Thèse

présentée pour obtenir le grade de

Docteur de l'Université de Lille

Spécialité : Informatique

présentée par

Franck Kamenga

Optimisation combinatoire intégrée de la gestion du matériel roulant et de la circulation ferroviaire dans les gares de passagers

Combinatorial optimization for integrating rolling stock management and railway traffic scheduling in passenger stations

soutenue publiquement le 20/11/2020 devant le jury d'examen:

Prof. Dr. Francesco Corman	Professor, ETH Zurich	Rapporteur
M. Fabio Furini	Researcher, Consiglio Nazionale delle Ricerche	Rapporteur
Mme. Sourour Elloumi	Professeure, ENSTA Paris	Présidente
M. Joaquín Rodriguez	Directeur de recherche, Université Gustave Eiffel	Directeur de thèse
M. Boubekeur Merabet	Docteur, SNCF Réseau	Examinateur
Mme. Paola Pellegrini	Directeur de recherche, Université Gustave Eiffel	Examinatrice
M. Bertrand Houzel	Directeur de département, SNCF Réseau	Membre invité

préparée à l'Université Gustave Eiffel COSYS-ESTAS
20 rue Elisée Reclus, 59650 Villeneuve D'Ascq
École Doctorale 72 Sciences pour l'ingénieur

Résumé

Les gares ferroviaires concentrant les fins et les débuts de trajets des trains structurent l'essentiel de l'exploitation de lignes pour passagers. En effet, on y programme des opérations de préparation du matériel roulant (nettoyage, accouplement des trains. . .) dites de « produit train » indispensables à la qualité de service. Ces opérations imposent des manœuvres et nécessitent de garer des trains.

Cette thèse aborde de manière intégrée la planification des opérations de produit train et la gestion de la capacité en gare. Elle introduit pour cela le Generalized Train Unit Shunting Problem (G-TUSP). Il s'agit plus précisément d'affecter des trains arrivant dans une gare à des départs et des voies de garage et d'ordonnancer leur maintenance et leurs manœuvres. Ces décisions sont prises afin de minimiser les retards au départ, les accouplements et désacouplements de trains et les annulations de départ ou de maintenance. Le G-TUSP possède des contraintes liées à des caractéristiques techniques du matériel roulant et de l'infrastructure ainsi qu'à la nature des opérations réalisées. Le G-TUSP comporte quatre sous-problèmes, souvent traités indépendamment dans la littérature. Cette thèse propose des algorithmes d'optimisation comme outils d'aide à la décision pour les planificateurs du produit train.

Une formulation en programme linéaire à variables mixtes est établie en considérant une représentation détaillée des aspects du G-TUSP. La formulation est testée sur des instances réelles de la gare Metz-Ville et des résultats pertinents sont obtenus en une heure de calcul. Nous proposons ensuite des algorithmes dans lesquels nous considérons différentes combinaisons d'approches séquentielles ou intégrées pour les sous-problèmes du G-TUSP. Dans une analyse expérimentale détaillée basée sur des instances de la gare de Metz-Ville, nous étudions la contribution de chaque sous-problème à la difficulté du G-TUSP et nous identifions le meilleur algorithme. Cet algorithme donne des résultats très satisfaisants en moins de vingt minutes.

Abstract

Railway stations that concentrate starts and ends of train journeys structure most of the passenger lines operations. Indeed, rolling stock preparation operations (cleaning, trains coupling. . .) which are called shunting are scheduled there. These operations are essential to ensure service quality. These operations require train movement and parking.

The thesis tackles an integration of shunting operation planning and capacity management in railway stations. The Generalized Train Unit Shunting Problem (G-TUSP) is introduced to consider this integration. In the G-TUSP we assign trains which arrive in a railway station to departures and parking tracks and we schedule their maintenance operations and their movements. These decisions are made to minimize departure delays, coupling and uncoupling operations and maintenance or departure cancellations. The G-TUSP has constraints due to rolling stock and infrastructure characteristics or related the nature of the operations carried out. The G-TUSP includes four sub-problems, often considered independently in literature. The thesis proposes optimization algorithms as decision support tool for shunting planners.

A mixed integer linear programming formulation which considers a detailed representation of G-TUSP aspects is set. The formulation is tested on instances representing traffic at Metz-Ville station. Relevant results are obtained within an hour of computation. Then, we propose algorithms in which we consider different combinations for the integrated or sequential solutions of the G-TUSP sub-problems. In a thorough experimental analysis, based on Metz-Ville station instances, we study the contribution of each sub-problem to the difficulty of the G-TUSP, and we identify the best algorithm. This algorithm returns very satisfying results in less than 20 minutes.

Contents

Chapter 1

Introduction

1.1 Context and motivation

Railway transportation in Europe is based on networks set in the 19^{th} and 20^{th} century. In France, the railway network covers most of the territory with 30 000 km of tracks. Because of its small ecological footprint per passenger transported, rail transport is a pillar of passengers mobility policies. European, national and regional policies aim to encourage modal shift from private car or plane to train. This modal shift needs an improvement of rail transport service quality. In the European norm (Comité Européen de Normalisation [2002]), service quality includes, among other factors, trip duration, frequency and comfort. Delays affect trip duration and can propagate easier in congested portions of a network. Comfort is improved with trains that have a suitable seat capacity. The cleanliness of trains and the proper functioning of interior equipment such as lights or air conditioning contributes to comfort. In stations, significant service quality features are involved. Indeed, trains are prepared in station to meet the comfort requirement, while avoiding delays with high traffic density.

Railway service is based on a timetable to satisfy passengers demand. It involves the use of an infrastructure, rolling stock and crews. There is an accurate planning so that required resources are in the right place at the right time. The use of each resource is usually complex to handle. This is why planning is mostly sequential and each resource is tackled separately. Efficient resources consumption is sought to minimize operation costs with a convenient service quality. On the one hand, rolling stock has significant purchase and maintenance costs. These costs are amortized by maximizing the journeys made by a train. On the other hand, major investments and operating costs of infrastructure encourage the passage of as many trains as possible to satisfy demand. Often, railway service planners face infrastructure capacity utilization issue. Rolling stock and infrastructure capacity are expensive resources handled in railway planning.

