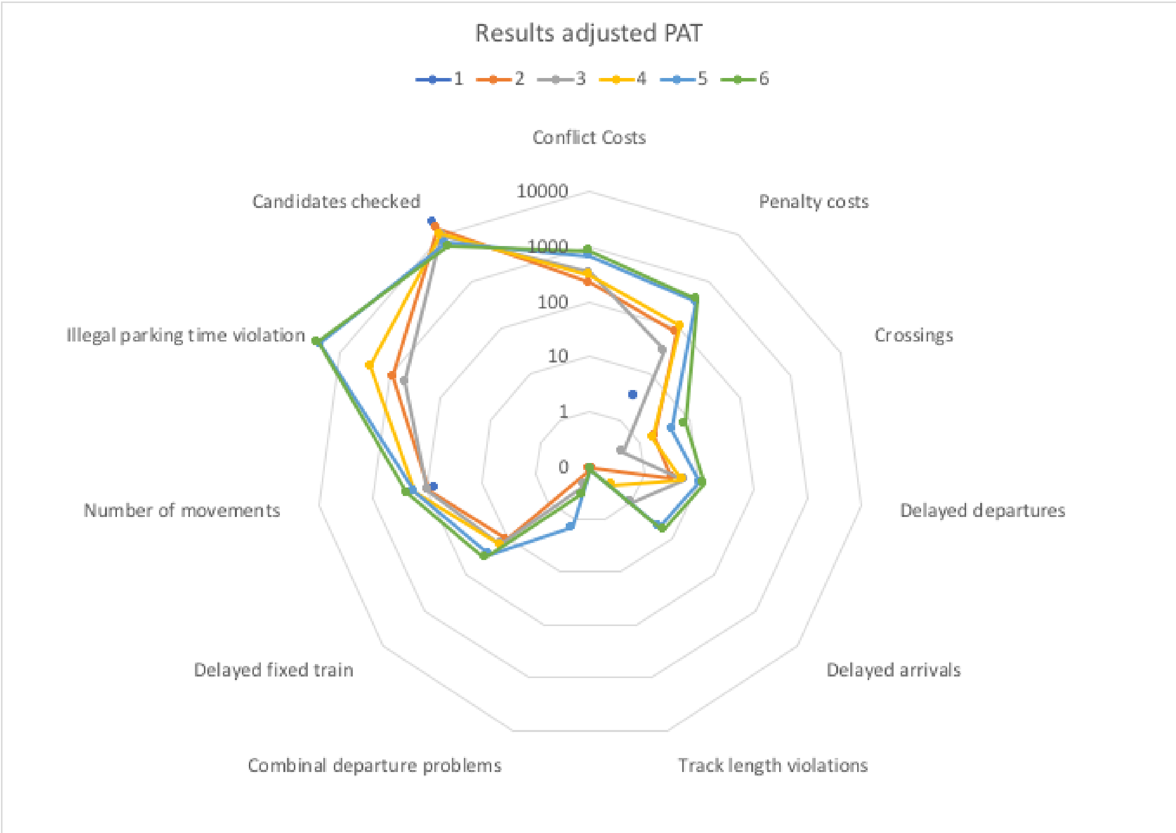

Figure 7.5: Results of adjusted HIP for the different planning assignments. The different planning assignments are marked by the different coloured lines. The axis is logarithmic.

indicator track length violations needs some extra attention, because the difference is either zero or (close to) a 100%. The 0% is presented when the original HIP already did not have any track length violations. Knowing this, it can be stated that the impact of the adjusted HIP on the track length violation can also be categorized as predictable. The same applies for the incorrect combinations. The only indicator, next to the number of movements, that scores negative for certain planning assignments is the delayed departures.

7.3.3. Conclusions of comparisons based on theoretical planning assignments

The theoretical experiment has proved that the adjusted HIP performs better on almost all indicators compared to the original HIP. The level of complexity of the planning assignments (or planning assignment) does not has a large effect on the level of improvement, unless the planning assignment is too simple and the original HIP can already create a feasible solution. The adjusted HIP does check less candidates compared to the original HIP if the complexity of the planning assignment is low. For more complex planning assignments, the number of candidates being checked exceeds the number of the original HIP. The steady negative impact of the adjusted HIP is the increase in the number of movements.

Based on the comparison, a set of indicators is characterized that show a stable trend regarding an increase of performance. For the adjusted HIP, the conflict costs decreases between the 25% and 50%, the amount of crossings with 30% to 80%, the track length violations almost with 100%, the combinal departure around 75%, the delayed fixed train between 15% and 45%. The number of movements increases between 0% and 25% if the adjusted HIP is used.

7.4. Answer on sub-question 5

The conclusion of this chapter is used to answer the following sub-question: **What is the impact of the integration of missing aspects to HIP?** To answer this question, the three different comparisons described above are concluded to describe the impact of the measures to HIP. First of all, the most important part to con-

Table 7.7: Differences between the results of the adjusted HIP and the original HIP. A positive percentage means that the values of the adjusted HIP are higher compared to the original HIP. A negative percentage means that the values of the adjusted HIP are lower compared to the original HIP.

Indicator	**1**	**2**	**3**	**4**	**5**	**6**	**Average**
Conflict Costs	**0%**	**-31%**	**-40%**	**-27%**	**-43%**	**-44%**	**-31%**
Crossings	0%	-51%	-81%	-72%	-65%	-42%	-52%
Delayed departures	0%	21%	-48%	36%	-38%	-37%	-11%
Delayed arrivals	0%	-89%	-36%	-79%	-32%	-37%	-45%
Track length violations	0%	-100%	0%	-100%	-96%	-98%	-66%
Incorrect combinations	0%	0%	-71%	0%	0%	-77%	-25%
Delayed fixed train	0%	-36%	-14%	-28%	-38%	-41%	-26%
Penalty costs	**0%**	**3894%**	**534%**	**2874%**	**1602%**	**773%**	**1613%**
Number of movements	0%	23%	2%	14%	11%	2%	9%
Illegal parking time violation	0%	2391%	409%	3913%	1502%	629%	1474%
Candidates checked	-9%	-1%	1%	3%	3%	2%	0%

clude is the qualitative impact of the aspects. The comparisons with experienced planners have shown that the measures based on the identified aspects have a positive impact on the qualitative performance of HIP. The solutions of HIP with these measures resemble the human plan more than the solutions of the original HIP. From the 60 differences identified in the experiment described in Chapter 4, 18 differences were solved completely and 25 were solved partially by the adjusted HIP. All human planners agreed that the solution produced by the adjusted HIP was a (large) improvement compared to the solution produced by the original HIP. Based on these empirical results, it can be stated that adding human aspects to a Planning Assistance Tool has a positive effect on the quality of the produced solution (or plan) from a planners perspective.

Next to a qualitative comparison, there was a quantitative comparison. Comparisons between the original HIP and the adjusted HIP with real-life planning assignments have shown that the adjusted HIP performs better on almost all performance indicators. The differences for the performance indicators differ quite much from planning assignment to planning assignment. No link can be laid between the differences in performance and the quantitative characteristics of the different planning assignments. This results in a broad range of differences of the performance of the adjusted HIP compared to the original HIP. However, the averages of the performance of the adjusted HIP on the different planning assignments are positive for almost every indicator. From the comparisons performed, it can be stated that the implementation of aspects have a positive impact regarding the performance of HIP based on real-life planning assignments.

From a theoretical perspective, the impact of the measures to the performance of HIP are positive as well. Using multiple fictional planning assignments that differed in complexity, a lot of information could be gathered by comparing the original HIP to the adjusted HIP. First of all, the increase of performance is for most of the indicators well-defined between certain boundaries. Next to this, relations can be laid between the level of complexity and the performance of the adjusted HIP. The adjusted HIP performs worse compared to the original HIP on the candidates checked indicator if the planning assignment is not complex. If a planning assignment is complex, the adjusted HIP is able to check more candidates compared to the original HIP. All the indicators perform better if the adjusted HIP is used, except the number of movements. The adjusted HIP seems to increase the amount of movements in the solutions of HIP.

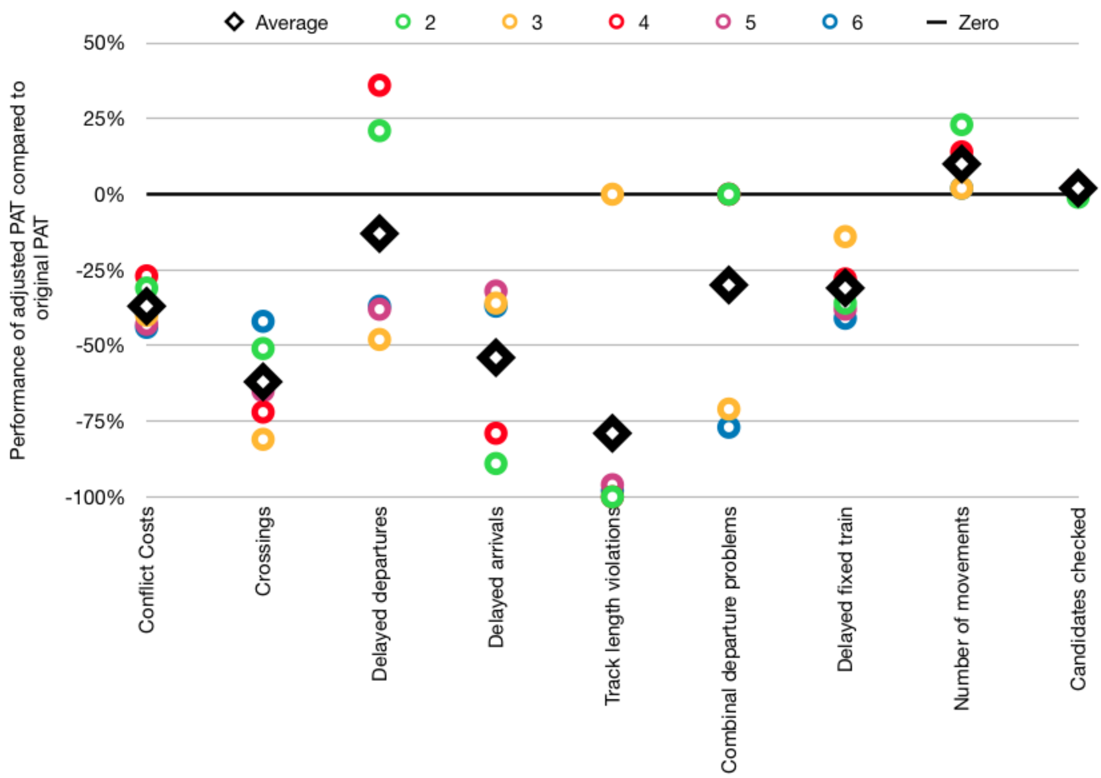

Figure 7.6: The performance (measured by the difference of values for different indicators between the adjusted HIP and the original HIP) of HIP after implementation of aspects based on theoretical planning assignments. Not taken into account is planning assignment 1 (difference is zero). The black diamond shaped points represent the average of the data set without planning assignment 1.

8
Conclusions

This Chapter contains the conclusion, the discussion and recommendations of this research.

8.1. Conclusion

The conclusion will be presented as follows: first, the sub-questions addressed in Section 1.5 are answered after which the research questions is answered.

Sub-question 1: **What is the best method is to research the qualitative difference between plans for complex planning assignments like railway node planning?** Complex planning problems are solved by dividing the large problem into sub-problems. For the Train Unit Shunting Problem (TUSP), six sub-problems can be identified: matching, job-shop scheduling, parking, routing, splitting and combining and crew assignment. Human planners currently solve this problem, but multiple algorithms have been developed that can (partially) solve the problem. However, the amount of algorithms that are implemented in real operation is low because human planners find it hard to understand and trust the solutions (the plans) of these algorithms.

The cooperation between humans and computers is beneficial for the planning process because humans are flexible and adaptable and they have intuition and the possibility to communicate and negotiate. These characteristics are preferred for coping with stated, non-stated, incomplete, outdated and missing information, which is often the case in ongoing operational processes. Computers are preferred above humans when dealing with information processing because of larger memory and fast computing times. The characteristics of humans and computers complement each other.

However, the implementation of algorithms in complex planning processes lacks because the solutions of the computer do not fit the standard of humans. Quantitative research to the differences between human and computer-generated plans is often performed. The qualitative research to these differences is lacking because these experiments can only be conducted with a high level of detail within a specific domain. Due to these outcomes, the best approach to research the qualitative difference between human and computer-generated plans is a mix of a field study and a case study. The field study part can be translated to the fact that experienced planners are needed to perform the comparison since they have much knowledge of planning. The case study part can be translated to that fact that the comparison of the plans needs to be scoped to an enclosed research area that needs to be real in order to use the experiences of the human planners.

Sub-question 2: **How is the integrated node planning of NS set up?** For all planning processes at the NS, the hierarchical production planning method is used to solve the node planning problem. This method is based on the fact that information available increases in regards to the execution time of the plan. Within the NS, there are three different phases where the level of information available increases. Based on these three phases, the level of detail needed of the plan increases as well. An example is that a plan for a general day needs to be adapted to a plan for a specific day. The combination of the amount of information available and the level of detail of the planning needed define the solution space for the different planning departments. As would be expected, the closer to the moment of execution of a plan, the smaller the solution space is. However, during the execution phase, the norms and regulations used for the planning are changed for the norms and regulations of the execution which are less strict compared to the norms and regulations used during the planning phases. This difference enlarges the solution space from the moment of execution.

Sub-question 3: **What are the characteristics of the developed Planning Assistance Tool (PAT)?** The Planning Assistance Tool that is developed by the NS is the Hybrid Integral Planning method (HIP) and includes all the sub-problems of the TUSP except crew assignment. The geographical scope of HIP is unique in a way that it contains the Service Locations (SL) and the station. This means that also movements between the station and the Service Location are planned. This geographical demarcation is called the node. HIP first generates an initial solution for the plan assignment, after which a local search method will be used to optimize the solution. The optimization is based on firstly reducing the conflict costs and secondly reducing the penalty costs. In order to run HIP, input is necessary that describes the timetable of the node, a list of trains that start or end their passenger service at the node, rolling stock characteristics and node characteristics. Next to the algorithm, a Graphical User Interface (GUI) is developed that makes it easy to adjust the input and presents the output in a way that makes it understandable for planners.

Sub-question 4: **How is it possible to identify lacking aspects of a Planning Assistance Tool using an empirical experiment?** In order to identify these aspects, an experiment set-up is presented where the comparison can be made of a human plan and a plan produced by the PAT for the same real-life plan assignment. The comparison is performed by an experienced planner, and an interview protocol is used to identify the lacking aspects. In order to perform this experiment using a real-life planning assignment, a case study is presented. The location of the case study is Eindhoven (station and Service Locations included). The research group consists of the planning department that is responsible for creating and adjusting the plan before it is going to the dispatchers. The solution space of this specific department is the smallest of all analyzed departments. The experiment is performed for six different planning assignments (real-life planning assignments). The proposed experiment model did its job: from the experiments, 59 differences were identified that could be regrouped in 37 aspects that were lacking the PAT.

Sub-question 5: **What is the impact of the integration of the missing aspects to the PAT?** First, the aspects were translated to measures that could be implemented to the algorithm. These measures consisted of three different types of adaptations: adjustments of the constraints, adjustments of input and adjustments of the objective function of the algorithm. 20 out of the 37 measures were selected for implementation based on expected value for the planner and level of difficulty to be implemented. The implementations to the original PAT resulted in the adjusted PAT. With a repetitive experiment including the same planners, it was possible to collect opinions of the planners of the adjusted PAT. The experiment compared the solutions of the adjusted PAT and the solution created by human planners. Next to this, the list of differences identified from the first experiment was used to find out if the differences were still present in the new solution of the adjusted PAT.

The opinion of the experienced planners is that the adjusted PAT produces a better solution compared to the original PAT. The solution contains fewer irrational choices in comparison with the solution created by the original PAT. The new solution better resembles the plan that was created by humans. An example of such a resemblance is the way trains are parked. Planners were able to better understand the solution of the PAT. Besides this, planners show more interest in the options the adjusted PAT presents. Where these options first were discarded because they were identified as not executable, the options of the adjusted PAT were feasible concerning all rules and norms. From a qualitative perspective, the impact of the integration of the missing aspects to the PAT is positive, because the solutions contained fewer mistakes and irrational choices. The human planners could better understand the new solution because it resembles the human plan.

Next to a qualitative comparison, a quantitative comparison was also executed between the adjusted and original PAT. The quantitative comparisons were executed on two different sets of planning assignments: real-life planning assignments and theoretical planning assignments. For the real-life planning assignments, the impact of the adjustments to the PAT was positive regarding the most important indicators. For most planning assignments, the adjusted PAT performed better for almost all indicators compared to the original PAT. This means that the adjusted PAT is able to make a better planning (regarding the indicators) compared to the original PAT. The second comparison using theoretical planning assignments showed something interesting. The theoretical planning assignments differed between each other on a level of complexity, with the most complex theoretical planning assignment being a simplified real-life planning assignment. First off all, the adjusted PAT performed better compared to the original PAT. Secondly, the comparison highlighted the fact that if the planning assignments became more complex to solve, the solution of the adjusted PAT was better compared to the solution of the original PAT. Based on these outcomes, it can be stated that the performance of the adjusted PAT is related to the complexity of the planning assignment. It can be deduced that the adjustments made to the PAT are specifically aimed at planning assignments with a high level of complexity.

The reason why the adjusted PAT seem to perform better in relation with the original PAt for complex planning assignments can be related to the extensions that are implemented. The extensions are based of empirical experiments using (real-life) complex planning assignment. These assignments have a higher level of complexity compared to the theoretical planning assignments. An advantage of using complex planning assignments is that the amount of feasible options decreases. Example: within a simple planning assignment, a train has a lot of freedoms when to start a movement, since the occupied track is not used within a short amount of time. This results in a large amount of choices. For the complex planning assignment, it might be possible that there are only 2 choices available. An extension, limiting the solution space, does have more effect on a complex planning assignment compared to a simple planning assignment. The less choices there are for the Local Search solver, the less candidates have to be checked. This means that the optimum of a planning assignment can be found in less optimization steps.

Main research question: **what does the addition of missing aspects based on human planning observations do to the solution of a Planning Assistance Tool?** It can be stated that the addition of aspects based on the human planning process does have a positive impact on the PAT if the same type of planning assignment is used. The qualitative differences are relatively more significant than the quantitative differences. In other words, the implementations to the PAT do affect the level of quality more compared to the quantitative indicators. From a qualitative perspective, the proposed solution better resembled the human solution. The human planners had the feeling to understand the plan and the choices made by the adjusted PAT better compared to the original PAT solution. From a quantitative perspective, the differences in performance were small compared to the original PAT regarding real-life planning assignments. The average results showed an overall positive outcome of the implementation of the aspects.