Operations research and combinatorial optimization, in particular, have been relevant fields for railway planning improvement. Indeed, they allow formal modeling of problems for which algorithms seek optimal or good solutions. In particular, scheduling and routing problems have been widely tackled. In such problems tasks have to be processed using specific resources so that no (or minimum) delay occurs. The literature contains many optimization approaches for infrastructure capacity or rolling stock utilization. Some of these approaches are used by planners thanks to software tools. Then manual processing and errors can be avoided and a further analysis of resource usage and quality of service is possible.

Sophisticated planning tools exist at the level of a railway line or of a region in France. However, they neglect the details of station operations. *Shunting*, which is the preparation of trains in a station between two trips, is planned manually between one month and few hours before operations. In this preparation, the management of rolling stock strongly interlinks with the use of station capacity. In shunting, trains can be split, coupled or kept as they are to cope with a rolling stock utilization defined at a larger scale. Such operations need specific tracks called shunting tracks and train movements between shunting tracks and between shunting tracks and platforms. In 2018, 13% of delay minutes in the French railway network involve shunting (Martin et al. [2019]). Indeed, shunting operations may be delayed because of failures. Moreover, the movement of a train from a shunting track to its departing platform has to be inserted in the rest of the station traffic. A such movement does not often have priority, which causes departure delays.

Shunting planning is a significant lever for service quality improvement. A great part of the combinatorial optimization literature tackles resource consumption in shunting. The contributions focus on either capacity consumption, rolling stock or shunting crews use. Nevertheless few approaches consider the whole shunting process from a train arrival to its departure. They do not consider the possibility of accepting delays or maintenance operations cancellations. The study of the whole shunting process highlights service quality precisely. Indeed several scheduling or routing decisions are linked and need to be consistent.

1.2 Research objectives

This thesis aims to propose **optimization approaches that integrate both rolling stock preparation and station capacity management for shunting planning in passengers stations**. Rolling stock is targeted in terms of train units, as they constitute the bulk of the passenger transport fleets in Western Europe. The approaches proposed consider practical aspects encountered in passenger stations. They aim at suitable computation times in order to be used in preoperational planning software systems.

The main objective of the research (O) is to **set a model and an algorithm for scheduling shunting operations such as coupling or splitting, maintenance and shunting movements**. A such model shall consider decisions that involve rolling stock and capacity utilization. The proposal must deal with movements between shunting tracks and between shunting tracks and platforms. Shunting movements have to be considered at microscopic scale in order to tackle infrastructure capacity consumption wisely. Service quality issues, such as delays or comfort of rolling stock, must be taken into account. The research output must meet two sub-objectives in order to facilitate their use by planners:

- The optimization approach must provide feasible solutions that meet practical requirements. The performance of the solutions must be satisfactory in the eyes of the planners. (O1) **Solutions have to be at least as efficient as the ones obtained manually by planners**.

- (O2) **The algorithms proposed must be able to deal with real-life instances** whose size can be significant because of microscopic modelling. Reasonable computation times have to be sought.

1.3 Contributions

The objectives mentioned in the previous section are achieved through two main contributions.

First, we introduce the Generalized Train Unit Shunting Problem (G-TUSP). It is an extension of the Train-Unit Shunting Problem (TUSP) introduced by Freling et al. [2005] to maintenance and movements scheduling. Such a comprehensive extension has never been tackled in the literature. In the G-TUSP, shunting movements are taken into account thanks to a microscopic representation of the infrastructure. The G-TUSP handles both departure assignment, shunting track assignment, path assignment, maintenance scheduling and shunting movements scheduling. All the mentioned decisions are integrated in a mixed-integer linear programming (MILP) formulation which considers a range of practical constraints typically encountered in reality. The formulation is tested on real instances of Metz-Ville station in France. The results show that the G-TUSP can be solved in a reasonable time with satisfying results. Moreover, experiments highlight the relevance of integrating shunting decisions. This formulation has made the object of a paper published in the proceedings of the international conference RailNorrköping 2019 - 8th International Conference on Railway Operations Modelling and Analysis (ICROMA) (Kamenga et al. [2019b]) and has been presented in the 20th congress of the French Operations Research & Decision Support Society (Kamenga et al. [2019c]).

Second, we propose sequential algorithms to solve the G-TUSP with lower computation times. These algorithms are based on the division of G-TUSP into sub-problems. In each algorithm a group of sub-problems is solved exactly in an integrated formulation, while the other sub-problems can be solved with a heuristic or another mathematical programming formulation. The algorithms are implemented to tackle instances that include several types of disturbances. Convenient results are obtained with relatively low computation times. The relevance of sub-problems integration is also investigated. These algorithms have been presented in the 12th World Congress in Railway Research (WCRR) (Kamenga et al. [2019a]) and in the 21th congress of the French Operations Research & Decision Support Society (Kamenga et al. [2020]). They make the object of an article currently under review for possible publication in an international journal.

1.4 Outline of the thesis

The rest of thesis is organized as follows.

In Chapter 2, we provide an overview of the railway system and present passengers rail transportation planning process and issues. We consider all resources that are needed to provide the railway service. Moreover, we highlight the characteristics of shunting in passengers operations planning in France.

In Chapter 3, we formally define specific problems related to shunting. These definitions are used in the whole thesis. We also provide a literature review of combinatorial optimization applied to shunting problems. The literature tackles each problem individually or integrate some of them.

In Chapter 4, we propose a MILP formulation for the G-TUSP. The formulation is based on a microscopic model of the infrastructure and formal train units in order to consider coupling and splitting. The relevance of the formulation is discussed with experimental tests on real-life instances at Metz-Ville. In particular, solutions obtained with the formulation are compared with planners decisions.

In Chapter 5, we deal with sequential algorithms for the G-TUSP. We propose heuristics and MILP formulations to solve groups of G-TUSP sub-problems. We propose experiments based on the real instances of Chapter 4 and fictive instances based on Metz-Ville traffic data. Thanks to this larger set of instances, we compare the performances of algorithms and we analyze the impact of sub-problem integrations.

Finally, in Chapter 6, we summarize the main contributions and results presented in the thesis and provide future research directions.