The quantitative comparisons using theoretical planning assignments have showed that the level of performance of the addition of missing aspects increases if the theoretical planning assignment resembles the real-life planning assignment. In other words, adjusting a PAT based on the observations of planning process A does not lead to the same improvements for planning process B. The aspects identified and added to the PAT are too much dependable on the available solution space and planning assignment. Unfortunately, it is not possible to research if these quantitative differences also occur on a qualitative field. Experienced planners are only used to real-life planning assignments. It would take a lot of time before planners have the same level of experience of solving the theoretical planning assignments.

The addition of missing aspects based on human planning observation improve the quality of the PAT. Because the solution of the tool is similar compared to the solution of the humans, it decreases hurdles that need to be tackled when implementing planning tools for complex planning processes. Planners seem to understand the solution of the PAT better, which indirectly provides a base of trust that is needed for planners to accept the solutions of the Planning Assistance Tool. In summary, this research on one side improved the algorithm in order to be better understood and used for human planners and on the other side prepared planners to the potential use of planning algorithms in their planning process in a constructive manner. Therefore it can be stated that the implementation of aspects based on the human planning process decreases the gap between the potential usage of Planning Assistance Tools and human planners for complex planning processes.

8.2. Discussion

During the research process, specific points of interest have to be noticed in order to discuss the research performed and the results. Nine points are mentioned that build up from research specific to broader academical knowledge.

Firstly, the aspects that were found within the complex planning process of the train shunting problem is an enrichment to the research field. Many researchers have come to the conclusion that the planning algorithm developed cannot be implemented, because key features are still missing. Within most of the research papers, the term key features or human factors or human element are not explained at all. What does it consist of or what are the points to focus on. Some research papers, provide a little bit more of detail. However, most of the time they are only focusing on one type of key feature. An example can be the report of Beerthuizen, where flexibility is wanted by operators of the algorithm (2018). This research provides a large list of aspects and examples, which can be related to planning processes in other industries.

Secondly, the method used to identify the aspects of a complex planning process proved to work. The application of this method can be used for many other purposes. However, the amount of time that is needed is high in order to perform the qualitative comparison at a level of detail that is wanted. It is needed to perform extensive and time-consuming research that is highly reliable of specific content knowledge. Next to this, time also needs to be invested in the relation between researcher and planners. This is crucial, because experienced planners are often used for input or experiments, however, they are not included within the larger research process. In order to capture the level of detail wanted, trust between the planner and researcher is needed.

Third, it has to be noticed that the planning assignments used are only focused on regular weekdays. Weekdays form not the most difficult planning assignments for planners. The weekends are more challenging, since the Service Locations will be occupied by trains that are not driving, creating extra constraints that make the problem more difficult. It would therefore be interesting to use the planning assignments of weekends now that the large problems are fixed in the PAT.

Fourthly, the theoretical planning assignments clearly have shown that there is a direct link between the complexity of the planning assignment and the performance of the adjusted Planning Assistance Tool. However, this is based on only one set of theoretical planning assignments that were build based upon one base planning assignment. The validity of testing scenarios is of great importance for developing a PAT. Therefore, it deserves more attention to design them. The level of complexity is based on simple characteristics (the number of trains and the difficulty of the type of trains). As stated in the Literature Review, complexity of a planning assignment cannot only be captured by numerical information, but also qualitative information. For the determination of the level of complexity of the theoretical planning assignments, only simple forms of numerical information is used. Other forms, like the time between arrivals or departure of trains from the same platform might be a more significant characteristic of the level of complexity. It might be good possible that not the number of trains is the problem, but trains depart or arrive within a limited period of time causes difficulties for planners.

Fifthly, it is noticed that planning algorithms are almost never tested using real-life planning assignments. Although the use of theoretical planning assignments can be useful to test the performance of algorithms on different specific problems, it is strange to see that within the literature discussed in this research algorithms are already called a success if they manage to plan simplified versions of real-life planning assignments. Especially with the fact that the results of an algorithm differ from planning assignment to planning assignment, it is needed that multiple real-life planning assignments are tested successfully before stating that the implementation of algorithms might be reasonable. This discovery might be a hidden reason why so many algorithms are created, but almost none have been implemented into operation. If an algorithm seems to work on quite a specific and small set of hypothetical planning assignments, this does not mean that the algorithm does the same thing for larger instances. The research described in this report is interesting because the results show that it is the other way around: if an algorithm is tested on a complex real life planning assignment, it does not mean that it performs as good for more simple planning assignments.

If more time was available from the planner side, it would have been very interesting to perform theoretical planning assignments with planners in order to capture the importance of aspects based on different planning assignments in a qualitative comparison. The challenge that arises is that these synthetic planning assignments must resemble normal planning assignments in order to have the same level of comfort a human planner is solving them. Planning assignments that are too hard might make the planner uncomfortable, while planning assignments that are too easy might not provide the level of choices that are interesting regarding the aspects human planners take into consideration. In order to test the planning assignments, the cooperation of the preparation of an experienced planners is definitely needed to determine whether a planning assignments meet the standards or not.

Second to last, from a larger perspective it is interesting to determine what the impact of this research is towards the potential implementation of PAT in complex planning processes. This method presents a manner to clearly identify aspects and measures per planning department. Since there is a relation between the solution space of a specific planning department, the planning assignment and the measures taken to improve the PAT, the method used to identify these aspects is a great way of determining the set of measures that have to be applied to a generic PAT to increase the level of acceptability of the planning departments.

The last thing that has to be noticed is the fact that this research aimed to resemble a human planning by a Planning Assistance Tool. It may be questionable if this is wanted looking purely to the efficiency of a planning. On the other side, strong and rational arguments were needed before extensions were implemented, meaning that the missing human aspects have the goal to increase the efficiency of the PAT. It is therefore interesting to determine what kind of indicators can best be used to measure the efficiency of a plan according to planners.

8.3. Recommendations

The recommendations for further research and the practical application of this research is discussed in this Section. First, the Academical points are presented in Section 8.3.1. The practical application of this research is described in Section 8.3.2

8.3.1. Academical

First of all, it is necessary to say that the research group used for this experiment is as large as possible, but the number of planners might not be very reliable for a general identification of aspects within this planning environment. From the six planners that can plan Eindhoven, five were able to collaborate on this research. Since the number of planners that can plan Eindhoven is small, it is necessary to increase the number of experiments in order to get a more reliable outcome. Because the experiments were time-consuming and planners did not have much time, the number of experiments performed were limited to 1 (and for one planner 2) per planner. For this iteration, almost all differences that were addressed by the planners were overlapping, meaning that more experiments were not needed in order to identify big problems. If differences are not overlapping and the opinions between plannes differ a lot, it is recommended to increase the number of experiments per planner. This will put the focus more to smaller differences that might include a certain kind of personal preference.

Second, it might be interesting to perform the same kind of identification in a complex planning environment where there are many planners in order to increase the number of individuals of the research group. Based on this, a more generalised list of aspects can be captured. Within the NS, there is no department that has a relatively large amount of planners. However, within the Hierarchical Production Planning environment, there are other good examples of large planning departments that deal with high complexity.

Third, the point described above leads directly to another big lack within this research field, which is the identification of planning assignments based on essential characteristics. A lot of research is focused on the identification of the differences between solutions (plans) of planning assignments, but does focus little on the differences between plan assignments. In other words, a lot of research is performed on the output and the model creating the output, while the input is neglected in the academical field. The limited quantitative characteristics that are presented do not say anything about the level of difficulty of the plan assignment. More quantitative indicators, including essential factors like time instances between arrivals and departures, need to be collected in order to measure the difficulty and relate this to complexity. For this research it is vital to include planners for their experience in order to measure the level of difficulty.

Fourth, if the level of difficulty of plan assignments can be measured, it is interesting to find out if the outcome of the PAT can be estimated based on the characteristics of the plan assignment. These relations might indicate the differences in results between the theoretical and the real-life planning assignments. The combination of planning assignment characterisation and the usage of theoretical planning assignments that focus on one specific problem makes it possible to find a relation between specific planning problems and the combination of aspects that are needed in order to create a solution that is qualitative of a high standard, but also performing quantitatively suitable.

At last, an academical recommendation that have to be noticed is that the identification of aspects can be used in a broader way within the algorithm use. Where for this research the algorithm used a Local Search Approach to solve the problem, the knowledge on the aspects is relevant for solving techniques that require a policy, like Machine Learning. A policy is needed in order to reward the computer when making a right (or better) decision. In order to create such a policy, real life examples (using human planners) can be used in order to stimulate and boost the performance of the machine learning solver. This means that the way how human planners plan need to be described using real-life planning assignments. The aspect identification method needs to be considered if machine learning solvers are used to solve complex production planning assignments.

8.3.2. Practical

In a practical way, this research can be applied and used within the specific sector that this research took place (Train Unit Shunting Planning), but also in wider context. This part will first focus on the practical application of the research for the NS (and other train operators), after which the focus will be widened to the complete hierarchical production planning perspective.

Recommendations for NS

Firstly, as described in Section 8.2, it is important to use real-life planning assignments in order to test developed algorithms. If the NS is seriously considering the possible application of the Planning Assistance Tool, it is needed to use real-life planning assignment more often to test the full applicability of the PAT to discover points of improvement. Especially since this research has uncovered that the performance of the PAT is related to the planning assignment, it is needed to increase the number of test planning assignments in order to find out how developments to the Planning Assistance Tool result using different planning assignments. At this moment, this point is a real focus point of the developing team of the NS: they are increasing the number of planning assignments and are more focused on real-life planning assignments.

Secondly, more iterations of the empirical experiment, the implementation of extensions and the evaluation of the adjusted PAT are needed in order to improve the level of quality of the PAT. Planners start to be curious about the options that the PAT provides, which means that planners are opening up to talk about the options since the less irrational choices are made by the PAT. An increase of iterations will on one hand improve the PAT, on the other hand these iterations can be used to better uncover and understand the human planning process. For these new iterations, new real-life planning assignments need to be used, since this can be interesting taking into regard that the level of importance of certain aspects differ from planning assignment to planning assignment. It is therefore needed to better analyze the planning assignments as described in the academical recommendations Section.

Thirdly, it is needed to carry out these experiments for nodes that have different characteristics compared to Eindhoven. This research is necessary to collect the set of aspects that differ between the nodes. It is necessary to relate conditions and motivations to aspects (and even to extensions). If research to multiple nodes is performed, it might be possible to relate aspects to node characteristics. An example: aspects can be linked to the fact that a Service Location is used as a shuffleboard, but do the same aspects come up when performing this research on other nodes with a shuffleboard set-up? These relations can be used to estimate a unique set of aspects per node without performing extensive and time consuming research.

At fourth, it is needed to widen the experiment to other planning departments within the NS. It would be interesting if aspects differ between planning departments in this planning chain. Secondly, it is also interesting to find out how other planning departments within the NS can use this tool. The PAT can, for example, be used by the strategic planners in order to determine more precisely the capacity of a node for different planning assignments. This feature is especially interesting for special days and for when significant maintenance works are scheduled that have not taken place before. The lack of experience or expectations for these occasions is something that often generates many problems solving the node planning and friction between the planning departments. With the help of this tool, a good prediction of the capacity can be generated so that the level of difficulty of the node planning can be managed better. The tool can also be used to test possible node layout changes in the future. By changing the layout for the tool, different options can be tested in order to calculate the capacity of the node. Based on these estimations, the cost benefits ratio per option can be determined.

The tool can also be used on the operational side of the plannings process. When disruptions occur an adjustment of the plan is necessary, which can be easily calculated by the PAT. Before the applicability of the tool in the operational phase, the aspects need to be gathered from this planning department. Because of the relaxation of norms and rules, it is expected that the aspects differ between the planning department focused on in this research and the operation planning department. Next to the aspect, the tool also needs to be able to make small adjustments to the planning, since operational planners like to have control of the planning and large adjustments will not be preferred.

Recommendations for production planning processes

Firstly, for industries that have already developed a PAT but still need to implement it, this experiment can be used. Firstly the performance of the PAT will be improved. Secondly the acceptance of human planners to the possible application of the PAT within their planning process will be higher. This method can also be perfectly used in order to describe the way human planners plan, which is for many hierarchical production planners unknown to other departments. The fact that the planning process might even differ between two the same type of planning departments of different locations is a perfect example of the irregularity of planning complex production processes. Although the method of this study is extensive and time-consuming, it seems that this is currently the only structured way that can be used to describe such complex planning processes at a detailed level. This research has proven that the level of detail is incredibly important considering these irregular planning processes.

However, if no PAT is developed yet for a production planning process but there is a wish to develop one, it is advised to developing teams to integrate human planners from the beginning of the development. First of all, an earlier participation of human planners within the developing period focuses on the real problems that planners have, meaning that a more efficient developing path can be created. The second point why it is important to have planners included to the developing team is that experienced planners do have a lot of power. Implementing a PAT within an ongoing and continuous planning process involves politics. Especially when the knowledge of planners is shared from experienced planner to newcomer, the knowledge transmission is completely (and only) owned by experienced planners. It is therefore important to include planners within the developing team in order to cooperate together and work as efficient as possible when creating such a tool. At last, planners know that within critical production processes, there will always be humans necessary to control and check solutions created by computers and to intervene if something is going wrong. This gives planners the power to decide if a costly implementation of a tool is going to be successful or a failure.

Bibliography

Emrah Arica, Jan Strandhagen, and Hans-Henrik Hvolby. Designing a Decision Support System for Production Scheduling Task in Complex and Uncertain Manufacturing Environments. pages 589–596, 2014. doi: 10.1007/978-3-662-44739-0{_}72{\"{i}}. URL https://hal.inria.fr/hal-01388601.

E Beerthuizen. Optimizing train parking and shunting at NS service yards. Technical report, TU Delft, Delft, 2018. URL http://repository.tudelft.nl/.

M. Berglund and J. Karltun. Human, technological and organizational aspects influencing the production scheduling process. *International Journal of Production Economics*, 110(1-2):160–174, 10 2007. ISSN 09255273. doi: 10.1016/j.ijpe.2007.02.024.

Gabriel R. Bitran, Elizabeth A. Haas, and Arnoldo C. Hax. Hierarchical Production Planning: A Single Stage System. *Operations Research*, 29(4):717–743, 8 1981. ISSN 0030-364X. doi: 10.1287/opre.29.4.717. URL http://pubsonline.informs.org/doi/abs/10.1287/opre.29.4.717.

BMI Projectteam. Adviesrapport Besturing Materieelinzet. Technical report, NS, Utrecht, 2018.

Milagros Castillo-Montoya. The Qualitative Report Preparing for Interview Research: The Interview Protocol Refinement Framework. Technical report, 2016. URL https://nsuworks.nova.edu/tqr/vol21/iss5/2.

Julien Cegarra and Wout Wezel. A Comparison of Task Analysis Methods for Planning and Scheduling. In *Behavioral Operations in Planning and Scheduling*, pages 323–338. 11 2011. doi: 10.1007/978-3-642-13382-4{_}13.

Alexandra Coman and Héctor Muñoz-Avila. Qualitative vs. Quantitative Plan Diversity in Case-Based Planning. 2011.

Jean François Cordeau, Paolo Toth, and Daniele Vigo. A survey of optimization models for train routing and scheduling. *Transportation Science*, 32(4):380–404, 1998. ISSN 00411655. doi: 10.1287/trsc.32.4.380.

S Crawford, B L Maccarthy, J R Wilson, and C Vernon. Investigating the Work of Industrial Schedulers through Field Study. Technical report, 1999.

Dagblad van het Noorden. Groningen krijgt 10 kilometer nieuw spoor op rangeerterrein De Vork. 2019.

J. C. De Man, T. Nehzati, E. Arica, and K. Kiil. Automating scheduling tasks; Experience of a real case. In *IEEE International Conference on Industrial Engineering and Engineering Management*, volume 2016-January, pages 1447–1451. IEEE Computer Society, 1 2016. ISBN 9781467380669. doi: 10.1109/IEEM.2015.7385887.

D de Ruijter, C Pustjens, and W Fleuren. De behandelcalculator. Technical report, NS, 2019.

Joris Den Ouden. Generating robust schedules for train maintenance staff. Technical report, Universiteit Utrecht, 2018.

Joris Den Ouden. The inner workings of HIP. Technical report, NS, Utrecht, 2019.

Manuel Dios and Jose M. Framinan. A review and classification of computer-based manufacturing scheduling tools. *Computers and Industrial Engineering*, 99:229–249, 9 2016. ISSN 03608352. doi: 10.1016/j.cie.2016.07.020.

Jose M. Framinan and Rubén Ruiz. Guidelines for the deployment and implementation of manufacturing scheduling systems, 4 2012. ISSN 00207543.

Richard Freling, Ramon M. Lentink, Leo G. Kroon, and Dennis Huisman. Shunting of passenger train units in a railway station. *Transportation Science*, 39(2):261–272, 2003. doi: 10.1287/trsc.1030.0076.

Rob M P Goverde. Lecture Railway operations II. Technical report, 2018.

Varun Grover and William Kettinger. *Process Think: Winning Perspectives for Business Change in the Information Age*. IGI Global, 1999.

Jørgen Thorlund Haahr, Richard M. Lusby, and Joris Camiel Wagenaar. Optimization methods for the Train Unit Shunting Problem. *European Journal of Operational Research*, 262(3):981–995, 11 2017. ISSN 03772217. doi: 10.1016/j.ejor.2017.03.068.

A.C. Hax and H.C. Meal. HIERARCHICAL INTEGRATION OF PRODUCTION PLANNING AND SCHEDULING. *Logistics*, pages 53–69, 1975.

Barbara Hayes-Roth and Frederick Hayes-Roth. A Cognitive Model of Planning. Technical report, 1979.

Jean Michel Hoc. From human – machine interaction to human – machine cooperation. *Ergonomics*, 43(7): 833–843, 7 2000. ISSN 13665847. doi: 10.1080/001401300409044.

Per Munk Jacobsen and David Pisinger. Train shunting at a workshop area. *Flexible Services and Manufacturing Journal*, 23(2):156–180, 6 2011. ISSN 19366582. doi: 10.1007/s10696-011-9096-1.

Dirk Jekkers. Train shunt planning using genetic algorithms. Technical report, 2009.

Anssi Käki, Katariina Kemppainen, and Juuso Liesiö. What to do when decision-makers deviate from model recommendations? Empirical evidence from hydropower industry. *European Journal of Operational Research*, 278(3):869–882, 11 2019. ISSN 03772217. doi: 10.1016/j.ejor.2019.04.021.

Leo G. Kroon, Ramon M. Lentink, and Alexander Schrijver. Shunting of passenger train units: An integrated approach. *Transportation Science*, 42(4):436–449, 2008. ISSN 15265447. doi: 10.1287/trsc.1080.0243.

Mickey Lauria and Jacob A. Wagner. What Can We Learn from Empirical Studies of Planning Theory? A Comparative Case Analysis of Extant Literature. *Journal of Planning Education and Research*, 25(4):364–381, 2006. ISSN 0739456X. doi: 10.1177/0739456X05282351.

John D. Lee and Bobbie D. Seppelt. Human Factors in Automation Design. In *Springer Handbook of Automation*, pages 417–436. Springer Berlin Heidelberg, 2009. doi: 10.1007/978-3-540-78831-7{_}25.

Ramon M Lentink, Pieter-Jan Fioole, Leo G Kroon, and Cor Van 't Woudt. Applying Operations Research techniques to planning of train shunting. Technical report, Erasmus University, Rotterdam, 2003.

Ramon Martijn Lentink. Algorithmic Decision Support for Shunt Planning. Technical report, Erasmus University Rotterdam, 2006.

Johan Lorentzon and Johan Fredlund. Application of Performance Measurement on Manufacturing Simulations for Knowledge-Based Decision Support. Technical report, 2017.

K McKay, J Buzacott, and F Safayeni. The scheduler's knowledge of uncertainty: The missing link. *Knowledge Based Production Management Systems*, 1989.

Kenneth N. McKay and Vincent C.S. Wiers. Integrated decision support for planning, scheduling, and dispatching tasks in a focused factory. *Computers in Industry*, 50(1):5–14, 1 2003. doi: 10.1016/S0166-3615(02)00146-X.

Kenneth N. McKay and Vincent C.S. Wiers. The human factor in planning and scheduling. In *International Series in Operations Research and Management Science*, volume 89, pages 23–57. Springer New York LLC, 2006.

Nasser Mebarki, Olivier Cardin, and Clément Guerin. Evaluation of a new human-machine decision support system for group scheduling. In *IFAC Proceedings Volumes (IFAC-PapersOnline)*, volume 12, pages 211–217, 2013. ISBN 9783902823410. doi: 10.3182/20130811-5-US-2037.00040.

Erik Meijer. Het verbeteren van plannings- en schedulingsprocessen bij Stach van Katoren. Technical report, Technische Universiteit Twente, 2012.

Philip G. Moscoso, Jan C. Fransoo, and Dieter Fischer. An empirical study on reducing planning instability in hierarchical planning systems. *Production Planning and Control*, 21(4):413–426, 6 2010. ISSN 09537287. doi: 10.1080/09537280903454172.

Karen L Myers and Thomas J Lee. Generating Qualitatively Different Plans through Metatheoretic Biases. Technical report, 1999. URL www.aaai.org.

NRC. ProRail verbouwt voor reizigersgroei, 6 2018.

NS. Aanschaf nieuwe Intercity's, 2016. URL https://nieuws.ns.nl/aanschaf-nieuwe-intercitys/.

NS. Functionele beschrijving OpstelPlanGenerator (OPG). Technical report, NS, 2018a. URL www.ortec.com.

NS. Kwaliteit Definities: Personeel. Technical report, Utrecht, 2018b.

Seamus O'Reilly, Anita Kumar, and Frédéric Adam. The role of hierarchical production planning in food manufacturing SMEs. *International Journal of Operations and Production Management*, 35(10):1362–1385, 10 2015. ISSN 17586593. doi: 10.1108/IJOPM-04-2014-0157.

ProRail. Netverklaring 2020. Technical report, Utrecht, 2019.

Rail Net Europe. Glossary 2019 Final. 2019.

Jan Riezebos, Jean Michel Hoc, Nasser Mebarki, Christos Dimopoulos, Wout Van Wezel, and Guillaume Pinot. Design of scheduling algorithms: Applications. In *Behavioral Operations in Planning and Scheduling*, pages 371–412. Springer Berlin Heidelberg, 2011. ISBN 9783642133817. doi: 10.1007/978-3-642-13382-4{_}15.

H.J. Rubin and I.S. Rubin. *Qualitative interviewing: the art of hearing data*. Sage, Thousands Oaks, 3rd edition, 2012.

P.M. Sanderson. The Human Planning and Scheduling Role in Advanced Manufacturing Systems: An Emerging Human Factors Domain. *Human Factors*, 31(6), 1989.

Joris Snijders. 'Proof of Concept' Behandelcalculator. Technical report, NS, Utrecht, 2019.

Sporenplan Online. Sporenplan Online, 2019. URL http://www.sporenplan.nl.

Martin Steinrücke. Tactical planning in supply chain networks with customer oriented single sourcing. *The International Journal of Logistics Management*, 23(2):259–279, 1 2012. doi: 10.1108/09574091211265387. URL https://doi.org/10.1108/09574091211265387.

Treinreiziger. NS en ProRail: het spoor is bijna vol, 2018. URL https://www.treinreiziger.nl/ns-en-prorail-spoor-is-bijna/.

Marjan Van Den Akker, Hilbrandt Baarsma, Johann Hurink, Maciej Modelski, Jacob Jan Paulus, Ingrid Reijnen, Dan Roozemond, and Jan Schreuder. Shunting passenger trains: getting ready for departure. Technical report, Universiteit Utrecht, 2008.

John van den Broek and Ramon Lentink. Rangeerplannen maken voor treinen. Technical report, Nieuwe Wiskrant, 2008.

Roel W Van Den Broek. Train Shunting and Service Scheduling: an integrated local search approach. Technical report, 2016.

M.T. Van Dommelen. Scheduling train service activities at service sites to determine the capacity. Technical report, VU Amsterdam, 2015.

Hans van Uden, Danielle Smolders, Peter Oosterbaan, Ron Verheugd, Ruud Rol, and Joris Snijders. De beste manier van knooppuntplanning. Technical report, NS, 2017.

Wout Van Wezel and René Jorna. Cognition, tasks and planning: Supporting the planning of shunting operations at the Netherlands Railways. *Cognition, Technology and Work*, 11(2):165–176, 6 2009. ISSN 14355558. doi: 10.1007/s10111-007-0106-8.

Jos Verlaan. NS: grote problemen dreigen op spoor en stations, 8 2019.

W H M Zijm. Towards intelligent manufacturing planning and control systems Perspektiven intelligenter Produktionsplanungs-und Produktionssteuerungssysteme. Technical report, 2000.

A

Scientific Paper

Adding human aspects to Planning Assistance Tools for complex logistical processes: a case-study of railway node planning in the Netherlands

M.P. Baurichter[a,*], Prof. Dr. R.M.P. Goverde[a], Dr. N. Bešinović[a], Dr. W.W. Veeneman[b], Ir. W. Fleuren[c]

[a] *Delft University of Technology. Faculty of Civil Engineering and Geosciences. Mekelweg 5, 2628 CD Delft, The Netherlands*
[b] *Delft University of Technology. Faculty of Technology, Policy and Management. Jaffalaan 5, 2628 BX Delft, The Netherlands*
[c] *Nederlandse Spoorwegen. Research & Development Node Logistics. Laan van Puntenburg 100, 3511 ER Utrecht, the Netherlands*

Abstract

Computers are preferred above humans when dealing with information processing because of the great memory and fast computing times. The use of algorithms within complex planning processes is an interesting research field. These algorithms should help human planners with the planning process and therefore are called Planning Assistance Tools (PAT). However, the implementation of a PAT often fails because of the lack of acceptance by human planners. The lack of acceptance is often linked to the fact of human aspects that are missing within the PAT. This research provides a methodology to identify the missing aspects using a developed PAT, real-life planning assignments and experienced planners within the railway node planning. Based upon empirical experiments, a list of aspects is formulated that are translated to extensions that are implemented in the PAT. The application of such a method to a real-life case has not yet been mentioned in literature before. The research proves that the implementations have a positive effect on the quality of the solution provided by the PAT, while not harming the performance of the solution. Planners seem to better understand the solution of the PAT and the decisions made by the PAT, which increases the level of trust and the level of acceptance. The method used to identify aspects and implement them might be applicable within the indus-

*Corresponding author
Email address: matthijsbaurichter@gmail.com (M.P. Baurichter)

try for planning tools that are already developed, but not yet implemented because of missing acceptance of human planners.

Keywords: Planning Assistance Tool, complex planning process, human factor, human aspects, planning assignment, empirical experiment, Train Unit Shunting Problem, node logistics, NS, qualitative comparison, quantitative comparison, Hierarchical Production Process

1. Introduction

Dutch Railways (in Dutch: Nederlandse Spoorwegen, NS) is the biggest Railway Undertaking (RU) in the Netherlands. In order to execute services on more than 4500 km of rail, the NS uses over 450 trains. If rolling stock is not used, it needs to be parked at a Service Location (SL) in order to keep the infrastructure available for other trains and services need to be executed to maintain the fleet. Customers expect clean trains that are safe to travel with.

Within the Netherlands, the amount of train passengers is growing every year. NS is warning the government about the rapid growth of passenger demand and the limited capacity of the rail system (Verlaan, 2019) (Treinreiziger, 2018). To respond to this problem, NS is expanding their fleet (NS, 2016) in order to move more passengers. But an expansion of the fleet can be directly linked to a necessity of more parking and service areas needed for trains. ProRail, the Infrastructure Manager (IM) of the rail network in the Netherlands is renovating and expanding Service Locations (NRC, 2018) (Dagblad van het Noorden, 2019). However, physical improvements of the network take up a lot of time and cost a lot. Therefore the availability of the existing infrastructure must be utilized es efficient as possible. Better use through better planning and control processes should increase the capacity of Service Locations.

The current planning process of the parking and servicing of trains is mostly executed by hand. The current plan process does not function as expected on specific nodes since not all the produced plans are always feasible (van Uden et al., 2017). The complex planning assignment can be partly solved by a computer, increasing the efficiency of the planning process as well as the efficiency of the Service Location. By implementing a Planning Assistance Tool (PAT), NS expects an increase of feasible plans. Next to that, NS expects that the plans are better measurable (within and between nodes), provide more control, generate more understanding of the complete planning process, increase the transparency and improve solutions regarding parking and service capacity (van Uden et al., 2017).

Introducing, implementing and working with a PAT comes with problems for actors in planning processes (McKay & Wiers, 2006). Whereas in theory a PAT should result in greater efficiency, lower workload and fewer human errors (Lee & Seppelt, 2009), the practical examples show something different. Experience from simulation projects has shown that plans from a PAT are difficult to interpret, difficult to trust, difficult to compare, mutually contradicting and difficult to act upon for human planners (Lorentzon & Fredlund, 2017). Computers tend to have strict rules and a large computing power to generate new plans and restructure existing ones fast while humans have the advantage of using soft skills for planning processes (Sanderson, 1989). It is because of these characteristics that terms like computer-based decision making and human-centred design are solutions that are based on the advantageous characteristics of computers and humans. However, there is a lack of fit between the human element and the formal or systematized element found in PAT (McKay & Wiers, 2006). McKay & Wiers (2006) call this lack of fit in computer-generated plannings the *human factor*. However, the term *human factor* is comprehensive and does not describe clear aspects. Besides, these human factors differentiate per planning process and planning structure. Also a lot of research is performed to the quantitative comparisons between humans and computer-generated planning processes, but almost no research is performed to the qualitative comparisons between these two different planning techniques. The qualitative differences can also influence the acceptance of Planning Assistance Tools by humans.

The following questions arise that are addressed in this paper: Is it possible to identify human factors by comparing qualitative differences between the human and the computerized way of solving planning assignments? And is a PAT better accepted if these human factors are known and integrated into the tool?

In this paper, an empirical experiment is designed in order to capture qualitative differences between an existing PAT and a human planning processes. Based on the outcomes of this experiment, extensions can be taken to improve the PAT. The PAT including the extensions is called the adjusted PAT In order to evaluate what the effect of the extensions taken to the solutions of the PAT and the reaction of human planners, quantitative and qualitative comparisons are performed between the plan of the adjusted PAT and of human planners. Based on the outcomes of these comparisons, a conclusion is drawn concerning the possibility to identify the human factor, the addition of this factor to the PAT and the effects of the extensions to the PAT and the reaction of human planners. The case study that is used to test this methodology is the shunting process of the NS, in which real-life planning assignments are used.

The main contributions of this research are:

1. the experiment framework;
2. a list of human aspects that are lacking PAT;
3. a set of extensions for PAT;
4. combining qualitative and quantitative approaches;
5. a real-life case study focused on shunting planning;
6. detailed content that should help increase the acceptance of PAT within planning processes.

Since qualitative comparisons between human and computer generated plans are never researched at a specified level, this research is filling this gap. The research is unique because it combines the identification of human aspects with the direct implementation and testing of the adjusted tool.

In order to structure the paper, the the research is explained using multiple sections. The paper start with a literature review in Section 2. The exact problem this research tackles is described in Section 3. The empirical experiment and the implementation of extensions is described in Section 4. The evaluation of the implementation of extensions and the reaction of human planners is presented in Section 6. Finally, the research is concluded and discussed in Section 7.

2. Literature Review

The literature review presents the reasons why it the cooperation between humans and computers are beneficial for complex planning processes and what the problems are regarding the implementation of algorithms. This section also describes the usage of empirical experiments in order to capture the qualitative differences between human and computer generated plannings.

Although planning tools can solve plannings problems on their own, it is not wanted to let these systems take over the task of human planners. The most important reason is the fact that computers might fail or might not work in which the human needs to take over the task (Hoc, 2000). There are also human characteristics that are beneficial and needed to solve complex problems (Sanderson, 1989). Sanderson is one of the first researchers that acknowledged that humans and algorithms seem to have complementary strengths which could be combined.

Beneficial characteristics of humans are flexibility, adaptability, intuition, the possibility to communicate and to negotiate (McKay et al., 1989). With these skills, humans are preferred above computer when cooping with stated, not-stated, incomplete, outdated and missing information (goals, constraints and data). Computers are preferred above humans when dealing with information processing because of its great memory and fast computing times (Sanderson, 1989). The mix of these characteristics of humans and computers are beneficial when solving complex problems.

The development of algorithms and optimization methods is a great innovation, but it has to be noticed that the implementation of the PAT faces a lot of difficulties. Cordeau et al. (1998) created an overview of 153 papers on train routing and scheduling problems and concluded that even though the proposed models are tested on realistic data instances, very few are actually implemented and used in operation. Cordeau et al. state in 1998 that: "Effort must be made to bridge the gap between theory and practice". Over time, not a lot of improvement has been made to fill this gap. In order to face the important reasons behind the implementation challenge of advanced scheduling solutions in practice, De Man et al. (2016) mention that: "recent studies indicate the need for scheduling tools that focus on the actual work routine, characteristics of the scheduling environment and cognitive tasks of the scheduler".

A Planning Assistance Tool is generally designed independently of the end-user, and psychological relevance is considered to be low (Cegarra & Wezel, 2011). Experience from simulation projects has shown that automatic generated (and optimal) plans by PATs are difficult to interpret, difficult to trust, difficult to compare, mutually contradicting and difficult to act upon for humans (Lorentzon & Fredlund, 2017). Sufficient knowledge has not been gathered about what planning really means and what people do when they plan (McKay & Wiers, 2003). Without knowing the key variables and attributes of this problem, it is difficult to craft solutions that consistently and clearly assist with problem-solving.

But why is the implementation of planning tools in practice hard? The lack of fit between the human element and the formal or systematized element found in the planning methods and systems is the biggest problem (McKay & Wiers, 2006). This lack of fit can be summarized in the following elements:

- Difference in the interpretation of constraints. Constraints implemented in the tool do not resemble the same constraints that human planners have (Dios & Framinan, 2016) (Arica et al., 2014). Dios &

Framinan (2016) state that a planning tool should provide the user a system to handle the soft constraints. Hidden constraints are not recognized during the development of the tool, but are considered by human planners nevertheless Framinan & Ruiz. The problem regarding the constraints result in a perceived problem and solution space for the model (Arica et al., 2014), while in real life it is reasonable to assume that there is more planning flexibility than any system can model (Moscoso et al., 2010).

- The lack of human expertise in PAT. Humans often contain expertise that is hard to implement in planning tools. The incorporation of human expertise into planning tools has been controversial for a long time (Dios & Framinan, 2016).

- Misinterpretation of reality and the difference in solutions space. Key data inserted in an algorithm does not always resemble the real data or the estimated data (Dios & Framinan, 2016) (De Man et al., 2016). Next to that, the solution space of a computer is different and less flexible compared to the solution space of planners. The key data inserted restricts the usability of the tool for the user, who has a better view of the reality and more flexibility in the solution space and therefore can act with greater flexibility and adaptability (McKay et al., 1989).

- The lack of user input and judgement in PATs. Although rescheduling is considered an integral part of planning tools, algorithms are simple or not driven by the user judgement (Dios & Framinan, 2016).

- A Graphical User Interface (GUI) that is not created together with end-users. Visualization plays an important role in the successful implementation of systems (Arica et al., 2014). The review of many production scheduling tools showed that a lot of tools were not implemented because of a malfunctioning or not desired GUIs (Dios & Framinan, 2016) (De Man et al., 2016).

Empirical experiments on planners have the goal to formulate and implement rule-based systems or heuristics with human-based rules of thumb (Van Wezel & Jorna, 2009). In order to determine what kind of empirical experiment is best to identify the human factors within a human planning process, a literature overview was created (see table 1). Experiments comparing results of human and computerized planning process have two forms:

Research	Laboratory study	Survey	Case study	Field study	Quant.	Qual.	Domain-specific	Domain
Riezebos et al. (2011)			x	x	x		x	Shunting problem
Mebarki et al. (2013)	x				x		x	Production process
Käki et al. (2019)		x			x		x	Hydro power plant
Myers & Lee (1999)	x					x		
Coman & Muñoz-Avila (2011)			x			x		
Berglund & Karltun (2007)				x		x	x	Production process
Baurichter (2020)			x	x	x	x	x	Shunting problem

Table 1: Overview of characteristics of empirical experiment of different scientific papers focused on the comparison between human and computer generated solution to complex planning assignments.

quantitative and qualitative. While quantitative comparisons are easy to execute because no domain-specific knowledge is needed, qualitative comparisons are based on domain-specific characteristics which human experts might use to differentiate between plans, thus being able to produce results of greater practical value (Coman & Muñoz-Avila, 2011). Berglund & Karltun (2007) have proved that field study set-up for the empirical experiment (e.g. sitting next to the planner) is the best way of obtaining the most detailed information. Since the goal is to compare the planning of the PAT with a real-life planning, a field study is the best option to obtain the data. Since the research area is limited, the experiment can also be seen as a case study, just as the experiment of Riezebos et al. (2011). A laboratory study is not chosen, because it lacks aspects of reality (Mebarki et al., 2013) (Myers & Lee, 1999), and surveys do not provide enough information in order to perform a thorough qualitative analysis because the content is too specific and it would require to much time for planners to answer (Käki et al., 2019). Other important lessons that are important for the experiment is the fact that experienced planners are needed if comparisons are needed in a real-life situation (Mebarki et al., 2013).