Chapter 2

Railway system and passenger transportation planning

2.1 Introduction to the railway system

The railway system is based on technical components that allow the execution of train traffic. These components include ground installations such as tracks, electrification systems and signaling installations. Ground installation form the *infrastructure*. The railway system also includes *rolling stock* which is designed to meet infrastructure characteristics and transportation requirements. Human actors and organizations are also involved in the system. They operate and plan railway services. In Europe, since 1991 and following a series of directives that aim at opening up the access to rail networks, infrastructure management missions and service operations have been separated. Previously, historical rail operators were in charge of all processes, typically one operator per country. *Infrastructure managers* own, maintain and develop a railway network on which they manage traffic. For example SNCF Réseau, ProRail, NetworkRail, Infrabel and RFI respectively are the main railway infrastructure manager of France, Netherlands, United Kingdom, Belgium and Italy. *Train operating companies* operate trains for passengers or freight transportation. Train operating companies have to pay fees to infrastructure managers which own the network where they operate. In this chapter we focus on passenger transportation.

Particular resources have to be deployed to allow the execution of rail services. Infrastructure, rolling stock and crews (drivers, guards or conductors, ground agents...) must be available when a service has to be performed. We note that rail services are based on a *timetable*. A timetable informs passengers of times at which trains stop at stations. A timetable can give more detailed information for infrastructure managers and train operating companies. The time at which trains pass signals can be indicated as well as the rolling stock used. Thanks to a timetable, it is possible to check if each necessary resource is available. For example, a train cannot be used for two trips at the same time and for safety reasons two portions of an infrastructure cannot be used by two trains at the same time. Slots have to be reserved for each section of the infrastructure a rail service has to use. This reservation is a *train path*.

In this section, we introduce concepts that deal with compatibility between resources and service quality. In particular, we deal with infrastructure and rolling stock.

2.1.1 Infrastructure capacity

Infrastructure capacity goes hand in hand with its ability to satisfy a timetable. A timetable may contain different numbers of trains per day or hour, with different speeds or stops. Capacity can be defined as the number of train paths that can be allocated in a time horizon to meet service requirements. This concept is actually hard to tackle, because a relevant part of infrastructure must be identified and service requirements must be defined. The UIC 406 UIC [2013] proposes methodological aspects to consider for lines and nodes. These are based on a timetable compression. In this section, we focus on capacity consumption due to train paths. We tackle characteristics of infrastructure and rolling stock that impact capacity consumption.

Capacity and signaling system

The signaling system is based on components such as signals, switch command systems and track equipment to ensure traffic safety Rétiveau [1987]. To avoid collision risk, train cannot enter a zone where another train is running. First, in order to satisfy this requirement, a signaling system needs the detection of trains presence on a specific zone. This detection is performed with *track-circuits*, which are electrical circuits that use the conductivity of a portion of the rail. An electrical shunt occurs when a train is on this portion. Then, the signaling system uses signals that have to be followed by train drivers. There is a *stop signal*, which has a red aspect [1], before entering an occupied zone. Upstream, a *warning signal*, which has a yellow aspect, indicates to the driver to reduce speed in order to be able to stop the train at the next signal if it still has a red aspect. A *clear signal*, which has a green aspect, indicates that a train can pass with the allowed maximal speed. These three types of signal ensure the separation of trains: this is *three aspects signaling.* Some signaling systems may contain a sequence of several different warning signals before a stop signal. In this case we can have a four or more aspects signaling. A *signaling control* commands switches and signals in a defined part of a railway network called *control area.* In a control area, a *route* links an origin signal and a destination signal through a sequence of track-circuits. A *block section* in a route is a set of track-circuits between two consecutive signals. If a route contains switches, their position must lead a train that follows the route to the right direction. Routes are set by a signaling control. A route can be set once switches are commanded and the origin signal is not a “stop" one. For safety reasons a switch should not be commanded, and an opening signal aspect can be clear, when a train uses the same or a conflicting block section. Signaling control includes the *interlocking system*, which ensures that no route can be set if logical safety conditions are not fulfilled. Many interlocking systems are based on electro-mechanical relays or electronic circuits.

When a train goes through a route, several stages are considered. Indeed, track-circuits of the route are successively utilized (Figure 2.1). To do so, first the route must be formed through the block section bs that contains a track-circuit tc: necessary switches must be set to the correct positions and the origin signal of the block section must not indicate to stop. The time needed to form a route through a block section is the *formation time.* If the route is not formed, the origin signal of the block section is a “stop", then according to the number of aspects of the signaling system, upstream signals indicate a “warning". In the case of a three aspect signaling, the signal preceding the origin signal has to be a “warning". More generally, with n aspects signaling the $n-2$th, $n-3$th,... signals before the origin signals have to be “warnings". Without loss of generality, hereinafter we consider a three aspects signaling. We also suppose that trains run passing through clear signals. Once the route is formed. The signal at the block section bs' which precedes bs is “clear". Then, the head of the train goes through bs'. The *reservation* of bs contains this time. Then, the head of the train runs through track-circuit tc. The duration of this step is the *running time.* Once the head of the train arrives at the next track-circuit, the tail of the train remains on track-circuit tc. A *clearing time* separates the moment when the head of the train arrives at the next track-circuit and the moment when the tail of the train leaves track-circuit tc. We note that track-circuit tc is occupied by the train between the moment when the head of the train enters in tc and the moment when the tail of the train leaves tc. The duration of this occupation is the *occupation time.* Finally, once the train leaves track-circuit tc, the state of the track-circuit changes after a *release time.* Therefore, the utilization of a track-circuit includes formation and reservation time, occupation time and release time. Two trains can use a same time track-circuits if their utilization intervals do not overlap.

[1] In mainlines, French signaling distinguishes two stop signals. The *carrée*, whose light signal has two red lights, protects switches and cannot be passed by a driver without an authorization. The *sémaphore*, whose light signal has one red light, ensures train spacing on a line. *Sémaphore* can be passed by a driver under some conditions.