3. Problem Description

It is clear that for the implementation of a PAT human acceptance of planners is needed. This acceptance can be increased if human planners are better able to understand, interpret and trust the solutions proposed by a PAT. A way of doing this is resembling the human solution as good as possible by the PAT. Adding human aspects to the PAT might be an effective way. However, these human aspects need to be identified using an empirical experiment.

This research has two objectives. The first objective is **to determine which aspects are missing in Planning Assistance Tools for com-**

plex planning processes. The term *aspects* refers to all objective and rational reasons why a human solution of a planning assignment is different compared to a computer-based solution. The second objective is **to find out what the implementation of aspects based on human planning observations in the PAT does to the solution of the PAT and the acceptance of human planners**. Does the solution of the PAT with aspects resemble the solution of a human planner? In order to structure the research, the following research question is proposed: **what does the addition of missing aspects based on human planning observations do to the solution of a Planning Assistance Tool?**

Since this research is multidisciplinary and wide in content, scoping is necessary to set-up the research field. The case study that is used for the empirical experiment and the PAT is a railway node planning process. This planning process requires to the characteristics of the complex planning process. The Railway Undertaking (NS) is currently working on a PAT that can be used. The focus of this research will be based on the operational phase of the planning process. This means that no long term (strategical) planning process is involved, just as real-time planning or control by dispatchers. Finally, it is needed to mention that this research is focused more on the human aspects within the planning process than on the implementation of these aspects in an algorithm. Therefore, some implementation might be executed in a simplistic form covering the aspect as good as possible, but not completely.

4. Methodology

Firstly, this Section describes the empirical experiment set-up for identification of missing aspects within the PAT. Secondly, the steps that are necessary to implement extensions based on the identified aspects are described.

4.1. Empirical experiment set-up

The goal of the empirical experiment is to gather aspects that planners think are important to include in planning options that the current algorithmic planning is not including. As stated in the literature review, a field study using a real-life case-study is necessary to capture the level of details necessary.

The conceptual model of the empirical experiment is presented in Figure 1. The input for the experiment is a real-life planning assignment. The planning assignment includes information about:

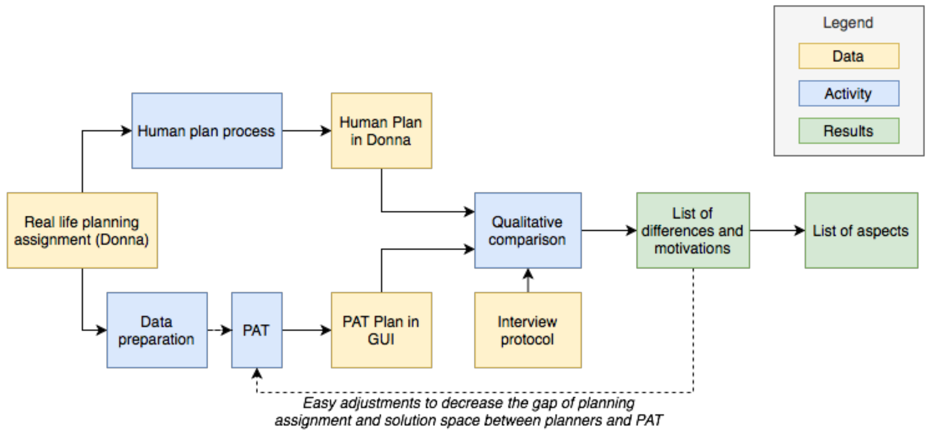

Figure 1: Conceptual model of the empirical experiment performed with human planners in order to compare human solutions of planning assignments with solutions created by the Planning Assistance Tool.

1. trains that end a passenger service at the node;
2. trains that start a passenger service from the node;
3. ongoing trains that call at the station but do not end a passenger moving service;
4. services that need to be applied to different train units;
5. tracks that are out of service.

This planning assignment is stored in the current planning database of the IM called Donna. The planning assignment can be directly used for the human planning process, since a human planner also uses Donna. Data preparation is necessary so it can be used within the PAT. The result of the human plan process is a human plan (in Donna). The result of the PAT is a computer generated plan called the PAT Plan. This plan is presented in a Graphical User Interface (GUI). A qualitative comparison is performed by an experienced planner using the human plan and the PAT plan. In order to structure this comparison and extract the most important information, an interview protocol is used (see Appendix B). This interview protocol is needed to fulfil to the following list of objectives:

1. Characteristics of planning assignment
2. Characteristics of human planning
3. Personal opinion of own planning or planning of PAT
4. Identification of differences between plannings

5. Why is the solution of PAT not technical feasible
6. Why is the solution of PAT preferable?
7. Why was human solution picked?
8. What is causing non-preferred solutions of PAT?
9. Improvement of the visualization and other characteristics of PAT
10. Grouping of differences and aspects

The result of the comparison is a list of differences between the human plan and the PAT plan. Motivations why a human planner preferred one option above the other are listed as well. Performing this experiment multiple times using different planners enlarges the list. Based on the list of differences and the motivation of multiple planners, the differences were able to be grouped in aspects. The aspects cover multiple differences with the same main motivation, like safety. There can be differences that are covered by multiple aspects.

In order to make sure the level of detail of the information captured was satisfied during the comparison, simple differences that were easy to fix were fixed during the experiments. These differences had to be generic and widely supported by planners, such as strict traffic rules.

4.2. Extension creation, selection and implementation

Based on the list of differences and the list of aspects, it is possible to create extensions that have to be added to the PAT. The process is made visible in Figure 2. A selection of the most important aspects is made based on the number of appearances during the experiment and the level of importance by human planners. This selection was discussed with the development team of the PAT to create extensions. These extensions have three different forms: extensions that adjust constraints, extensions that adjust the objective function and extensions that adjust the input of the PAT. Based on the level of difficulty and the level of importance of the aspect, extensions are implemented in the PAT.

5. Results

This Section consist out of an explanation of the case study. Secondly, the initial results of the empirical experiment are presented. Finally, information about the selected extensions is presented.

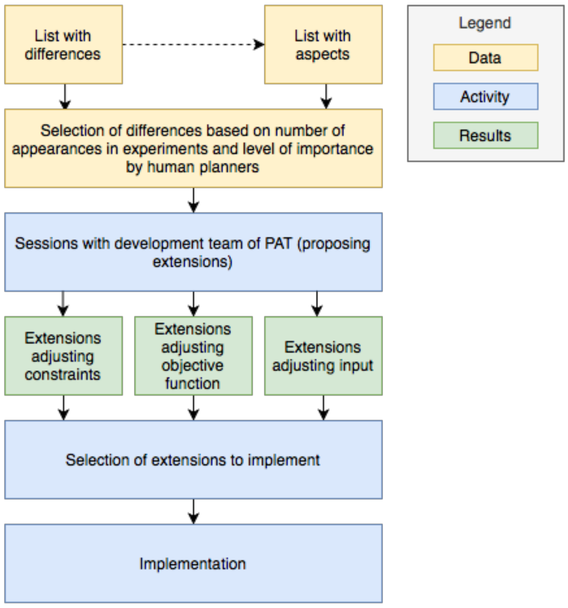

Figure 2: The process of defining, selecting and implementing extensions based on the list of differences and aspects defined by the empirical experiment

5.1. Case study

In order to execute the experiment and the selection and implementation of extensions, we used a case study focused on the planning assignments of the node Eindhoven in the Netherlands. Eindhoven is a relatively large node that consist of a large station with 6 platforms and two separated Service Locations (see Figure 3). A planning assignment at Eindhoven consist on average of the planning of 218 trains per 24 hours. Such a planning assignment consist out of six sub-problems: the matching of incoming and outgoing trains, the job shop scheduling for non-human resources (servicing like interior cleaning, mechanically checking and exterior washing of the train), parking, routing, splitting and combining in order to create the wanted outgoing train combinations and crew assignment. The combination of the large planning assignment and the many sub problems that are part of the planning assignment makes Eindhoven a perfect example of a complex planning assignment. In total, six different real-life planning assignments were used of Eindhoven. The characteristics of the planning assignments are presented in Table 2.

The research group that was used consisted out of 5 planners that are

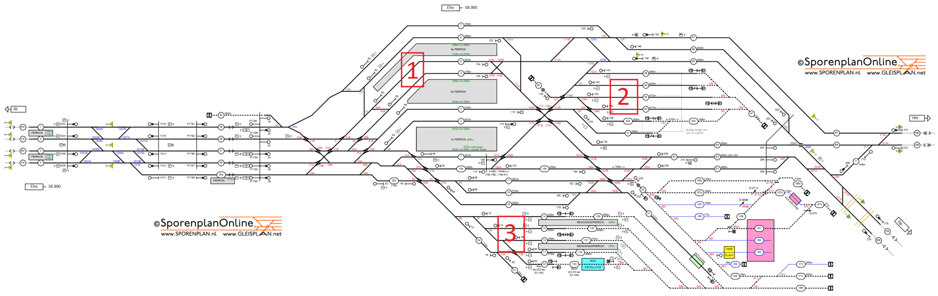

Figure 3: Layout of Eindhoven. The red 1 presents the station with 6 platforms, the 2 and 3 are the seperated Service Locations (Sporenplan Online, 2019)

Table 2: Characteristics of the different real-life planning assignments used. Ongoing trains are trains that call at the station, but do not end the passenger moving service.

Indicators	**A1**	**B**	**A2**	**C**	**D**	**E**
Date	2-3 Jan	5-6 Dec	30-31 Dec.	26-27 Feb	24-15 Feb	6-7 Feb
Sum of train units	191	214	203	237	249	215
Arriving trains	148	148	138	152	154	144
Train units per arriving train	1,29	1,45	1,47	1,56	1,62	1,49
Departing trains	154	156	147	161	163	148
Train units per departing train	1,24	1,37	1,38	1,47	1,53	1,45
Tracks out of service	0	0	0	13	0	0
Ongoing trains	536	534	512	547	556	511
Level of difficulty	Easy	Medium	Medium	Hard	Hard	Medium

responsible for the creation of the planning from four weeks before operation up and until 56 hours before operation. This planning department was selected because the level of detail needed within the plan is high, while the solution space that planners have is the smallest of all planners. The planners that took part of the experiment were all experienced, which is wanted in order to collect information with a high level of quality.

The PAT used for this research is called HIP (Hybrid Integral Planning method). The tool can be used to solve all the sub problems of a node logistics planning assignment except crew assignment. HIP is a unique tool for node logistics, because it integrates the complete node: ongoing traffic as well as traffic between the station and the Service Locations are planned as well. The optimization of the plan is performed using a local search approach. HIP is a multi-objective problem solver. The objectives are solved sequentially: first minimizing conflict costs and then penalties. The solution of HIP is made visible using a Graphical User Interface (GUI). The GUI also presents the list of conflicts that are still present in the plan. The running time of the PAT in order to generate a solution is 10 minutes (600 seconds).

5.2. Results empirical experiment

From the experiments 59 differences were identified between the human plan and the PAT plan. The observed differences could be categorized in six aspects: simplification of the planning assignment, capacity optimization, increasing flexibility, optimizing plan for KPIs, norms and agreements and safety. It is possible that certain differences relate to multiple aspects.

The simplification of the planning assignment is necessary for human planners to structure the big planning assignment into smaller parts. By making rational choices, the planning assignment is simplified because the number of needed actions is decreased. The rational choices make the solution space for human planners smaller, because the amount of possibilities is decreased.

The capacity optimization is described as the best usage of the of infrastructure and resources. For example, certain combinations of trains are preferred in order to fill a parking track to its maximum or limited movements create more freedom of movement for other trains.

By increasing flexibility, planners try to create a plan that takes into account the effects of disruptions. The choices made by planners are subordinate to choices that ease the planning assignment.

The optimization of the plan for KPIs means that a planner makes certain choices because they are expected to lead to better service in general. KPIs can be train related or passenger related. An example of a train related KPI is the amount of clean trains that start a passenger service.

The norms and agreements relate to rules that the planning department has with other planning departments or dispatchers. Certain norms or agreements are very strict, while other are more flexible.

Finally, safety is an aspect that is very important for planners. Safety contains traffic safety but also safety to the operational crew members. General safety regulations are already part of the PAT, but node-specific information is missing, such as tracks where train drivers can exit or enter a train safely. These rules limit the solution space of the PAT because more constraint limit the amount of choices that can be made.

5.3. Selection of extensions

Based on the six aspects and the list of differences, 37 extensions were proposed in cooperation with the development team of the PAT. From these 37 extensions, 20 were selected to be implemented. The total list of extensions and whether they are selected can be found in Table 6.1 in the main report.

The extensions can be grouped into three groups. First, there are extensions adjusting the constraints of the PAT. These adjustments have a direct effect on the solution space of the PAT since they reduce or increase the space with the adjustments of constraints. An example of such an extension is the change of parking constraints.

Secondly, there are extensions that adjust the objective function of the PAT. These extensions affects the optimization function by adding an extra element. It is important that the difference created by this addition can be measured in order to identify the differences between the original and adjusted PAT. The addition of parking weights is the only example of an extension that adjust the objective function in the case-study.

Thirdly, there are extensions that adjust the input of the PAT. These extensions have a direct impact on the initial solution space, since this information describes the initial boundaries of the solution space. A good example of such an extension is a change in the layout.

6. Evaluation of adjustments to PAT

In order to find out what the impact of the extensions to the PAT is and what the reaction of the human planners is to the plan of the adjusted PAT, three comparisons are performed that are presented in Figure 4. With a real-life scenario, a qualitative comparison is performed in order to capture the reaction of the planners of the solution of the adjusted PAT. The solution of the real-life planning assignment of the PAT were presented to the same experienced planners that were used for the experiment. They compared the solution of the adjusted PAT with the solution they have created for the same planning assignment. A quantitative comparison could compare the performance of the PAT plan in relation to the original PAT for two different planning assignments: real-life and theoretical. The theoretical planning assignments are synthetic planning assignments that are created based upon a real-life planning assignment. The theoretical assignments differ between themselves in a level of complexity.

First, the results of the qualitative comparisons are presented. The overall opinion of the planners was that the adapted PAT produced a better solution compared to the original PAT in two ways. First there were less irrational choices made in the solution of the adjusted PAT. Second there was a higher level of resemblance between the human solution and the solution of the PAT. Planners were better able to understand the solution of the adjusted PAT, as well as the motivation behind certain choices the PAT made. On a more node-specific detail, a lot of improvements were notified

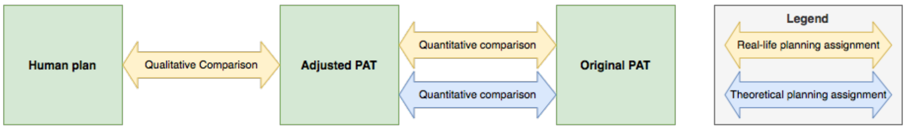

Figure 4: The comparisons made in order to test the adjusted PAT. The comparison between the human plan and the adjusted PAT is based on real-life planning assignments and has to goal to capture the qualitative differences. The two comparisons between the adjusted PAT and the original PAT have the goal to capture the quantitative differences based on real-life planning assignments and theoretical planning assignments.

by the planners regarding the parking and servicing process on the node. Whereas in the initial experiment planners reacted mostly with statements that the solution of the original PAT was not possible in real life, the evaluation with the adjusted PAT led to a lot of questions regarding certain choices the adjusted PAT made. This change indicates the increase of a willingness to understand. This willingness can be translated to a higher level of acceptance.

In order to quantify differences between HIP plans, three indicator groups are defined: conflict costs, penalty costs and candidates checked. The **conflict costs** indicates the costs of conflicts of the plan. The conflict costs score consist of number of crossings, delayed departures, delayed arrivals, track length violations, incorrect combination and delayed fixed trains. The **penalty costs** consist of the number of shunt movements and the illegal parking time. **Candidates checked** indicates the amount of candidate solutions that are checked by the HIP in order to optimize the plan. If the same running time is used and the number of candidates checked has increased, it is positive for the solution approach of the algorithm. In other words, the plan assignment might be a little less complex, since more candidates were able to be checked within the same time limit. The number of candidates is not part of one of the two costs functions.

The quantitative comparison performed on the real-life planning assignments are presented in table 3. The average conflict costs decreases generally with 19% when the adjusted HIP is used. The average penalty costs increases with 29% when the adjusted HIP is used. However, it has to be noticed that the penalty costs of planning assignment A1 increases with a relatively large number compared to the other. The average of the penalty costs of all planning assignment except A1 is -18%. The reason why the penalty costs of A1 rises with such a high number is because the penalty costs of the original HIP solution are very low. Based on an analysis of all penalty costs, it can be stated that the variety of penalty costs between different runs (seeds) of the HIP is high. More runs will result in a more stable average of the penalty costs. The number of candidates checked decreased for the adjusted HIP, which means the adjusted HIP investigates less candidates for the local search optimum. It can be stated that for real-life planning assignments, the adjusted HIP performs better compared to the original HIP.

Indicators	**A1**	**B**	**A2**	**C**	**D**	**E**	**Average**
Conflict costs	-16%	-90%	-13%	35%	1%	-31%	**-19%**
Penalty costs	262%	-23%	27%	-39%	-26%	-29%	**29%**
Candidates checked	-26%	12%	-1%	12%	-4%	4%	**-1%**

Table 3: Differences between adjusted HIP and original HIP on real-life planning planning assignments. The six different planning assignments are presented in the columns, and the average column at the end represents the average of these six planning assignments. A positive percentage means that the values of the adjusted HIP are higher compared to the original HIP.

The quantitative comparison performed on the theoretical planning assignments are presented in table 4. A notice has to be made that the running time of these planning assignment was limited to 2 minutes (120 seconds). The theoretical planning assignments increase in level of complexity from 1 to 6, where the level of complexity of 6 resembles a simple real-life planning assignment. The conflict costs decreases with 31% if the adjusted HIP is used. However, the penalty costs increase with an extremely large number, especially for the less complex planning assignments. This is due to the fact that the illegal parking time violations increases if the adjusted HIP is used (1474% on average over these six theoretical planning assignments). However, as stated above, the variety of the penalty costs is large, meaning that it is hard to interpret something based on this indicator. There is no significant difference between the amount of candidates checked.

Indicator	**1**	**2**	**3**	**4**	**5**	**6**	**Average**
Conflict costs	0%	-31%	-40%	-27%	-43%	-44%	**-31%**
Penalty costs	0%	3894%	534%	2874%	1602%	773%	**1613%**
Candidates checked	-9%	-1%	1%	3%	3%	2%	**0%**

Table 4: Differences between the results of the adjusted HIP and the original HIP. A positive percentage means that the values of the adjusted HIP are higher compared to the original HIP. A negative percentage means that the values of the adjusted HIP are lower compared to the original HIP.