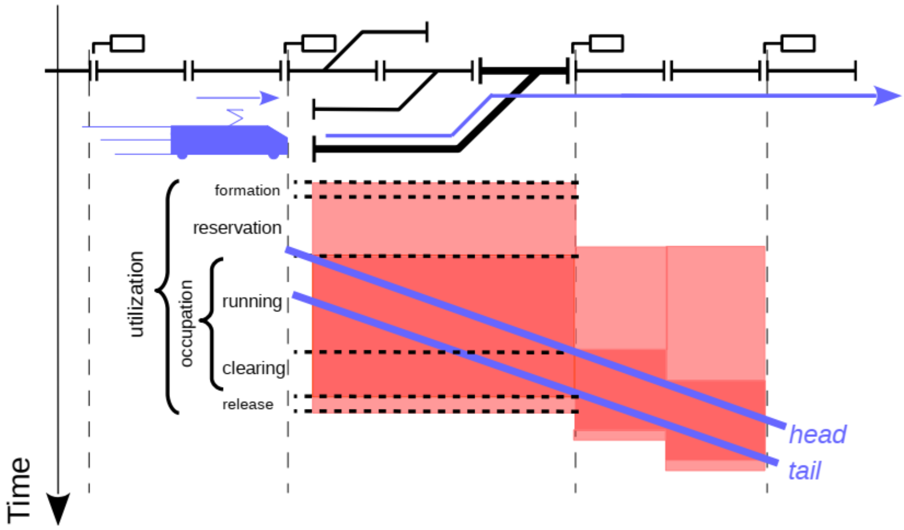


Figure 2.1: Utilization of track-circuits by a train that follows the blue route in the case of a three aspect signaling system

Railway station capacity

In passengers railway stations, trains can pass through, stop, be coupled to other trains or parked. A station has platform tracks on which trains can stop to board and unboard passengers. *Switch areas*, in green in Figure 2.2, gather switches in order to move trains from and to the platform tracks. Platform areas, in red in Figure 2.2, and switch areas are part of *main tracks*. Main tracks are used for traffic along a line or stops at platforms. Stations also contain *shunting tracks*, which lead to facilities (maintenance equipment, refueling...) or are used for parking. A group of parallel shunting tracks is a shunting yard, as the ones represented in blue in Figure 2.2. When trains have to move in, from or to a shunting yard,they perform a *shunting movement*. A shunting movement may pass through main tracks. According to the regulations of many European networks, a shunting movement is executed at a low speed (e.g., less than 30 km/h in France and Belgium) so that the driver can be able to stop at any obstacle as soon as he sees it.

In order to tackle capacity in a station we need to consider:

- regular train movements that only pass through main tracks and are set in a timetable.
- shunting movements whose schedule is typically not planned in advance.

Shunting movements, that run slower, usually consume more capacity than regular train movements. We can also note that capacity consumption includes stopping times on platforms or shunting tracks. Therefore capacity at stations is based on track-circuit utilization by regular or shunting movements in switch areas. In platform and shunting area, the simultaneous use of a same track is also taken into account. Indeed, under some conditions defined by a station regulation, a train can enter a platform or shunting track when it is occupied in order to get coupled or to use allow the use of fewer tracks.

2.1.2 Rolling stock management

A train is made of a motive power system and passenger cars. One or several locomotives can contain the motive power system and pull or push passenger cars. Some trains with locomotives can be driven at both sides, those are *push-pull* trains. There can be a locomotive at each end of the train, or one locomotive and a cab in the passenger car at the opposite end. If a train is not push-pull, a turnaround

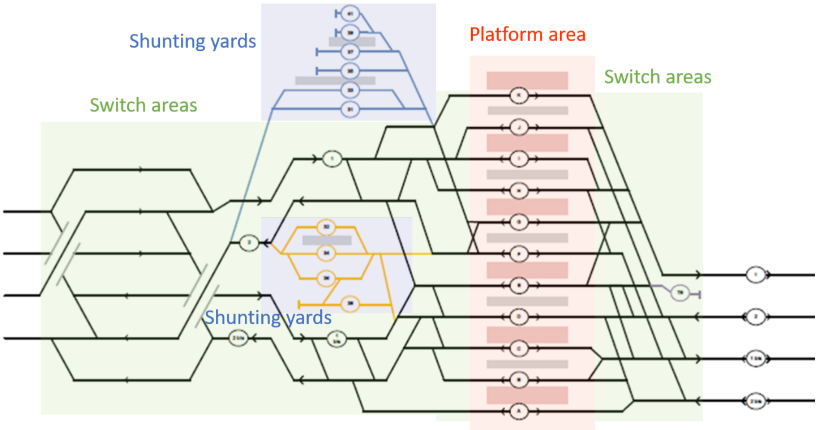


Figure 2.2: Example of a station layout: Lyon Part-Dieu station, in France

usually requires to remove the locomotive at one end of the train and hang another locomotive at the other end. A *train unit* (also called multiple unit) gathers motive power system and passenger cars which cannot be uncoupled during operations. Train units can often be driven at both ends and can be coupled with other train units. Locomotives or train units can be powered by electricity or fuel (usually diesel). In particular, electric trains can only move on an electrified infrastructure with specific characteristics. European countries can have different electrification systems which are not compatible with all trains. For example, France has two electrification systems. Trains must also respect gauge or length restrictions depending of the network they run on. Trains that have same characteristics belong to a same rolling stock *type*. Train units that belong to different rolling stock types, can be coupled if their rolling stock types share some technical features: rolling stock types can be *compatible*.

Once a train is acquired by a train operator, maintenance must be done. Maintenance makes rolling stock able to ensure service quality. Indeed, it prevents incidents and provides clean trains with comfort equipment (seat, light, air conditioner...) in good conditions. A great part of rolling stock maintenance is preventive: components are checked, repaired or replaced according to nominal wear indications rather than actual failure. Operations are done according to the total distance traveled or to specific temporal frequencies. In France, maintenance activities are structured into five levels. The higher the level of maintenance, the heavier the tasks performed. Interventions are more frequent for low level maintenance. Level 1 maintenance is performed daily and does not need any specific facility. In level 1 maintenance, drivers or conductors check that safety devices work at the beginning of a day. Level 2 maintenance contains operations performed in a shunting yard that has specific facilities. Level 2 maintenance is made between two rail services that a train operates, so during off-peak hours or during the night. Level 2 maintenance contains light operations such as cleaning, technical checks of mechanical components, motorization and interior fittings. The frequency of most these operations varies from 4 times a week to once every 15 days. Level 3 maintenance gathers heavy maintenance operations. Such operations require immobilizing a train for several days for in-depth examinations. Level 3 maintenance is executed on a train every one or two years in a maintenance center. Level 4 and level 5 maintenance gather industrial maintenance operations. These operations are performed on a train every 10 years or more in industrial maintenance centers. Industrial maintenance contains complete overhaul or modernization operations.