7. Conclusion

Within the conclusion, the main research questions is answered: What does the addition of missing aspects based on human planning observations do to the solution of a Planning Assistance Tool? This questions is answered by describing the solutions of the two objectives stated in the Problem description. First of all, the empirical experiment in a field study set-up using real-life planning assignments and experienced planners proved to be a successful way to determine which aspects are missing in a PAT. The literature revealed that such an intensive experiment set-up was never used in order to identify the human factors or aspects. The experiment is however time consuming since the level of detail needed requires that the researchers have a lot of knowledge about the planning environment. It can be stated that that the objective to determine which aspects are missing in Planning Assistance Tools for complex planning processes has been achieved.

The extensions made to the PAT in order to cover the missing aspects identified by the empirical experiments do have a positive impact on the solution of the PAT, both in a qualitative and quantitative manner. The qualitative differences are relatively more significant than the quantitative differences. In other words, the implementations to the PAT do affect the level of quality more compared to the quantitative indicators. From a qualitative perspective, the proposed solution better resembled the human solution. The human planners had the feeling to understand the solution and the choices made by the adjusted PAT better compared to the original PAT solution. This indirectly provides a base of trust that is needed for planners to accept the solutions of the Planning Assistance Tool. The increase of interest to the way the PAT is solving planning assignments and the increase of resemblance between the human solution and the PAT solution can be directly related to a higher level of acceptance of the PAT by human planners. From a quantitative perspective, the differences in performance were small compared to the original PAT regarding real-life scenarios. The average results showed an overall positive outcome of the implementation of the aspects. In summary, it can be stated that the second objective is achieved as well: to find out what the implementation of aspects based on human planning observations in the PAT does to the solution of the PAT and the acceptance of human planners.

The relevance of this work is important for industries that are concerning the use of planning tools in order to make the planning process more efficient. This research has proven that the identification of human aspects and the implementation of extensions based on the aspects has a positive effect on the acceptance of the tool by human planners. Within all production industries, a lot of money has been spent to expensive development of planning tools, but the amount of successful implementations of tools within the planning process is low because human planners did not accept the tool. It is needed to admit that within critical planning processes, human planners will always stay in charge of controlling the complete process and there will always be a need for human actions in case computer systems fail. Therefore, it is crucial that human planners are involved in the development of planning tools as soon as possible. If tools are already created, the approach described in this research might be a fast and efficient way to include human planners within the process and to increase the level of acceptance by human planners.

However, there are some marks to make regarding the results of this research. First of all, the comparison with the theoretical planning assignments have showed that the implementations added to the PAT are clearly related to the level of complexity of the planning assignments. If aspects are identified using complex planning assignments, the implementations to cover these aspects perform best if planning assignments with the same level of complexity are used.

Build upon this, the current research focus within this field is often too much on the creation of PAT and the way the solutions are presented. However, this research has proven that characteristics of the PAT has to be build upon the characteristics of the input (in this case, the planning assignment). More research to the characteristics of planning assignments and the differences between them is necessary.

Thirdly, the scope of this research is extremely specific. It is necessary to perform this research on other nodes that differ (but also on a node that resembles Eindhoven) in order to find out if the same type of aspects are identified and if the same type of extensions can be added to the PAT. Based upon the current difference of planning departments between nodes, an assumption is made that it will be hard to determine generality between different nodes.

Fourthly, it has to be stated that this research aimed to create a solution of the PAT that resembles a human solution of a complex planning assignment. This does not mean that the human solution is the best solution that can created. It is by far more logical to expect that a computer is able to create better solutions of complex planning assignment in less time. However, before optimization processes can be applied for computers, it is of course necessary that the solution provided by the PAT is accepted by the planners in order to integrate it. Besides, there are no general indicators that determine the level of performance of solutions to planning assignments. These often differ from planning department to planning department and between industries as well.

References

Arica, E., Strandhagen, J., & Hvolby, H.-H. (2014). Designing a Decision Support System for Production Scheduling Task in Complex and Uncertain Manufacturing Environments, . (pp. 589–596). URL: https://hal.inria.fr/hal-01388601. doi:10.1007/978-3-662-44739-0{_}72{\"{i}}.

Berglund, M., & Karltun, J. (2007). Human, technological and organizational aspects influencing the production scheduling process. *International Journal of Production Economics*, *110*, 160–174. doi:10.1016/j.ijpe.2007.02.024.

Cegarra, J., & Wezel, W. (2011). A Comparison of Task Analysis Methods for Planning and Scheduling. In *Behavioral Operations in Planning and Scheduling* (pp. 323–338). doi:10.1007/978-3-642-13382-4{_}13.

Coman, A., & Muñoz-Avila, H. (2011). Qualitative vs. Quantitative Plan Diversity in Case-Based Planning, .

Cordeau, J. F., Toth, P., & Vigo, D. (1998). A survey of optimization models for train routing and scheduling. *Transportation Science*, *32*, 380–404. doi:10.1287/trsc.32.4.380.

Dagblad van het Noorden (2019). Groningen krijgt 10 kilometer nieuw spoor op rangeerterrein De Vork, .

De Man, J. C., Nehzati, T., Arica, E., & Kiil, K. (2016). Automating scheduling tasks; Experience of a real case. In *IEEE International Conference on Industrial Engineering and Engineering Management* (pp. 1447–1451). IEEE Computer Society volume 2016-January. doi:10.1109/IEEM.2015.7385887.

Dios, M., & Framinan, J. M. (2016). A review and classification of computer-based manufacturing scheduling tools. *Computers and Industrial Engineering*, *99*, 229–249. doi:10.1016/j.cie.2016.07.020.

Framinan, J. M., & Ruiz, R. (2012). Guidelines for the deployment and implementation of manufacturing scheduling systems. doi:10.1080/00207543.2011.564670.

Hoc, J. M. (2000). From human – machine interaction to human – machine cooperation. *Ergonomics*, *43*, 833–843. doi:10.1080/001401300409044.

Käki, A., Kemppainen, K., & Liesiö, J. (2019). What to do when decision-makers deviate from model recommendations? Empirical evidence from hydropower industry. *European Journal of Operational Research*, *278*, 869–882. doi:10.1016/j.ejor.2019.04.021.

Lee, J. D., & Seppelt, B. D. (2009). Human Factors in Automation Design. In *Springer Handbook of Automation* (pp. 417–436). Springer Berlin Heidelberg. doi:10.1007/978-3-540-78831-7{_}25.

Lorentzon, J., & Fredlund, J. (2017). *Application of Performance Measurement on Manufacturing Simulations for Knowledge-Based Decision Support*. Technical Report.

McKay, K., Buzacott, J., & Safayeni, F. (1989). The scheduler's knowledge of uncertainty: The missing link. *Knowledge Based Production Management Systems*, .

McKay, K. N., & Wiers, V. C. (2003). Integrated decision support for planning, scheduling, and dispatching tasks in a focused factory. *Computers in Industry*, *50*, 5–14. doi:10.1016/S0166-3615(02)00146-X.

McKay, K. N., & Wiers, V. C. (2006). The human factor in planning and scheduling. In *International Series in Operations Research and Management Science* (pp. 23–57). Springer New York LLC volume 89.

Mebarki, N., Cardin, O., & Guerin, C. (2013). Evaluation of a new human-machine decision support system for group scheduling. In *IFAC Proceedings Volumes (IFAC-PapersOnline)* (pp. 211–217). volume 12. doi:10.3182/20130811-5-US-2037.00040.

Moscoso, P. G., Fransoo, J. C., & Fischer, D. (2010). An empirical study on reducing planning instability in hierarchical planning systems. *Production Planning and Control*, *21*, 413–426. doi:10.1080/09537280903454172.

Myers, K. L., & Lee, T. J. (1999). *Generating Qualitatively Different Plans through Metatheoretic Biases*. Technical Report. URL: www.aaai.org.

NRC (2018). ProRail verbouwt voor reizigersgroei.

NS (2016). Aanschaf nieuwe Intercity's. URL: https://nieuws.ns.nl/aanschaf-nieuwe-intercitys/.

Riezebos, J., Hoc, J. M., Mebarki, N., Dimopoulos, C., Van Wezel, W., & Pinot, G. (2011). Design of scheduling algorithms: Applications. In *Behavioral Operations in Planning and Scheduling* (pp. 371–412). Springer Berlin Heidelberg. doi:10.1007/978-3-642-13382-4{_}15.

Sanderson, P. (1989). The Human Planning and Scheduling Role in Advanced Manufacturing Systems: An Emerging Human Factors Domain. *Human Factors*, *31*.

Sporenplan Online (2019). Sporenplan Online. URL: http://www.sporenplan.nl.

Treinreiziger (2018). NS en ProRail: het spoor is bijna vol. URL: https://www.treinreiziger.nl/ns-en-prorail-spoor-is-bijna/.

van Uden, H., Smolders, D., Oosterbaan, P., Verheugd, R., Rol, R., & Snijders, J. (2017). *De beste manier van knooppuntplanning*. Technical Report NS.

Van Wezel, W., & Jorna, R. (2009). Cognition, tasks and planning: Supporting the planning of shunting operations at the Netherlands Railways. *Cognition, Technology and Work*, *11*, 165–176. doi:10.1007/s10111-007-0106-8.

Verlaan, J. (2019). NS: grote problemen dreigen op spoor en stations.

B
Interview protocol

The interview protocol underneath is used in order to structure the interview that took place when comparing the human solution and the solution of the PAT (Hip in this case) of a planning assignment. Since all the employees of the NS speak Dutch, the interview protocol is also presented in Dutch. An English translation of the protocol is also added to the appendix.

Interview protocol voor vergelijkingsanalyse

Duur: 75 minuten

Benodigdheden: Plan van planner, plan van HIP

Introductie (5 minuten):

Allereerst bedankt voor jouw medewerking en tijd. Zoals je waarschijnlijk weet ben ik benieuwd wat de verschillen zijn tussen jouw planning en een uitkomst van een planning gemaakt door een computer. Het doel van dit onderzoek is om deze verschillen aan te duiden en de oorzaak ervan te vinden. Het programma van de computer waarin dit gedaan wordt, heet HIP, dus vandaar zal ik deze naam vanaf nu gebruiken. Aangezien HIP zelf de planning maakt en deze verbetert, kan ik weinig zeggen over waarom HIP bepaalde keuzes heeft gemaakt, behalve dat het in de ogen van een algoritme een betere oplossing produceert. Ik ben benieuwd naar wat jij van de oplossing van HIP vindt, aangezien jij veel ervaring hebt met het maken van planningen op deze knoop. Zoals je misschien ook al weet, is er heel veel kennis die niet in boeken staat, en is er veel kennis welke dynamisch (over tijd) is veranderd. Jij bent als planner een expert, en het is daarom dat ik graag van jou wil leren!

Daarnaast wil ik zeggen dat ik van plan ben om dit interview ongeveer 12 keer af te nemen. Hierbij zal ik alle respondenten anoniem vermelden in mijn onderzoek. Het is mogelijk dat ik iemand twee keer zal benaderen. In dat geval zal ik de respondent op dezelfde anonieme manier aanduiden met een ander volgnummer (dus 2 voor de tweede keer, eventueel 3 voor de derde keer).

Dit onderzoek zal bestaan uit een beginsectie van 15 minuten waarin ik benieuwd ben wat jij van je eigen planning vindt en waarin we de eigenschappen van het planvraagstuk aanduiden. In het middengedeelte zullen we ingaan op de verschillen tussen jouw planning en HIP. Het middengedeelte zal ongeveer 45 minuten duren. In het laatste gedeelte zullen we conclusies gaan trekken. Dit zal 10 minuten duren.

Evaluatie planvraagstuk (15 minuten)

1. Hoe vond je het planvraagstuk?
    a. Was het moeilijker of makkelijker dan normaal?
    b. Wat waren de lastige punten in een planning?
2. Ben je meer of minder tijd kwijt geweest om deze planning te maken dan normaal?
    a. Hoeveel tijd precies? Welke delen / problemen kostten meer tijd dan normaal?
3. Hoeveel conflicten waren er die je moest oplossen en is dit meer of minder dan normaal?
    a. Waren er conflicten die lastig waren?
        i. Zo ja, kan je deze beschrijven?
        ii. Hoeveel?
    b. Heb je enig idee hoe deze conflicten in de planning zijn gekomen?
    c. Waren er conflicten die makkelijk waren?
        i. Zo ja, kan je er één van beschrijven?
        ii. Hoeveel?
4. Is de planning nu helemaal conflict vrij?
5. Wat zijn de sterke punten van jouw plan? Welke punten heb je voldoende gedekt met jouw plan? [Zie tabel 1 voor mogelijke doelstellingen van planners]

6. Welke punten zouden extra aandacht kunnen krijgen in jouw plan? Als je tijd over zou hebben, naar welke punten zou je kijken om je plan te verbeteren? [Zie tabel 1 voor mogelijke punten]

Plan vergelijking (45 minuten)

We gaan nu jouw planning vergelijken met de planning die geproduceerd is door HIP. HIP heeft hetzelfde beginprobleem gekregen, echter is een perfecte translatie tussen Donna en HIP niet mogelijk, dus zijn er een aantal dingen aangenomen. Deze aannames zijn:

- *Leeg-materieel-ritten worden niet naar vastgestelde sporen gepland, maar komen via de vrije baan de knoop binnen. HIP mag zelf bepalen waar deze trein op de Service Locatie wordt geplaatst. Want als knoopplanner kun je dit aankomst- (of vertrek)spoor ook zelf aanpassen.*
- *Elke trein wordt voor een vaste tijd aangeboden (per materieel type -> check dinsdag met planner) voor een technische check en interne reiniging.*
- *HIP mag treinen ver van tevoren voorbrengen (parkeren) op het station, zolang deze trein geen andere bewegingen in de weg zit.*
- *HIP houdt geen rekening met personeel. Dit houdt in dat er geen personeelsplanning gemaakt wordt.*

We gaan nu chronologisch door de twee planningen lopen om te kijken welke verschillen er zijn. Ik kijk naar verschillen, maar geef ook aan als jij er één ziet!

1. Bij een verschil:
    a. Onder welke noemer kunnen we het verschil plaatsen? [zie tabel 2]
    b. Is de optie van HIP mogelijk? Mag hetgeen wat HIP doet?
        i. Als de oplossing van HIP niet mag, waarom niet?
    c. Is de optie van HIP wenselijk? Zou jij ook deze optie gekozen kunnen hebben?
        i. Zo ja, waarom wel?
        ii. Zo nee, waarom niet?
    d. Als we kijken naar de oplossing die jij voorstelt in jouw plan, waarom heb jij gekozen voor deze oplossing?
        i. Wat waren de andere opties?
        ii. Zijn al deze opties wenselijk? Waarom wel of waarom niet?
        iii. Heb je het gevoel dat je de beste keuze hebt gemaakt?
    e. Wat vind jij van de optie van HIP?
        i. Negatief: Waarom?
        ii. Positief: Waarom?
2. Wat zijn de overeenkomsten tussen jouw planning en de planning van HIP?

Conclusie (10 minuten)

Nu we de twee planningen helemaal zijn nagelopen, is het tijd om de balans op te maken en enkele dingen te concluderen. Hiervoor gaan we nog 10 minuten enkele vragen bespreken!

1. Wat is jouw eerste indruk van HIP?
    a. Positieve punten?
    b. Negatieve punten?
    c. Denk je dat er potentie in zit om deze tool te gebruiken voor (delen van) het planvraagstuk? Waar denk je dat HIP nog in moet verbeteren?

2. Wat vind je van de visualisatie van de oplossing?
    a. Welke vertoningen vind jij fijn?
    b. Welke dingen kunnen er nog worden toegevoegd aan de visualisatie?
3. We hebben de volgende verschillen tussen de plannen waargenomen: [lijst met verschillen]. Ben ik er nog een vergeten?
    a. Zou je deze verschillen kunnen groeperen?
4. De volgende belangen (aspecten) hebben de verschillen (deels) veroorzaakt: [lijst met belangen (aspecten) van de verschillen]. Ben ik er nog een vergeten?
    a. Zou je deze belangen (aspecten) kunnen groeperen?
        i. Eigen groepering
        ii. Tussen knooppunt specifiek en generiek (dus toepasbaar op elke knoop)
        iii. Welke belangen (aspecten) vind zijn het belangrijkst voor de kwaliteit van het plan?
5. Vanaf jouw perspectief, wat zou het belangrijkst zijn om nu te moeten verbeteren aan HIP?
6. Zijn er nog dingen die je zelf kwijt wil over jouw planning, de planning van HIP of het algehele plan proces?

Nogmaals bedankt voor jouw input! Aangezien de knooppunt planning erg verschilt tussen knopen, en kennis over deze precieze knopen zo belangrijk is, is jouw expertise cruciaal in dit onderzoek. Daarvoor wil ik je enorm bedanken. Ik ga zo nog voor mezelf dit interview evalueren, en zal dan een samenvatting naar jouw toe sturen! Het zou fijn zijn als je deze zou kunnen checken, voor de zekerheid. De input van jouw en van je collega's zal worden gebruikt om HIP te verbeteren. We willen extra aspecten toevoegen om te kijken of de planning van HIP makkelijker aansluit bij de planning van planners. Ik heb echter geen idee of dit daadwerkelijk zal lukken. Echter zou ik wel graag nog een keer met je willen zitten om jouw mening over de nieuwe variant van HIP weer te registreren. Daarom zal ik binnenkort weer contact met je opzoeken, en zal ik het enorm waarderen als je dan misschien weer even zou willen kijken naar de vernieuwde versie van HIP. Als er nog dingen zijn, die je nog kwijt wilt, kan dat natuurlijk nu, maar ook nog per mail of in het vervolg als ik weer hier ben! En als je nog vragen hebt over mijn onderzoek in het algemeen, kan je dat natuurlijk altijd vragen! Ik houd je in ieder geval op de hoogte, tenzij je dat niet meer wenst natuurlijk!