Our study focuses on push-pull trains or train units, that we both call train units. Light maintenance, i.e., level 2 maintenance in France, is tackled in the rest of the thesis.

Table 2.1: Phases in the tactical planning of rail services (France)

phase	preconstruction	construction	adaptation	preoperational
time horizon	1 to 5 years	6 months to 1 year	6 days to 1 month	few hours to 6 days
focus	line planning periodic timetabling	timetabling day time rolling stock circulation	timetabling rolling stock circulation crew scheduling	timetabling rolling stock circulation crew scheduling
local scale		platforming	platforming shunting	platforming shunting

2.2 Passengers railway operations planning

2.2.1 Overview of passengers operations planning

Railway services are planned so that necessary resources which are infrastructure capacity, rolling stock and crews can be deployed. At a strategic level resources are sized. The decision horizon is often larger than five years. Strategic decisions include the purchase of rolling stock and infrastructure building or upgrading. At tactical level, rail services are defined and resources are assigned. The decision horizon is between five years and one day before operations. Tactical decisions include timetabling as well as rolling stock and crew assignment. At operational level, traffic is managed in real-time. The decision horizon is often less than few hours. Operational decisions may change timetable or resource assignment to cope with disturbances.

In France, the timetable for a year is called the *annual service*. It is obtained thanks to tactical decisions which occur in four stages (SNCF Réseau [2018]) reported in Table 2.1.

The first stage is the *preconstruction*. Preconstruction decisions are made between five years and one year before the annual service. The basic framework of the railway service is defined. During this stage a systematic timetable diagram for a 2 hour period and for a 24 hour period is made. Diagrams are obtained by defining first systematic paths with stops. Such paths are called lines and defining lines is *line planning*. The frequency of lines is set in order to satisfy an expected demand. Then a periodic *timetabling* is made. During periodic timetabling, train paths of the lines are scheduled.

The second stage is the *construction* and occurs between one year and six months before operations. During this stage train operating companies request train paths to the infrastructure manager. If a request is accepted, it is scheduled and appears in the annual service: *timetabling* is performed. At stations a platform track is assigned to each train path: this is *platforming*. Usually, the requests are consistent with the timetable diagrams obtained in the preconstruction phase. Nevertheless, timetabling in construction phase is not periodic, for example holidays may change the passenger service. Once train paths are examined, train operating companies make a daytime *rolling stock circulation* plan. In a rolling stock circulation plan, a train is assigned to each train path.

Adaptation is the third stage of annual service planning. It takes place between six months and six days before operations. Train paths defined in the construction phase can be modified because of infrastructure works or special events. In adaptation phase, a complete rolling stock circulation is set. It consider maintenance as well as parking in nighttime or off-peak hours. In particular, *shunting* planning considers trains without passengers, which are parked or maintained in stations. A *crew scheduling* is also set.

The final stage is the *preoperational* one and occurs between six days and few hours before operations. In this stage, last minute requests and planned or expected disturbances such as failures or strikes are considered. A final timetable and resource assignment is set to be used by dispatchers.

Railway planning is mostly resource centered. A planning problem deals with the use of specific resources. For example timetabling deals with the use of infrastructure capacity while rolling stock circulation concerns the use of rolling stock. Resources can be considered in a regional or national scale in which services are treated all along their path. Local scale planning details resource assignment or scheduling at major stations.

Many decision support tools or methods are used by infrastructure managers and train operating companies for service planning. Planners often use software tool which check the feasibility and precise the outcome of an assignment or a timetable. Software tools can go further and explore a set of possible decisions to provide an optimal decision. Such tools integrate combinatorial optimization techniques. The interest for operation research in passenger railway planning has grown in the last decades: as train operating companies face competition in Europe, they aim to improve their performances. Moreover, infrastructure managers aim to optimize the use of capacity. Operations research have proposed relevant modeling approaches and algorithms to solve planning problems.

In Section 2.2.2 we give some details about capacity use problems, and in Section 2.2.3 we focus on rolling stock. We refer the interested reader the literature (e.g., Caprara et al. [2007], Huisman et al. [2005] and Kroon et al. [2009]) for a deep review of operations research applications to European passengers rail services.

2.2.2 Train timetabling

In timetabling, train path requests have to be satisfied while meeting capacity constraints due to infrastructure and rolling stock features. Capacity constraints state that every train path must run with "clear" signals. Train path requests are made by train operating companies and specify rolling stock used as well as desired schedules at stops. A train path request also includes commercial requirements such as dwell times at stops or time windows for connections. A train path request is accepted if a path in the infrastructure and conflict-free schedule can be found. There may be a deviation between schedules desired by the train operating companies and schedules proposed by the infrastructure manager.

In France, timetabling is made by the capacity allocation division (DAC) of SNCF Réseau. The DAC completely handles long distance train paths and devolves regional train path requests to regional timetabling divisions. Local timetabling divisions focus on platforming at major stations. Planners use descriptive software tools which check traffic conflicts.

Timetabling can be cyclic or acyclic. Infrastructure can be represented at different scales to model traffic conflicts. Macroscopic scales consider lines and stations in which a maximum flow is specified. Microscopic scales involve signals and track-circuits and take into account track-circuit utilization. A such scale increases the size of the problem tackled but provides an accurate model closer to reality. Intermediate scales aggregate some elements of the infrastructure in order to focus on bottlenecks. Timetabling can also be considered in discrete time or continuous time. In French, national network schedules have up to 10 seconds precision.