Mogelijke doelstelling	Mogelijke antwoorden planner
Robuustheid verbeteren [Duidelijk aangeven wat de planner ziet als robuust]	Doorgang naar service locatie altijd vrij
	Elk type materieel binnen minimale aantal bewegingen beschikbaar
	Wisselveranderingen minimaliseren
	Riskante bewegingen (i.v.m. potentiele verstoringen) minimaliseren
	Aankomst of vertrek perron wijzigen om plan robuuster te maken
Capaciteit optimaal gebruiken	Parkeer- en servicesporen maximaal gebruiken
	Treinen met service dicht bij elkaar zetten
	Aanbieden van 'service tijd' (tijd dat service verleend kan worden aan een treinstel) maximaliseren.
	Toegangsmogelijkheden (i.v.m. met doorgaande bewegingen) optimaal gebruiken

	Het specificeren van exacte parkeerplek op een spoor om de capaciteit van dit parkeerspoor zo goed mogelijk te gebruiken.
Inzet van machinisten minimaliseren	Minimaal aantal machinisten inplannen
	Plan maken op minimale looptijd machinisten
	Meedenken met de routes voor machinisten (meereizen met andere treinen, etc.)
Rijtijd minimaliseren	Kortste route kiezen
	Dichtstbijzijnde parkeerplek kiezen
	Aankomst of vertrek perron wijzigen om routes te verkorten
Opstelmethode handhaven	Treinen opstellen zodat ze snel weer in de dienst gebracht kunnen worden (treinen die als eerst gebruikt moeten worden dichtbij (of via een makkelijke route) naar de perrons.
	Treinen materieel specifiek opstellen
	Treinen zo snel mogelijk kunnen af-rangeren (dus treinen via de makkelijkste route of op de dichtstbijzijnde plek parkeren)
	Treinen op perron parkeren en behandelen.
Aankomst of vertrekperron wijzigingen minimaliseren	Minimale wijzigingen met het originele plan
Aantal bewegingen minimaliseren	Treinen op parkeer & service sporen parkeren
	Treinen eerder laten combineren (of extra laten combineren, als dezelfde route wordt afgelegd)
	Bestaande routes aanpassen om (zaag)bewegingen te verminderen.
	Aankomst of vertrek perron wijzigen om bewegingen te minimaliseren

Tabel 1 Lijst met mogelijke doelstellingen van planners en mogelijke antwoorden van planners

Verschil factoren	Mogelijke verschillen
Totale opstelling	Treinen staan anders verdeelt op de knoop
	Services worden anders aangeboden op de knoop
Parkeren	Een trein staat op een ander spoor geparkeerd
	Een trein staat op meerdere verschillende sporen geparkeerd
Beweging	De routes verschillen
	De begin tijdstippen van een beweging verschillen
Matching	Treinen worden anders gematched (inkomende treinen worden aan uitgaande treinen gekoppeld)
Service toewijzing	Treinen worden een ander service spoor toegewezen
	Treinen worden op een ander moment aangeboden voor service
Splitsen en combineren	Treinen worden op een ander moment gecombineerd of gesplitst
Personeel toewijzing	Personeel wordt ander toegewezen

Tabel 2 Factoren die kunnen verschillen en voorbeelden van mogelijke verschillen

Interview protocol for comparative analysis

Duration: 75 minutes

Supplies: Plan of planner, plan of HIP

Introduction (5 minutes):

First of all thanks for your cooperation and time. As you probably know, I'm curious to know what the differences are between your planning and an outcome of a planning made by a computer. The purpose of this study is to identify these differences and to find the cause of them. The application of the computer in which this is done is called HIP, so I will use this name from now on. Since HIP makes the planning itself and improves it, I cannot provide information about why HIP has made certain choices, except that it produces a better solution in the eyes of the algorithm. I am curious to hear what you think of HIP's solution, since you have a lot of experience with planning on this node. As you may already know, there is a lot of knowledge that is not in books, and there is a lot of knowledge that has changed dynamically (over time). As a planner, you are an expert, and that is why I would like to learn from you!

In addition, I would like to say that I plan to do this interview about 12 times with different respondents. I will mention all respondents anonymously in my survey. It is possible that I will approach someone twice. In that case I will indicate the respondent in the same anonymous way with a different sequence number (i.e. 2 for the second time, possibly 3 for the third time).

This survey will consist of a starting section of 15 minutes in which I am curious what you think of your own planning and in which we indicate the characteristics of the plan assignment. In the middle section we will discuss the differences between your planning and HIP. The middle part will take about 45 minutes. In the last part we will draw conclusions. This will take 10 minutes.

Evaluation plan issue (15 minutes)

1. *What did you think about the plan assignment?*
    a. *Was it harder or easier than usual?*
    b. *What were the hard parts of the planning?*
2. *Did you spend more or less time making the plan than normal?*
    a. *How much time? Which parts / problems took more time than normal?*
3. *How many conflicts did you have to solve and is this more or less than normal?*
    a. *Were there conflicts that are difficult?*
        i. *If so, can you describe them?*
        ii. *How many?*
    b. *Do you have any idea how these conflicts appeared in the planning?*
    c. *Were there conflicts that were easy?*
        i. *If so, can you describe any of them?*
        ii. *How many?*
4. *Is the planning completely conflict-free now?*
5. *What are the strengths of your plan? Which points have you covered sufficiently with your plan? [See table 1 for possible objectives of planners]*
6. *What points could get extra attention in your plan? If you had extra time, what points would you look at to improve your plan? [See table 1 for possible points]*

Plan comparison (45 minutes)

We are now going to compare your planning with the planning produced by HIP. HIP had the same initial problem, however a perfect translation between Donna and HIP is not possible, so a number of things have been assumed. These assumptions are:

- *Empty-material trips are not planned to fixed tracks, but enter the node via the open track. HIP may decide for itself where this train will be placed on the Service Location. Because as node planner you can also modify this arrival (or departure) track yourself.*
- *Each train is offered for a fixed time (per rolling stock type -> check Tuesday with planner) for a technical check and internal cleaning.*
- *HIP is allowed to park trains at the platform far in advance, as long as this train does not get in the way of other movements.*
- *HIP does not take staff into account. This means that no personnel planning is made.*

We will now go through the two schedules chronologically to see what differences there are. I look for differences, but also indicate if you see one!

1. *If there is a difference:*
    a. *Under what difference factor can we categorize the difference? [see table 2]*
    b. *Is the HIP option technical possible and feasible? Is what HIP proposed allowed?*
        i. *If the solution of HIP is not allowed, why not?*
    c. *Is the HIP option desirable? Could you have chosen this option as well?*
        i. *If so, why?*
        ii. *If not, why not?*
    d. *If we look at the solution you propose in your plan, why did you choose this solution?*
        i. *What were the other options?*
        ii. *Are all these options desirable? Why or why not?*
        iii. *Do you feel that you have made the best choice?*
    e. *What do you think of the HIP option?*
        i. *Negative. Why?*
        ii. *Positive: Why?*
2. *What are the similarities between your planning and the planning of HIP?*

Conclusion (10 minutes)

Now that we have gone over the two plans completely, it is time to take stock and conclude a few things. For this we are going to discuss some questions for another 10 minutes!

1. *What is your first impression of HIP?*
    a. *Positive points?*
    b. *Negative points?*
    c. *Do you think there is potential to use this tool for (parts of) the plan issue? What do you think HIP should improve on?*
2. *What do you think of the visualization of the solution?*
    a. *Which parts of the visualization do you like?*
    b. *Which things can still be added to the visualization?*
3. *We have observed the following differences between the plans: [list of differences]. Have I forgotten one?*
    a. *Could you group these differences?*

4. *The following aspects have (partly) caused the differences: [list of causes of the differences]. Have I forgotten one?*
    a. *Could you group these aspects?*
        i. *Own grouping*
        ii. *Between node specific and generic (i.e. applicable to each node)*
        iii. *What aspects are most important for the quality of the plan?*
5. *From your perspective, what would be the most important thing to improve HIP now?*
6. *Are there still things you want to say about your planning, the planning of HIP or the overall planning process?*

Again thanks for your input! Since node planning is very different between nodes, and knowledge about these exact nodes is so important, your expertise is crucial in this research. For that I would like to thank you very much. I'm about to evaluate this interview for myself, and I'll send you a summary! It would be nice if you could check it, just to be sure. The input of you and your colleagues will be used to improve HIP. We want to add extra aspects to see if the planning of HIP is more in line with the planning of planners. However, I have no idea if this will actually work. In order to find out I would like to sit down with you again to register your opinion and thoughts about the new version of HIP. That is why I will contact you again soon, and I would really appreciate it if you would like to take another look at the new version of HIP. If there are still things you want to say, you can do that now, but also by email or in the future when I'm back here! And if you have any questions about my research in general, you can always ask me! I will keep you informed, unless you don't want it anymore!

Possible objectives	Possible answers planner
Improve robustness [Clearly state how the planner describes robustness]	Routes to service location always free
	Any type of equipment available within minimum number of movements
	Minimize switch changes
	Minimize risky movements (due to potential disturbances)
	Change arrival or departure platform to make plan more robust
Optimal use of capacity	Maximum use of parking and service tracks
	Placing trains with service close together
	Maximize 'service time' (time that service can be provided to a trainset).
	Make optimal use of access possibilities (due to ongoing train movements).
	Specifying the exact parking spot on a track in order to make the best use of the capacity of this parking track.
Minimize train-operator deployment	Minimum number of drivers to be scheduled
	Plan at minimum walking time for drivers
	Thinking along with the routes for train drivers (travelling with other trains, etc.)
Minimising driving time	Selecting the shortest route
	Choose the nearest parking space

	Change arrival or departure platforms to shorten routes
Maintain certain parking method	Park trains so that they can be quickly brought back to the platform (trains that must be used first near (or via an easy route to) the platforms.
	Park train based on rolling stock equipment
	Park trains as fast as possible (i.e. park trains on the easiest route or at the nearest spot).
	Parking and servicing of trains at platform.
Minimize arrival or departure platform changes	Minimize changes with the original plan
Minimize number of movements	Park trains on parking & service tracks
	Combine trains earlier (or combine them extra, if the same route is used)
	Adjust existing routes to reduce (saw-) movements.
	Modify arrival or departure platform to minimize movements

Tabel 1 Possible objectives and possible answers of planners

Difference factors	Possible differences
Different arrangement of trains at node	Trains are distributed differently over the node
	Services are offered differently over the node
Parking	A train is parked at another track
	One train is parked at several different tracks
Movement	Routes differ
	The starting times of a movement differ
Matching	Trains are matched differently (incoming trains are coupled to outgoing trains)
Service allocation	Trains are allocated a different service track
	Trains are offered for service at another time
Splitting and combining	Trains are combined or split at another time
Staff allocation	Staff will be allocated differently

Tabel 2 Factors that can be different and possible examples of these differences

C
Results of Experiment

The results of the experiments are summarized in a table. This table contains a lot of information, and can therefore be used as a reference based on the text provided in the main part of the report. The difference between the plan of the PAT and the human planner is stated in the column *difference*. This information might contain a real difference in the planning, but it might also contain a more general difference. The next column describes the conditions when this difference occurs. It states when these actions take place, and what specifics are included in this difference. The next column describes whether this difference is generic (G) or node specific (NS). There are a lot of differences that are generic, but the exact data is node specific. An example can be the fact that parking is forbidden on certain tracks (generic), in the case of Eindhoven on tracks 40 and 172 (node specific). The fifth column with the name *motivation* describes the motivation of the planner behind his or her action or choice. This information is important, because based on this information, the aspects can be determined. These are presented in the sixth column. Some differences can be related to multiple aspects. The last column presents the amount of times a difference is notified in the experiments. The number (#) represents the amount of experiments where the specific difference is notified. The amount of experiments executed is 6.

№	Difference	Condition	G / NS	Motivation	Aspects	#
1	PAT brings trains earlier to the station compared to planner.	If the train is completely serviced	Generic	If the train stays for a longer time at the service location, the flexibility to change the exact service allocation increases. This has an indirect result on the KPI's. Concerning the planning of train drivers, bringing trains earlier to the station relaxes the crew planning.	Increasing flexibility in the operational phase, Optimizing plan based on KPI's	3
2	Rolling stock is not placed at the correct service tracks, because these services have to be assigned one by one in the PAT. The identification what services are needed for a train unit cannot be done based on the input information	Services are assigned to trains that stay the night at a service location. For different type of rolling stock types, different services have to be assigned.	Generic	Services have to be assigned in the PAT per train unit number, in real life these services are assigned based on if these train unit stay at the service location	Technical limitation, optimizing plan based on KPI's	6
3	Planners like to park rolling stock types and forms (amount of carriages) on the same track. The PAT parks these randomly over the shunting yard	If there are multiple train units of the same rolling stock type and form at the same time at the SL	Generic	This eases the potential match problem because train numbers can be swapped as long as the swapped train units are of the same rolling stock type and form. Indirectly, this decreases the amount of movement necessary	Simplification of future planning assignment, increase of capacity	5
4	The PAT does not take into account time for passengers to get out of trains, resulting in trains that are immediately moved towards the SL after arrival at the platform.	If a train arrives, it takes time before all the passengers have left the train and the train is checked on sleeping passengers	Generic [Data dependent on train services]	These exact times are not yet fully implemented by planners, so these times can differ from planner to planner. There are however certain agreements or norms for certain train services.	Norms and agreements, Data limitations	2
5	The PAT does not take into account train movement of third parties. Human planners do have to take this into account limiting the movements over the node.	Train movement of third parties are set in real life.	Generic	These movements do limit the planner in real life, so they have to be taken into account in the PAT as well	Data limitation	6

№	Difference	Condition	G / NS	Motivation	Aspects	#
6	After arrival, trains are moved later to the SL in human plans compared to the plan of the PAT	If a train arrives, the passengers need time to leave the train	Generic [Data dependent on train services]	Exact time are not set yet, but at least some time is needed	Norms and agreements, Data limitations	1
7	Human planners prefer certain tracks for certain rolling stock combinations because of optimal usage of the tracks	When deciding parking spots for train units	Generic [Data dependent on layout]	By parking certain combinations of rolling stock types at certain tracks, the occupancy of the tracks can be optimized, resulting in an optimal usage of the complete SL. By using these rules from the begin on the chance on replanning decreases	Increase of capacity, simplification of the future planning assignment	6
8	Human planners take into account the fact that cleaning teams and mechanics are present at the SL during certain time periods. In the PAT, the trains can be serviced 24/7	When service assignments are planned	Generic [Data is node specific]	Service cannot be executed if crew is not present at the SL. The exact times are based on an agreement with the dispatchers of the SL. Indirectly, this information increases the KPI's	Norms and agreements,Optimization of KPI's	2
9	The PAT contains a lot of difficult matches between train units, while in real life these matches are set.	If a sprinter service consist of two different train numbers	Generic [Depending on train services]	The matching of an ongoing sprinter service is set, meaning that node planners do not need to look at this.	Technical limitation	4
10	The PAT does not assign services to trains that only need to be serviced before moving on to another node or does assign services to trains that have already had services	Some trains are only serviced at the SL and then sent to another node.	Generic	For human planners, these trains need to identified manually. These trains get priority when assigning services, since they often have to be moved to another node within limited time	Data limitation, Optimization based on KPI's	3
11	The PAT parks trains at parking tracks that are not servicing track, while service tracks are still available. Human planners will always first fill the service tracks, before using only parking tracks	When deciding parking spots for train units	Generic	In order to increase flexibility during the operational phase, it is preferred to park as many trains at a service track. Indirectly this will aim at a higher KPI and increases capacity because the amount of movement (in order to service trains) are limited.	Increasing flexibility in the operational phase, Optimizing plan based on KPI's, increase of capacity	4

№	Difference	Condition	G / NS	Motivation	Aspects	#
12	The PAT plans to change directions (performs saw movements) at tracks where this is not allowed because of (traffic) safety.	If a train needs to change direction in order to move from A to B	Generic [Data dependent on safety rules]	It is unsafe for train drivers to exit the train at these points. Even if train drivers are able to walk from one cabin to another cabin (singe train unit), traffic risks are to high to change directions.	Safety	2
13	It is possible to change driving direction at certain tracks, only if this is according to the traffic safety rules	If a train exist out of only one single train unit and the train driver is able to change cabins without leaving the train	Generic	This option is only used by human planners if other tracks where saw movements are allowed are occupied or cannot be reached.	Safety	2
14	The PAT brings multiple trains to the platform in the early morning, exceeding the platform length. In human plans, this is never identified	If the train is completely serviced	Generic	In theory, it is possible to already place multiple arriving and departing trains at the platform as long as traffic safety conditions are fine. However, for passengers this might result in longer walking times to trains (because they are on the edge of the platform), which is unwanted. Therefor, this option is highly unfavorable, since problems will arise during the execution of this plan	Feasibility of the plan, Optimizing plan based on KPI's	2
15	For certain rolling stock types or train series, planners prefer certain tracks to place these train units	If certain rolling stock types or certain train services need to be pared and serviced at a SL	Generic [Data dependent on layout and rolling stock]	Empty trains can be parked directly to the shunting yard, some combination of different rolling stock increase the usage of certain tracks	Simplification of future planning assignment, optimization plan based on KPI's	5
16	Some planners park train units of a different rolling stock type at a track that is not yet occupied	During daytime and if there are tracks that are not yet occupied by train units	Generic [Data dependent on train services]	By doing this, it is possible to call back a train unit of every type immediately, since these are not blocked by train units of other rolling stock types. Especially during disruptions, this can be useful	Increasing flexibility in the operational phase	2

№	Difference	Condition	G / NS	Motivation	Aspects	#
17	For certain rolling stock types or train series, planners prefer certain tracks to place these train units	If certain rolling stock types or certain train services need to be parked and serviced at a SL	Generic [Data dependent on layout and rolling stock]	Empty trains can be parked directly to the shunting yard, some combination of different rolling stock increase the usage of certain tracks	Simplification of future planning assignment, optimization plan based on KPI's	2
18	The time to change directions for certain times are different between the PAT and the human planner	If a train needs to change direction this takes up time according to rolling stock type and train combination	Generic [Data dependent on rolling stock]	This time is needed in order to create a feasible plan. Exceptions (using train drivers in both cabins) are ways to decrease the amount of time.	Data limitation	1
19	Some human planners park this train at a track that is already occupied	During daytime a train unit needs to be parked at the SL	Generic	By doing this, there is freedom in the operation to easily park more train because of a disruption and still operate the original plan, since these new trains do not block the planned trains. Indirectly, parking trains like this optimizes the utilization of parking tracks.	Increasing flexibility in the operational phase, increase of capacity	3
20	Human planners try to optimize the utilization rate of parking tracks	If train units need to be parked	Generic	This results in an optimal usage of the complete shunting location.	Increase of capacity	4
21	The human plan has (a lot) less (saw)movements compared to the PAT plan	When a train needs to be routed on a node	Generic	Less movements are wanted, because the are less complex to execute, they keep the planning simple (less conflicts because of movements) and they make it easy to add an extra movement, since there is more solution space	Simplification of the future plan assignment, Increasing flexibility in the operational phase, Increase of capacity	4
22	Planners combine these trains, then move them and then split them, whereas the PAT moves these trains separately	If two trains need to be moved to (relatively) the same destination.	Generic	This reduces the amount of train drivers that are needed, which increases the availability of the train drivers	Increase of capacity	1