Platforming

Platforming is timetabling at a station scale. The infrastructure considered is often a station control area. As running times through such areas are often less than ten minutes, little flexibility exists for scheduling decisions while the opposite holds for routing decisions, for which there is a significant number of paths. Platforming is based on a regional timetable which indicates time at which train paths leave and enter the control area. Platforming must also satisfy commercial requirements based on passenger flow or connections. For example, at some stations two trains should not stop on two adjacent tracks of a same platform at the same time, unless the platform is large enough to carry the passenger flow. Otherwise, as major stations are often destination of rail services, trains stay at destination station until their next trip. Therefore, rolling stock assignment affects track occupation at stations. The longer trains stay in a station, the more platforms or shunting tracks are needed.

A platforming can be represented with a track occupancy diagram, which indicates the time at which a platform track is used by trains (Figure 2.3). Each row of the diagram represents a platform track and the horizontal axis represents time. A rectangle is drawn when a train is parked on a platform track. At

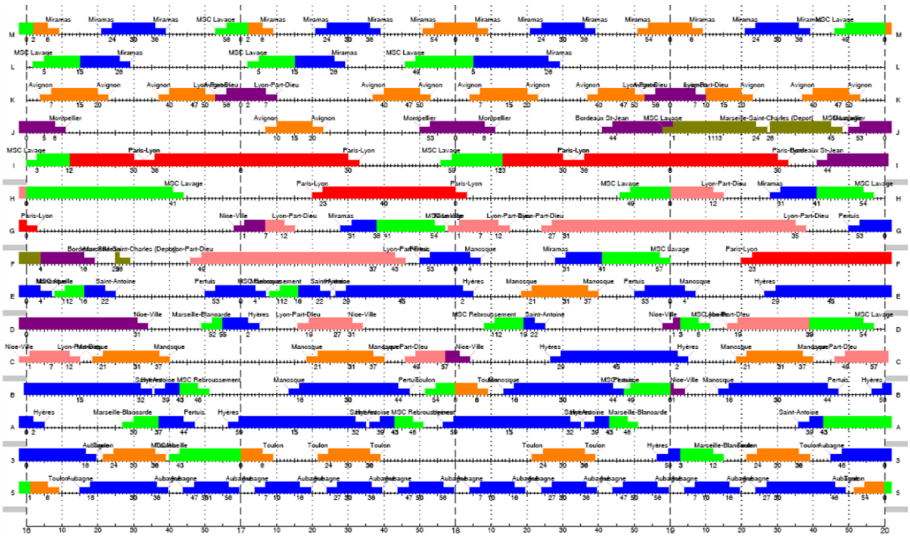


Figure 2.3: Track occupancy diagram at Marseille Saint-Charles station: each row represents a platform track; a rectangle indicates the time at which the platform track is occupied by a train (SNCF Réseau)

the left of each rectangle, a smaller rectangle indicates the travel time in the control area to the platform track, while at the right another small rectangle indicates the travel time in the control area from the platform track.

The software tool `OpenGOV`https://www.sncf-reseau.com/fr/solutions-innovantes-exploitation-ferroviaire developed by SNCF Réseau is used by around twenty local timetabling services in France. OpenGOV is based on stable set model solved with either linear programming or local search (Kamenga [2016]).

2.2.3 Rolling stock circulation planning

In rolling stock circulation planning, train operating companies need to cover all train paths with an available number of trains. A train can be assigned to a trip if it respects technical characteristics such as gauge, electrification or length and its seats can carry the expected demand. In a fleet made of train units, trains can be coupled or uncoupled during or between trips to adapt seat capacity. The train units coupled have to belong to identical or compatible rolling stock types. When a train finishes a trip at a station, its next trip can start either at the same station or at another one, then trains may run empty between these stations. The cost of rolling stock circulation integrates the distance traveled by trains. Rolling stock circulation contains maintenance, namely preventive maintenance that is operated depending on distance traveled or time frequency. Capacity of shunting yards used to park trains in daytime or nighttime is considered. In a fleet made of train units, constraints due to position of train units in a train are also taken into account. When train units are coupled and uncoupled, the trip assignment must be consistent with the position of train units. For example, if a train made of two train units finishes a trip and gets uncoupled at a terminus station, which has dead end platform tracks, the original head train unit is at the bottom of the station. Therefore, after uncoupling, the head train unit has to leave the station last. The head train unit cannot be assigned to a train path that leaves the station before the train path assigned to the tail train unit.

In the construction phase, a year before operations (Table 2.1), a daytime rolling stock circulation is made. Each day, trains are assigned to train paths. Nevertheless, the position of trains at the beginning and at the end of the day is not considered. In a complete rolling stock circulation, days are linked by including night time parking or maintenance. In France, fleet managers define maintenance tasks that have to be performed on trains. Light maintenance (level 2) can be performed on shunting yards of stations as well as in maintenance centers. Heavy maintenance (level 3) can only be performed in maintenance centers. Maintenance tasks are grouped into maintenance cycles. All the tasks of a maintenance cycle can be performed in few hours in case of light maintenance or several days in case of heavy maintenance. Fleet managers set maintenance cycles each time a train is parked at a shunting yard or a maintenance center. In the adaptation or preoperational phase, they must also change the rolling stock circulation when a train needs an extra maintenance because of an incident. This adaptation is the *maintenance routing*.

2.3 Shunting and local scale planning

Shunting deals with local scale issues. A limited space around a station is available. This space can be delimited by a station control area which includes platform area, switch and shunting yards. A maintenance center close (less than 1 km) to the station can also be included. In the common definition, shunting gathers operations that require shunting movements. A shunting movement, as mentioned in Section 2.1.1, is restricted to the local scale and is made so that trains change tracks for operations (coupling, maintenance...). This capacity centered definition is used by SNCF Réseau and is translated in French as *manœuvre*. SNCF Voyageurs, the main French passenger train operating company, prefers a rolling stock centered definition. This definition is translated in French as *produit train* and refers to the set of technical operations carried out at a station to prepare trains for departure. These technical operations include coupling and uncoupling as well as light maintenance. As such operations are needed, trains have to be parked on shunting tracks and shunting movements have to be performed.

If we combine the rolling stock centered definition and the capacity centered one, shunting includes coupling and uncoupling operations, light maintenance, parking and shunting movements. Hence, shunting planning must consider several resources:

- Infrastructure capacity, since parking and shunting movements require capacity
- Crews, maintenance crews are involved for light maintenance, drivers and ground agents are involved for coupling, uncoupling and shunting movements
- Rolling stock, since trains must be available to ensure departures.