№	Difference	Condition	G / NS	Motivation	Aspects	#
23	Human planners parks train as long as possible at the SL, so that services can be executed to the train as long as possible	During weekdays, when the amount of services that need to be executed is higher compared to the weekends	Generic	If the train stays for a longer time at the service location, the flexibility to change the exact service allocation increases. This has an indirect result on the KPI's. Concerning the planning of train drivers, bringing trains earlier to the station relaxes the crew planning.	Increasing flexibility in the operational phase, Optimizing plan based on KPI's	2
24	PAT brings trains earlier to the station compared to human planners, which is not preferable.	During weekdays, when the amount of services that need to be executed is higher compared to the weekends	Generic	If the train stays for a longer time at the service location, the flexibility to change the exact service allocation increases. This has an indirect result on the KPI's. Concerning the planning of train drivers, bringing trains earlier to the station relaxes the crew planning.	Increasing flexibility in the operational phase, Optimizing plan based on KPI's	2
25	The PAT plans empty material first to a platform track, after which it is moved the the SL. A human planners uses these parking tracks or plans the train immediately to the SL	Empty trains (needed for the rolling stock balance in the country) are not planned to or from the platform track to or from parking tracks	Generic	All train drivers can drive a train on CBG area, meaning that these trains can be parked immediately on CBG tracks at the SL. This increases the capacity of the train drivers (since the arriving train driver takes over a driving task) and less movements are necessary. Also, the platform tracks are not used, which increases the flexibility during operation. This data is however not available.	Data limitation, technical limitation, increase of capacity, increasing flexibility in the operational phase	3
26	Some tracks are highly used by the PAT and almost never by the human planner.	Looking at the usage of the tracks at the SL	Node specific	By removing this track, the "garden" area changes from a carousel type to a shuffleboard type. By doing this, the FILO strategy is necessary which decreases the amount of movements (no driving direction movements necessary). This results in a simplification of the planning assignment. Another reason is the fact that train drivers get the assignment to drive the train to the end of the track. If all the trains are parked at the end of the track, there is not room to carousel (only from and to the same track).	Simplification of the planning assignment, Feasibility	4
27	Human planners park the train at a track where it does not hinder other trains and therefore also does not have to move	If a train needs to be parked for a long duration at the node.	Generic [Data dependent on layout]	This reduces the amount of movements and therefore simplifies the future planning assignment	Simplification of future planning assignment	1

№	Difference	Condition	G / NS	Motivation	Aspects	#
28	The PAT used tracks that are not assigned to the NS or that cannot be used for planners	When parking trains and assigning services	Generic	There are strict rules with other parties which parts can be used by NS. Using other parts is not preferred and can only be used when agreed upon with the third parties.	Norms and agreements	1
29	The PAT plans a long parking break between certain saw movements. Human planners always try to perform a saw movement with an as small break as possible.	Within a saw movement, the break between the first movement and the second movement must be	Generic	First of all, these breaks take place at important tracks that for the accessibility of the SL. Secondly, these these breaks take up a lot of time of the train driver in which no additional value is added.	Increasing flexibility in operational phase, increase of capacity	1
30	This platform change cannot be done by the PAT, while the planner is able to change arrival (and departure) platforms	If another platform track reduces the complexity	Generic	To reduce complexity, changing an arrival (or departure) platform can be execute, as long as passengers are not negatively influenced by the change.	Technical limitation, simplification of future planning assignment	1
31	Within the PAT, trains are allowed to have a delayed departure	When looking to the differences between hard and soft constraints	Generic	This is a hard constraint in real life	Technical limitation	2
32	In real life, trains are allowed to enter service dirty, whereas in the PAT, this is not allowed	When looking to the differences between hard and soft constraints	Generic	This is a soft constraint for planners (it negatively affects the KPI)	Technical limitation	2
33	The PAT used tracks that are not assigned to the NS or that cannot be used for planners	When parking trains and assigning services	Generic	There are strict rules with other parties which parts can be used by NS. Using other parts is not preferred and can only be used when agreed upon with the third parties.	Norms and agreements	1
34	The PAT plans empty material first to a platform track, after which it is moved the the SL. A human planners uses these parking tracks or plans the train immediately to the SL	Empty trains (needed for the rolling stock balance in the country) are not planned to or from the platform track to or from parking tracks	Generic	All train drivers can drive a train on CBG area, meaning that these trains can be parked immediately on CBG tracks at the SL. This increases the capacity of the train drivers (since the arriving train driver takes over a driving task) and less movements are necessary. Also, the platform tracks are not used, which increases the flexibility during operation. This data is however not available.	Data limitation, technical limitation, increase of capacity, increasing flexibility in the operational phase	3

№	Difference	Condition	G / NS	Motivation	Aspects	#
35	Not all trains are serviced over night in the PAT, while the human planner offers a service for every train unit that stays at the SL overnight.	Services are assigned to trains that stay the night at a service location. For different type of rolling stock types, different services have to be assigned.	Generic	Services have to be assigned in the PAT per train unit number, in real life these services are assigned based on if these train unit stay at the service location	Technical limitation, optimizing plan based on KPI's	2
36	The PAT plans to change directions (performs saw movements) at tracks where this is not allowed because of (traffic) safety.	If a train needs to change direction in order to move from A to B	Generic [Data dependent on safety rules]	It is unsafe for train drivers to exit the train at these points. Even if train drivers are able to walk from one cabin to another cabin (singe train unit), traffic risks are to high to change directions.	Safety	1
37	The human plan has (a lot) less movements compared to the PAT plan	When a train needs to be routed on a node	Generic	Less movements are wanted, because the are less complex to execute, they keep the planning simple (less conflicts because of movements) and they make it easy to add an extra movement, since there is more solution space	Simplification of the future plan assignment, Increasing flexibility in the operational phase, Increase of capacity	3
38	Routes in the human plan are different compared to the plan of the PAT	For ongoing trains	Node specific	Routes of ongoing trains are given for the human planner. These routes are precisely described, but it is not possible to extract this exact data from the database. Therefor, the PAT creates its own routes based on the starting and end point of these routes. Some of these routes cross important shunting tracks (like saw tracks) minimizing the possibility of using these tracks. In real life, these tracks will never be used for ongoing trains.	Data limitation	3
39	Human planners park the train at a track where it does not hinder other trains and therefore also does not have to move	If a train needs to be parked for a long duration at the node.	Generic [Data dependent on layout]	This reduces the amount of movements and therefore simplifies the future planning assignment	Simplification of future planning assignment	1

№	Difference	Condition	G / NS	Motivation	Aspects	#
40	In order to optimise the use of the tracks, planners prefer to place certain rolling stock types at certain tracks	When deciding parking spots for train units that are used for special train services	Generic [Data dependent on train services]	By parking certain combinations of rolling stock types at certain tracks, the occupancy of the tracks can be optimized, resulting in an optimal usage of the complete SL. By using these rules from the begin on the chance on replanning decreases	Increase of capacity, simplification of the future planning assignment	3
41	In order to optimise the use of the tracks, planners prefer to place certain rolling stock types at certain tracks	When deciding parking spots for long train units	Generic [Data dependent on rolling stock]	By parking certain combinations of rolling stock types at certain tracks, the occupancy of the tracks can be optimized, resulting in an optimal usage of the complete SL. By using these rules from the begin on the chance on replanning decreases	Increase of capacity, simplification of the future planning assignment	1
42	Planners prefer to fill first all the service tracks, before the parking tracks without service resources are used	If parking spots have to be determined for trains	Generic [Data dependent on layout]	This reduces the complexity of the planning, since no movements are necessary in order to service trains	Simplification of future planning assignment, Optimization based on KPI’s	3
43	The PAT reserves complete routes for ongoing trains, meaning that these tracks are all not possible to be used, while the human planner has more freedom because the exact location of a train is used and not the complete route.	Looking at the reservation of track for movements of trains	Generic	These high reservation norms result in a smaller solution space, since less movements are possible.	Data limitation, technical limitation, Safety	3
44	The PAT only moves split-off train units of ongoing trains after the ongoing train has left the station. Human planners plan the movement of the split of train unit sometimes earlier, by moving it in the opposite direction of the ongoing train to the SL.	When train units are split off of ongoing trains	Generic	Increase the time that train units can be moved to a SL increases the possibilities to do so.	Capacity optimization	2

№	Difference	Condition	G / NS	Motivation	Aspects	#
45	Rolling stock is not placed at the correct service tracks, because these services have to be noticed by a planner and have to be assigned by the planner	Services are assigned to trains that stay the night at a service location. For different type of rolling stock types, different services have to be assigned.	Generic	Services have to be assigned in the PAT per train unit number, in real life these services are assigned based on if these train unit stay at the service location	Technical limitation, optimizing plan based on KPI's	1
46	The PAT reserves complete routes for ongoing trains, meaning that these tracks are all not possible to be used, while the human planner has more freedom because the exact location of a train is used and not the complete route.	Looking at the reservation of platforms of stations	Generic	These high reservation norms result in a smaller solution space, since less movements are possible.	Data limitation, technical limitation, Safety	3
47	The PAT identifies a conflict if a train is parked at a track where this is not possible, even if this duration is short. For human planners, this is not a conflict, which means that no further actions are necessary to make the plan feasible.	In between two movement (saw movements) it might be necessary for a train to wait a little bit before the next movement can start	Generic	In real life, this is not a conflict, but for the PAT, it is.	Technical limitation	3
48	Certain tracks can be used for short parking in order to combine and split trains or to wait before the next movement is possible.	In between two movement (saw movements) it mighty be necessary for a train to wait a little bit before the next movement can start	Generic [Data dependent on layout]	In real life, this is not a conflict, but for the PAT, it is.	Technical limitation	3
49	These train units can be moved in opposite direction to empty the platform track in real life. The PAT did not use this opportunity, causing conflicts that were not needed	When train units are split-off of ongoing trains	Generic	Increase the time that train units can be moved to a SL increases the possibilities to do so.	Capacity optimization	2

№	Difference	Condition	G / NS	Motivation	Aspects	#
50	In the garden area of the shunting yard, the PAT uses track 47 a lot in order to bring trains from the right side of the yard to the station. These movement include a saw movement. Human planners never apply these kinds of movements, because their set-up is based on the FILO (First In Last Out) system.	Looking at the usage of the tracks at the SL	Node specific	By removing this track, the "garden" area changes from a carousel type to a shuffleboard type. By doing this, the FILO strategy is necessary which decreases the amount of movements (no driving direction movements necessary). This results in a simplification of the planning assignment. Another reason is the fact that train drivers get the assignment to drive the train to the end of the track. If all the trains are parked at the end of the track, there is not room to carousel (only from and to the same track).	Simplification of the planning assignment, Feasibility	2
51	Because of the high utilization rate of track 47, simple rematching of rolling stock is not executed. These simple rematches reduces the amount movements necessary.	Looking at the usage of the tracks at the SL	Node specific	By removing this track, the "garden" area changes from a carousel type to a shuffleboard type. By doing this, the FILO strategy is necessary which decreases the amount of movements (no driving direction movements necessary). This results in a simplification of the planning assignment. Another reason is the fact that train drivers get the assignment to drive the train to the end of the track. If all the trains are parked at the end of the track, there is not room to carousel (only from and to the same track). The less movements necessary increase the availability of the train drivers	Simplification of the planning assignment, Feasibility, Increase of capacity	2
52	Empty rolling stock is planned by the PAT over platform tracks, which is not favored by human planners	If empty rolling stock needs to be serviced / parked at the shunting yard	Generic	Platform tracks are normally highly utilized and therefor, extra movements are unwanted on these track. Because of the high utilization, the flexibility decreases in the operational phase, which is also unwanted	Simplification of the future plan assignment, increasing flexibility in the operational phase	1
53	The human planner uses the parking tracks that can be entered from the main tracks immediately (no saw movement) and from both directions.	Sometimes empty rolling stock passes through the node that needs to split-of or combine with a train unit. For these cases	Generic [Data dependent on layout]	It is unwanted to use platform tracks for these kind of trains / actions, and a lot of movements are most of the time not possible because of the timetable. Therefore, these tracks are the most easy to use considering these type of trains	Simplification of future plan assignment	1

№	Difference	Condition	G / NS	Motivation	Aspects	#
54	The PAT plans to change directions (performs saw movements) at tracks where this is not allowed because of (traffic) safety.	If a train needs to change direction in order to move from A to B	Generic [Data dependent on safety rules]	It is unsafe for train drivers to exit the train at these points. Even if train drivers are able to walk from one cabin to another cabin (singe train unit), traffic risks are to high to change directions.	Safety	1
55	The PAT parks this rolling stock in the garden, whereas the human planner prefers the south side of the shunting yard.	If empty rolling stock needs to depart from the node	Generic [Data dependent on layout]	The garden is a parking location that is easy accessible from all the platform tracks. Therefore, the garden must be highly utilized for passenger trains that need to start up. If Empty rolling stock trains start up from the garden, this increases the complexity of the planning assignment, since the platform tracks must be used, which are already highly utilized.	Simplification of future plan assignment, Increasing flexibility in operational phase.	1
56	The PAT allows services at tracks where service resources are not present	When trains need to be serviced	Generic	This is not possible	Optimizing plan based on KPI's	2
57	The PAT parks trains at tracks, where long term parking is not permitted	When parking trains	Generic [Data dependent on layout]	These track are important for the network (saw tracks) and therefor must not be occupied by a long term parked train	Increase of flexibility of operational phase, simplification of future plan assignment	2
58	It is possible to change driving direction at certain tracks, only if this is according to the traffic safety rules	If a train exist out of only one single train unit and the train driver is able to change cabins without leaving the train	Generic [Data dependent on safety rules]	This option is only used by human planners if other tracks where saw movements are allowed are occupied or cannot be reached.	Safety	1
59	Some tracks can only be used if other (comparable) tracks are out of service and specific rules are made with the original user of these tracks (third parties)	If important tracks are out of service	Generic	In this case, the third party is the maintenance department of NS, so agreements are easily made (or sometime not even necessary)	Norms and agreements	1

D

Implementing measures

In order to add aspects to the PAT, measures are added one by one. The test planning assignment was a real life planning assignment that was used to identify the aspects with a planner. The running time for the PAT was 60 seconds, and three seeds were used. This means that three different solutions were calculated with one run. The average of these three solutions was used to compare the old PAT and the updated PAT. These solutions were compared using different indicators, that are described in Section 7.1.

Within this Appendix, the results of the measures are presented in the same sequence as they were implemented in the PAT (Section D.1). The *basic* test environment refers to last used test environment. For the first test environment, this means that the original PAT is used. For the second test environment, the basic resembles the adjusted PAT of the first test environment. At the end, the total difference between the original PAT and the PAT including the measures will be presented in Section D.2.

D.1. Sequential differences of measures

The implementation of the measures is performed once or twice at a time in order to identify the affect of the measure to the algorithm. For the complete results. The most important affects will be described per measure underneath. The real quantitative differences will be presented in the Subsections of this section. The affects of the measures are tested using a real-life planning assignment with 3 seeds, meaning that one run provides three different solutions. The running time for the PAT was 1 minute (60 seconds). The following affects were identified for the measure(s) indicated in bold:

- **Manual assignment of services:** The amount of delayed departing trains decreased. Since indicators of clean trains and services performed are not yet identified by the PAT, no overview can be given of the service processes (ex. amount of trains cleaned or mechanically checked).
- **Change and addition of parking constraints and driving directions constraints:** The removal of parking possibilities and the possibility to change the driving direction increased the complexity of the plan assignment. This increase is however small. The addition of these rules clearly showed that some tracks are not used anymore for parking or changing driving directions. The parking constraints (not allowed to park at a certain track) has a direct link with the illegal parking time violations, which is likely to increase, because certain tracks will sometimes still be used to park trains for a limited time.
- **Removal of tracks in the layout of the node:** The removal of tracks that cannot be used by NS decreases the plan complexity, since less options are possible. This results in a clear increase of the amount of candidates that can be checked by the PAT (+14%).
- **Change driving direction times:** The increase of these times do not have serious outcomes, only that this train is brought later to the platform because it needs more time to change directions. No new or other problems occurred.
- **Changing the layout of the garden:** By removing track 47, the PAT automatically changes the parking method to a First In Last Out (FILO) method, since retrieving trains from the back of the garden yard is

not possible any more. The change in the layout of the garden comes with a big decrease of the penalty costs (-80%). This decrease is caused by a small decrease of movements (-3%) and a big decrease of illegal parking time (-80%). The amount of crossings increased with 47% and the value of track length violation as well with 24%.

- **Removal of mobile service resources:** By removing the mobile resources (since planners do not plan with these utilities), the conflicts regarding services being executed at tracks where this is not possible decreased to zero.

- **Addition of certain fixed matches to input:** By fixing the matching for the sprinter service, the plan assignment becomes less complex. This is good visible by the increase of number of candidates checked (5%). The irrational matches that the PAT created earlier also came with a larger amount of movements that were necessary: the number of movements decreased with -4%. The number of delayed departures and delayed arrivals decreased with -25% and -13% respectively. The amount of track length violations also decreased with -20%. All these decreases result in lower conflict costs (-14%) for the planning assignment with fixed matching. However, the penalty cost increases with 52%, which is caused by a higher illegal parking time (53%). The combinal departure conflicts (departure of trains in the wrong sequence) increase also with 50%.

- **Add opening times for Service Location:** By adding opening times for service locations, meaning the times that employees are present at the SL to perform services, the plan assignment should be a little bit more complex. However, this increase in complexity falls into thin air when comparing cases with and without the opening times. The number of candidates checked even increased by a small percentage (2%). The only interesting part is the fact that the penalty costs decreased with -13%. An explanation cannot be found for this change.

- **Removal of switch in the layout:** By removing the switch, track 35A would be inaccessible for ongoing trains, but can still be used for saw movements in order to shunt trains from the station to one of the shunting yards. This adjustment results in smaller conflict costs (-5%) and a lower penalty costs (-12%). A decrease of track length violations (-33%) and illegal parking time (-18%) also have to be noticed. Although the change in the location file is small, this causes for a more complex problem, because the number of candidates being checked decreased with -12%.

- **Addition of parking weights:** Both the conflict costs as the penalty costs decreases with -13% and -54% respectively. These decreases are based on the decrease of the amount of crossing conflicts, the decrease of track length violation conflicts and the decrease of illegal parking time violation. The only large increase is the amount of delayed fixed trains with 33%.

D.1.1. Removal of track 47

Table D.1: The difference of the outcomes of the PAT after removing track 47.

	Basic			**Track 47**			**Averages**		**Difference**
Measurements	**0**	**1**	**2**	**0**	**1**	**2**	**Basic**	**Track 47**	
Conflict costs	2004	2011	1761	1945	1796	1937	1925,3	1892,7	-2%
Penalty costs	560	623	703	236	88	78	628,7	134,0	-79%
Crossings	16	14	15	27	17	23	15,0	22,3	49%
Delayed departures	26	25	21	24	22	26	24,0	24,0	0%
Delayed arrivals	20	19	14	15	17	12	17,7	14,7	-17%
Track length violations	7	9	13	14	7	15	9,7	12,0	24%
Incorrect combination	1	2	4	1	2	1	2,3	1,3	-43%
Delayed fixed train	78	81	64	65	69	66	74,3	66,7	-10%
Number of movements	251	245	274	254	239	254	256,7	249,0	-3%
Illegal parking time violation	55205	61640	69474	22839	8039	6999	62106,3	12625,7	-80%
Candidates checked	2657	2771	2819	2863	2957	3045	2749,0	2955,0	7%

D.1.2. Removal of tracks that NS cannot use

Table D.2: The difference of the outcomes of the PAT after the removal of tracks that are not predetermined to use for NS.