Shunting occurs once a train ends its trip at a station and all its passengers get off. The train can be moved to a shunting yard and may be uncoupled if it is made of several train units. Then maintenance tasks are performed on the train units at specific shunting tracks which have appropriate equipment. Finally, after possible coupling operations, a train is moved from a shunting yard to a platform before its departure.

Shunting planning follows rolling stock circulation planning and involves several decision makers. In France, ground operations services work for train operating companies. They are, among other things, in charge of trains once they get empty. They supervise most of shunting operations. Ground operations planners are in charge of shunting planning. Ground operations services deploy ground agents for coupling and uncoupling and drivers, who can be assisted by a ground agent, for shunting movements. Maintenance is carried out by agents who belong to a maintenance service or an external service provider. Maintenance operation managers schedule maintenance operations. Shunting movements have also to be scheduled and routed. If a shunting movement goes through main tracks it can meet regular movements. In this case, routing and scheduling shunting movements can be considered as a part of platforming. Therefore, they can be treated by station traffic schedulers. In France, station traffic schedulers are local timetabling divisions. A shunting movement is *timetabled* if it is routed and scheduled at least the day before operations. Shunting movements which do not go through main tracks are not precisely planned. Often, shunting movements are scheduled so that the number of simultaneous movements is limited. Then, the shunting movement is routed in real-time according to the current situation. A such movement is not timetabled. There are stations in which even shunting movements that go through main tracks are not timetabled. Organizational variants exist in French stations for planning shunting operations.

Performance of shunting is traditionally tackled by two points of view. For infrastructure managers, efficient shunting minimizes its use of capacity. The performance indicators can be the number of shunting tracks that remain free during a period or the number of shunting movements planned during a period. For train operating companies, the service quality indicators matters, i.e., trains must be in good condition and on time. Recently in France, SNCF Réseau and train operating companies tend to agree on common objectives. Service quality for passengers has to be a priority. In particular, punctuality is a priority that leads to the "H00" program. The objective is an increased punctuality in the whole French national network. The "H00" program includes actions and studies carried out on practices and tools. This program tackles operations scheduling in particular when several actors interact. A part of the program focuses on shunting operations which can be a cause of delay. Therefore the improvement of shunting planning has recently become a research issue. Most of shunting planning is currently manual. Each actor has its own tool which does not always interact efficiently with other ones. A global vision of shunting can help to know if there is room for improvement. Therefore, it is necessary to avoid a non-integrated approach for which it is difficult to asses the quality of shunting planning in terms of punctuality, comfort and use of infrastructure capacity.

2.4 Conclusion

In this chapter, we introduced the main resources and planning process involved in passenger rail service. We focused on planning at stations and we remarked that resources such as rolling stock, infrastructure capacity and crews are all involved in shunting operations. We also presented the organization of shunting planning in France as well as performance issues.

In Chapter 3, we will review operations research approaches that tackle rolling stock preparation and station traffic scheduling in shunting problems.

Chapter 3

Overview of shunting problems modelling approaches

In passenger railway transportation, shunting is the management of rolling stock when it is not used for a trip. Shunting usually takes place during nights or off-peak hours. Shunting concerns rolling stock rotation given the layout of a station. We focus on train units, since they are widely used for suburban, regional and high speed transit in Europe. We do consider locomotives and coaches management.

In this chapter, we focus on shunting planning in a railway station. As discussed in Chapter 2, in France shunting operations are considered during two phases of service planning:

- The adaptation phase (six months to six days before operations) which uses a nominal timetable,
- The preoperational phase (six days to few hours before operations) in which planned disturbances are considered.

We present four main problems related to shunting operations. We depict a shunting system set up and provide a literature review on modeling approaches for solving shunting problems.

3.1 Shunting problems description

3.1.1 Shunting problems setup

Shunting is the management of train units which stay long enough in a station to be parked at specific yards called *shunting yards*. Shunting problems are based on several characteristics of infrastructure or rolling stock as well as planning inputs.

First, stations have main tracks where regular train movements go through. At some of these main tracks there are platforms for passengers boarding. Stations also contain *sidings* which are tracks on which trains move at a lower speed and can be parked. These tracks may have technical equipment for maintenance, cleaning or refueling (for diesel trains): Maintenance may require a pit for engine examination or a gangway for roof examination. External cleaning can be performed at a track with a washing machine. It is an automatic washing installation, similar to automatic car wash, set along a track which has rollers and liquid jets. Sidings can have various topologies. There are dead-end tracks on which train units must respect *last-in first-out (LIFO)* rule. These sidings are called LIFO tracks. There are other tracks with two access sides: FIFO tracks and regular tracks. In *first-in first-out (FIFO)* tracks, train units can enter at one side and leave at the opposite side. In *regular* tracks, train units can enter or leave at both sides. Tracks also have a limited length and are not necessarily electrified. This can prevent electrified train units from parking or moving through. A *shunting yard* contains several parallel sidings with common access sections. Then, sidings which belong to a shunting yard can also be called *shunting tracks*. A *shunting neck* is a siding on which a turnaround can be performed so that trains can change track in a shunting yard. Figure 3.1 shows the layout of Metz-Ville station, in eastern France. The station contains four shunting yards dedicated to passenger trains. Shunting yards are green squared in Figure 3.1 and shunting tracks are represented with a bold line. For example, track 90 at the top of Figure 3.1 is a LIFO track, while track 15 at the bottom is a regular track. Track 82, which is drawn in brown in Figure 3.1, has a gangway and track 24, drawn in light blue, has a washing machine. There are two shunting necks on the north side and two shunting necks on the south one, which are red squared in Figure 3.1.