	Basic			Removal of tracks			Averages		Difference
Measurements	**0**	**1**	**2**	**0**	**1**	**2**	**Basic**	**Rem. tracks**	
Conflict costs	2140	2200	1887	1936	1788	2307	2075,7	2010,3	-3%
Penalty costs	545	120	246	737	880	258	303,7	625,0	106%
Crossings	30	25	28	26	28	51	27,7	35,0	27%
Delayed departures	23	23	20	26	24	25	22,0	25,0	14%
Delayed arrivals	26	24	20	14	12	16	23,3	14,0	-40%
Track length violations	2	14	4	6	3	11	6,7	6,7	0%
Incorrect combination	1	1	2	1	1	3	1,3	1,7	25%
Delayed fixed train	83	82	72	72	67	74	79,0	71,0	-10%
Number of movements	250	258	246	245	248	256	251,3	249,7	-1%
Illegal parking time violation	44621	11234	23764	72885	87230	25038	26539,7	61717,7	133%
Candidates checked	2443	2469	2427	2892	2820	2674	2446,3	2795,3	14%

D.1.3. Fixed matching

Table D.3: The difference of the outcomes of the PAT when the matching of the sprinter service is fixed.

	Basic			Fixed matching			Averages		Difference
Measurements	**0**	**1**	**2**	**0**	**1**	**2**	**Basic**	**Fixed matching**	
Conflict costs	2004	2181	1828	1481	1555	1520	2004,3	1518,7	-24%
Penalty costs	305	128	83	374	742	499	172,0	538,3	213%
Crossings	21	34	40	19	17	21	31,7	19,0	-40%
Delayed departures	26	26	20	18	20	18	24,0	18,7	-22%
Delayed arrivals	20	19	12	11	12	12	17,0	11,7	-31%
Track length violations	4	8	6	2	5	5	6,0	4,0	-33%
Incorrect combination	0	1	1	3	1	2	0,7	2,0	200%
Delayed fixed train	79	78	65	60	62	58	74,0	60,0	-19%
Number of movements	252	238	246	220	231	225	245,3	225,3	-8%
Illegal parking time violation	29722	11993	7538	36625	73440	49077	16417,7	53047,3	223%
Candidates checked	2950	2885	3100	3284	3197	3371	2978,3	3284,0	10%

D.1.4. Opening time of SL

Table D.4: The difference of the outcomes of the PAT when the opening times of the Service Location are added to the algorithm

	Basic			Opening time of SL			Averages		Difference
Measurements	**0**	**1**	**2**	**0**	**1**	**2**	**Basic**	**Opening time**	
Conflict costs	1481	1555	1520	1298	1598	1566	1518,7	1487,3	-2%
Penalty costs	374	742	499	419	391	443	538,3	417,7	-22%
Crossings	19	17	21	12	17	19	19,0	16,0	-16%
Delayed departures	18	20	18	14	22	17	18,7	17,7	-5%
Delayed arrivals	11	12	12	12	12	13	11,7	12,3	6%
Track length violations	2	5	5	6	2	4	4,0	4,0	0%
Incorrect combination	3	1	2	4	2	3	2,0	3,0	50%
Delayed fixed train	60	62	58	50	64	65	60,0	59,7	-1%
Number of movements	220	231	225	224	221	213	225,3	219,3	-3%
Illegal parking time violation	36625	73440	49077	41140	38276	43563	53047,3	40993,0	-23%
Candidates checked	3284	3197	3371	3188	3433	3551	3284,0	3390,7	3%

D.1.5. Routing

Table D.5: The difference of the outcomes of the PAT when routes of ongoing trains are changed.

	Basic			Routing			Averages		Difference
Measurements	**0**	**1**	**2**	**0**	**1**	**2**	**Basic**	**Routing**	
Conflict costs	1298	1598	1566	1240	1419	1395	1487,3	1351,3	-9%
Penalty costs	419	391	443	193	527	264	417,7	328,0	-21%
Crossings	12	17	19	15	15	20	16,0	16,7	4%
Delayed departures	14	22	17	16	19	20	17,7	18,3	4%
Delayed arrivals	12	12	13	10	9	10	12,3	9,7	-22%
Track length violations	6	2	4	3	2	1	4,0	2,0	-50%
Incorrect combination	4	2	3	2	3	1	3,0	2,0	-33%
Delayed fixed train	50	64	65	47	58	51	59,7	52,0	-13%
Number of movements	224	221	213	228	222	233	219,3	227,7	4%
Illegal parking time violation	41140	38276	43563	18562	51890	15571	40993,0	28674,3	-30%
Candidates checked	3188	3433	3551	2639	2755	2550	3390,7	2648,0	-22%

D.1.6. Track preference

The addition of the track preference function in the PAT required a small change to the objective function, since this one had to include the optimization of the track preference score. Because of this, the PT (parking track) measure is added. This parking track measure is multiplied by 0.0002 and subtracted from the penalty costs. This penalty score will be presented in the row *Penalties inc. pref. tracks*.

Table D.6: The difference of the outcomes of the PAT when track preferences are added to the algorithm

	Basic			Track preference			Averages		Difference
Measurements	**0**	**1**	**2**	**0**	**1**	**2**	**Basic**	**Track preference**	
Conflict costs	1240	1419	1395	1259	1268	997	1351,3	1174,7	-13%
Penalty costs	193	527	264	109	101	241	328,0	150,6	-54%
Penalties inc. pref. tracks	-	-	-	-52,42	-40,73	110	-	5,6	-
Crossings	15	15	20	9	9	6	16,7	8,0	-52%
Delayed departures	16	19	20	20	18	16	18,3	18,0	-2%
Delayed arrivals	10	9	10	9	11	10	9,7	10,0	3%
Track length violations	3	2	1	2	2	0	2,0	1,3	-33%
Incorrect combination	2	3	1	2	3	3	2,0	2,7	33%
Delayed fixed train	47	58	51	47	49	34	52,0	43,3	-17%
Number of movements	228	222	233	242	235	246	227,7	241,0	6%
Illegal parking time violation	18562	51890	15571	10034	9156	23288	28674,3	14159,3	-51%
Preference track points	-	-	-	806875	704466	657210	-	722850,3	-
Candidates checked	2639	2755	2550	2896	2879	2726	2648,0	2833,7	7%

D.2. Total difference of measures

This Appendix concludes with the total affects of the measures to the PAT. Table D.7 presents the quantitative differences between the original PAT and the adjusted PAT. The affects of the measures are tested using a real-life planning assignment with 3 seeds, meaning that one run provides three different solutions. The running time for the PAT was 1 minute (60 seconds). The table shows that both the conflict costs as the penalty costs decreased for the adjusted PAT. Especially the penalty costs decreased with a large number. These decreases are based on a decrease on almost all the different conflicts and violations. The only conflict that is negatively affected by the measures is the combinal departure conflict, meaning that the combination of train unit within a train is not as planned. The increase in the amount of candidates indicates that the adjusted PAT is able to perform more iterations for the local search optimization within the same time window, meaning that the planning assignment is less complex. However, it must be stated that these results are only based on one real-life planning assignment and that only three seeds were used for these comparisons..

Table D.7: The difference of indicators between the original PAT and the new PAT

	Orignal			New			Averages		Difference
Measurements	**0**	**1**	**2**	**0**	**1**	**2**	**Original**	**New**	
Conflict costs	2004	2011	1761	1259	1268	997	1925,3	1174,7	-39%
Penalty costs	560	623	703	109	101	241	628,7	150,6	-76%
Penalties inc. Pref. Tracks	-	-	-	-52	-40	110	-	6,0	-
Crossings	16	14	15	9	9	6	15,0	8,0	-47%
Delayed departures	26	25	21	20	18	16	24,0	18,0	-25%
Delayed arrivals	20	19	14	9	11	10	17,7	10,0	-43%
Track length violations	7	9	13	2	2	0	9,7	1,3	-86%
Incorrect combination	1	2	4	2	3	3	2,3	2,7	14%
Delayed fixed train	78	81	64	47	49	34	74,3	43,3	-42%
Number of movements	251	245	274	242	235	246	256,7	241,0	-6%
Illegal parking time violation	55205	61640	69474	10034	9156	23288	62106,3	14159,3	-77%
Preference Tracks points	-	-	-	806875	704466	657210	-	722850,3	-
Candidates checked	2657	2771	2819	2896	2879	2726	2749,0	2833,7	3%

E

Quantitative comparison real-life

In this appendix, the quantitative results are presented concerning the original PAT and the adjusted PAT for the real-life planning assignments. The results are presented per planning assignment. The running time for the PAT was 10 minutes (600 seconds). The amount of seeds differs between 2 for A1 and B, and 3 for the rest of the planning assignments. This differences was based on the amount of seeds that were used for the experiments with planners to identify aspects.

The table all contain the same information. In the first column, the indicators (described in Section 7.1) are presented. The following columns contain the results of either the two or three seeds of the original PAT. The following columns contain the information of the two or three seed of the adjusted PAT. The average section presents the average results of the two or three seeds of either the original PAT or the adjusted PAT. The last column presents the difference between the average of the original PAT and the average of the adjusted PAT in a percentage. The indicator parking weights points shows the points that the PAT has received for parking weights. Since these points were not present in the original PAT, no values were provided in the original columns and the average original column.

E.1. Results A1

Table E.1: Results of planning assignment A1

Indicators	**Original**		**Adjusted**		**Averages**		**Difference**
	0	**1**	**0**	**1**	**Original**	**Adjusted**	
Conflict Costs	466	426	400	351	446,0	375,5	**-16%**
Penalty costs	84	22	187	196	53,0	191,9	**262%**
Crossings	31	5	4	3	18,0	3,5	**-81%**
Delayed departures	3	4	3	2	3,5	2,5	**-29%**
Sum of departure delay	717	1076	516	392	896,5	454,0	**-49%**
Delayed arrivals	2	1	0	0	1,5	0,0	**-100%**
Sum of arrival delay	122	30	0	0	76,0	0,0	**-100%**
Track length violations (TLV)	0	0	0	0	0,0	0,0	**0%**
Duration of TLV	0	0	0	0	0,0	0,0	**0%**
Maximum overshoot of TLV	0	0	0	0	0,0	0,0	**0%**
Combinal departure problems	0	0	0	0	0,0	0,0	**0%**
Delayed fixed train	5	26	27	27	15,5	27,0	**74%**
Sum of delayed fixed train	520	3718	4068	3788	2119,0	3928,0	**85%**
Number of movements	272	337	282	276	304,5	279,0	**-8%**
Illegal parking time violation	7676	1293	1498	2584	4484,5	2041,0	**-54%**
Parking weight points			821576	811939		816757,5	
Candidates checked	54163	48970	38340	37518	51566,5	37929,0	**-26%**

E.2. Results B

Table E.2: Results of planning assignment B

Indicators	Original		Adjusted		Averages		Difference
	0	1	0	1	Original	Adjusted	
Conflict Costs	2861	3054	312	295	2957,5	303,5	**-90%**
Penalty costs	276	504	305	294	390,0	299,2	**-23%**
Crossings	11	18	1	3	14,5	2,0	**-86%**
Delayed departures	40	39	5	5	39,5	5,0	**-87%**
Sum of departure delay	78920	83148	946	978	81034,0	962,0	**-99%**
Delayed arrivals	44	41	1	1	42,5	1,0	**-98%**
Sum of arrival delay	30888	26035	60	60	28461,5	60,0	**-100%**
Track length violations (TLV)	1	4	4	0	2,5	2,0	**-20%**
Duration of TLV	2940	6323	27024	0	4631,5	13512,0	**192%**
Maximum overshoot of TLV	79	226	384	0	152,5	192,0	**26%**
Combinal departure problems	1	0	2	2	0,5	2,0	**300%**
Delayed fixed train	105	122	7	9	113,5	8,0	**-93%**
Sum of delayed fixed train	70455	78690	524	718	74572,5	621,0	**-99%**
Number of movements	321	328	319	336	324,5	327,5	**1%**
Illegal parking time violation	26734	49522	9436	7619	38128,0	8527,5	**-78%**
Parking weight points			1008598	1043847		1026222,5	
Candidates checked	30514	29461	33725	33521	29987,5	33623,0	**12%**

E.3. Results A2

Table E.3: Results of planning assignment A2

Original	Original			Adjusted			Averages		Difference
	0	1	2	0	1	2	Original	Adjusted	
Conflict Costs	738	789	810	613	697	716	779,0	675,3	**-13%**
Penalty costs	79	78	387	247	237	208	181,3	230,5	**27%**
Crossings	9	9	17	5	9	8	11,7	7,3	**-37%**
Delayed departures	8	9	9	9	11	11	8,7	10,3	**19%**
Sum of departure delay	2232	2052	2034	2277	1878	2933	2106,0	2362,7	**12%**
Delayed arrivals	4	9	5	6	5	6	6,0	5,7	**-6%**
Sum of arrival delay	784	1125	620	1156	1304	705	843,0	1055,0	**25%**
Track length violations (TLV)	6	5	6	0	0	0	5,7	0,0	**-100%**
Duration of TLV	5211	2527	23524	0	0	0	10420,7	0,0	**-100%**
Maximum overshoot of TLV	277	206	215	0	0	0	232,7	0,0	**-100%**
Combinal departure problems	2	1	0	1	1	1	1,0	1,0	**0%**
Delayed fixed train	24	26	26	22	23	24	25,3	23,0	**-9%**
Sum of delayed fixed train	5376	4862	5330	5683	4876	4642	5189,3	5067,0	**-2%**
Number of movements	255	258	254	275	250	277	255,7	267,3	**5%**
Illegal parking time violation	7040	6997	37863	6504	6554	3613	17300,0	5557,0	**-68%**
Parking weight points				868411	817593	821074		835692,7	
Candidates checked	33101	34394	34441	33403	34224	33378	33978,7	33668,3	**-1%**

E.4. Results C

Table E.4: Results of planning assignment C

Original	Original			Adjusted			Averages		Difference
	0	1	2	0	1	2	Original	Adjusted	
Conflict Costs	1398	1231	1621	1888	1845	2021	1416,7	1918,0	**35%**
Penalty costs	1621	631	365	273	770	566	872,3	536,3	**-39%**
Crossings	22	19	16	15	24	28	19,0	22,3	**18%**
Delayed departures	14	12	20	25	22	20	15,3	22,3	**46%**
Sum of departure delay	6048	4884	9900	10412	10626	10068	6944,0	10368,7	**49%**
Delayed arrivals	15	13	11	14	16	18	13,0	16,0	**23%**
Sum of arrival delay	2730	1690	2398	4497	4702	5562	2272,7	4920,3	**117%**
Track length violations (TLV)	8	7	10	19	14	16	8,3	16,3	**96%**
Duration of TLV	39520	28632	66524	48389	47836	29227	44892,0	41817,3	**-7%**
Maximum overshoot of TLV	628	422	502	970	548	762	517,3	760,0	**47%**
Combinal departure problems	2	2	1	1	2	3	1,7	2,0	**20%**
Delayed fixed train	44	42	64	58	54	66	50,0	59,3	**19%**
Sum of delayed fixed train	8360	9114	17152	25350	19969	25176	11542,0	23498,3	**104%**
Number of movements	291	299	294	331	314	298	294,7	314,3	**7%**
Illegal parking time violation	161228	62216	35636	9453	63238	42276	86360,0	38322,3	**-56%**
Parking weight points				848655	643554	676562		722923,7	
Candidates checked	27778	26813	27288	30600	29557	31162	27293,0	30439,7	**12%**

E.5. Results D

Table E.5: Results of planning assignment D

Original	Original			Adjusted			Averages		Difference
	0	1	2	0	1	2	Original	Adjusted	
Conflict Costs	1017	937	892	1054	848	998	957,7	966,7	**1%**
Penalty costs	424	569	272	333	290	313	421,7	312,0	**-26%**
Crossings	15	15	13	13	8	17	13,5	12,7	**-6%**
Delayed departures	13	11	13	13	9	11	11,7	11,0	**-6%**
Sum of departure delay	2093	2662	2301	3052	2884	3168	2693,3	3034,7	**13%**
Delayed arrivals	8	13	7	11	9	10	9,7	10,0	**3%**
Sum of arrival delay	1120	2340	791	1764	1268	1760	1507,2	1597,3	**6%**
Track length violations (TLV)	8	3	5	11	6	7	6,7	8,0	**20%**
Duration of TLV	49114	7372	9526	23256	24776	5128	19862,0	17720,0	**-11%**
Maximum overshoot of TLV	488	310	441	887	499	448	512,2	611,3	**19%**
Combinal departure problems	0	1	1	0	2	1	0,8	1,0	**20%**
Delayed fixed train	31	24	25	27	28	28	27,2	27,7	**2%**
Sum of delayed fixed train	6231	3888	3575	4354	7884	4824	5126,0	5687,3	**11%**
Number of movements	280	272	278	253	268	272	270,5	264,3	**-2%**
Illegal parking time violation	41564	56069	26338	16807	11971	15368	28019,5	14715,3	**-47%**
Parking weight points				783120	810872	756483		783491,7	
Candidates checked	29760	30768	29389	28040	27519	27763	28873,2	27774,0	**-4%**

E.6. Results E

Table E.6: Results of planning assignment E

Original	**Original**			**Adjusted**			**Averages**		**Difference**
	0	**1**	**2**	**0**	**1**	**2**	**Original**	**Adjusted**	
Conflict Costs	1250	1362	1375	658	701	762	1018,0	707,0	**-31%**
Penalty costs	336	249	340	211	266	181	308,3	219,4	**-29%**
Crossings	23	17	16	5	3	4	11,3	4,0	**-65%**
Delayed departures	15	17	20	11	12	13	14,7	12,0	**-18%**
Sum of departure delay	5250	7310	8660	5643	4155	5308	6054,3	5035,3	**-17%**
Delayed arrivals	10	14	10	3	3	4	7,3	3,3	**-55%**
Average of arrival delay	1270	3206	1310	347	300	582	1169,2	409,7	**-65%**
Track length violations (TLV)	1	1	2	0	1	0	0,8	0,3	**-60%**
Duration of TLV	2100	10793	16630	0	10428	0	6658,5	3476,0	**-48%**
Maximum overshoot of TLV	107	55	35	0	53	0	41,7	17,7	**-58%**
Combinal departure problems	0	0	1	0	0	0	0,2	0,0	**-100%**
Delayed fixed train	46	54	52	26	29	31	39,7	28,7	**-28%**
Average of delayed fixed train	9246	10908	10816	10200	9659	6451	9546,7	8770,0	**-8%**
Number of movements	282	283	286	278	311	256	282,7	281,7	**0%**
Illegal parking time violation	32682	24085	33154	5230	12269	3115	18422,5	6871,3	**-63%**
Parking weight points				757217	673770	710426		713804,3	
Candidates checked	28421	28975	28167	32015	28973	32239	29798,3	31075,7	**4%**