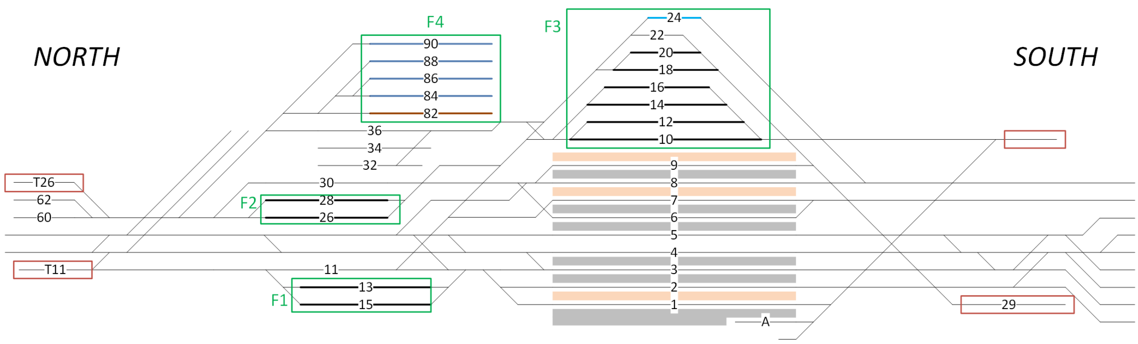


Figure 3.1: Simplified layout of Metz-Ville station

A timetable provides the set of trains that have to enter or leave the shunting yards. These trains are called *shunted trains*. The timetable also contains trains that simply go through the station. These trains are called *passing trains*. Each train departure requires a certain number of train units which belong to specific rolling stock types. A group of train units that arrive coupled to be shunted is an *arriving train*. Each departure is ensured by one or several train units stored in the shunting yard that can be coupled if necessary. This group of train units is called a *departing train*. We remark that a such group, for both arriving and departing trains, can include only one train unit if suitable. Maintenance and cleaning operations can be planned for arriving trains. These operations require specific tracks, as mentioned above, and specialized crews. These crews are available at specific periods of the day or the night. A rolling stock rotation can also be provided as an input for shunting problems. In France this input is mandatory and is a reference that shunting planners seek to stick to.

3.1.2 Characteristics of shunting problems

Shunting contains four basic problems which consider the management of a specific resource:

- *Train matching Problem (TMP)* in which train units are assigned to departing trains. Train units are the resource managed.
- *Track Assignment Problem (TAP)* in which train units are assigned to shunting tracks. The capacity of shunting tracks is the major constraint.
- *Shunting Maintenance Problem (SMP)* in which maintenance operations are scheduled. The resources needed are maintenance crews and shunting tracks.
- *Shunting Routing Problem (SRP)* in which shunting movements are set and scheduled. Infrastructure capacity in main tracks and shunting tracks is the main constraint.

These problems are illustrated with an example based on the Metz-Ville infrastructure shown in Figure 3.1 and the timetable in Tables 3.1 and 3.2. These tables contain a part of train units that have to be shunted at Metz-Ville during a week day. In Table 3.1, trains arriving at the station are listed. The time at which they arrive at their platform (column “Platform") is mentioned in column “Hour". A train has been parked in the shunting yard during the previous night, its Id is 01 and its hour is “D-1". As this

Table 3.1: Example of arrivals at Metz-Ville in daytime

Arrivals				Operations
Type	Id Number	Hour	Platform	
ZGC3	01	D-1	*	
ZGC3	23722	05:36	8	
ZGC4	23734	07:56	8	NS4,WC,LV,MAL
ZGC4	837511	08:21	4	NS4,WC,LV,MAL
ZGC3	834356	08:36	9	
Z2	834013	08:44	5	VAR,NS4,WC,LV
ZGC3	23738	08:46	8	VAR,NS4,WC,LV,MAL
TER2N	837519	08:57	6	ES,NS4
TER2N-TER2NL	88719	09:14	1	
TER2N-TER2N	88721	09:27	1	VAR,NS4,WC,LV,MAL
ZGC3-ZGC3	834358	12:36	8	VAR,NS4,WC,LV,MAL
ZGC3-ZGC3	88727	12:59	3	EAU
REG6	830304	13:15	6	EAU, MAL
Z2-Z2	834023	13:58	3	NS1
REG4	830306	14:15	8	EAU

Table 3.2: Example of departures at Metz-Ville in daytime

Departures			
Type	Id Number	Hour	Platform
ZGC3-ZGC3	834014	07:13	1
ZGC3	701005	10:20	7
TER2N	88730	10:33	9
Z2	837534	11:02	7
TER2N	88506	11:33	7
ZGC4	837536	11:39	2
TER2N	88734	11:42	3
TER2NL	88738	13:06	6
ZGC3	88740	14:46	1
ZGC3	837554	16:02	2
Z2	837556	16:09	1
ZGC4	23755	16:38	6
Z2	837560	16:39	1
REG4	830307	16:43	8
REG6	830315	17:25	7
ZGC3-ZGC3	837568	18:02	2
ZGC3	837536	18:26	A
TER2N	88766	18:33	2

train has been stored in the shunting yard, it did stop at a platform, therefore no platform is mentioned in Table 3.1. The configuration of each train is mentioned in column “Type". In this example, there are six types of train units:

- Regiolis B83500 with 4 carriages (REG4)
- Regiolis B83500 with 6 carriages (REG6)
- Z24500 with 3 carriages (TER2N)
- CFL 2200 with 3 carriages a version of Z24500 owned by Luxembourg National Railway Company (TER2N)
- Z11500 with 2 carriages (Z2)
- Z27500 with 3 carriages (ZGC3)
- Z27500 with 4 carriages (ZGC4)

If a train unit of type X and a train unit of type Y are coupled, then the configuration is denoted X-Y in Table 3.1 and 3.2. Some cleaning or maintenance operations are performed at the shunting yard:

- Arriving technical check (VAR)
- Engine inspection (ES)
- Water filling (EAU)
- Internal cleaning of type 1 (NS1)
- Internal cleaning of type 4 (NS4)
- Glass cleaning (LV)
- External cleaning with a washing machine (MAL)
- Toilet emptying (WC)

The operations that have to be performed on each arriving train are mentioned in column “Operations" in Table 3.1. The operations performed on each train are listed in the order in which they have to be done.

Train matching problem (TMP)

In the TMP, a main input to be considered is that a departure requires a specific configuration. Any train unit which belongs to a rolling stock type of the suitable configuration can ensure the departure. Train units must also arrive at the station sufficiently earlier than the departure time.

The configuration of departure trains is fixed as shown in Table 3.2. Local planners must respect these configurations when they match train units with departing trains.

Departures are ensured by train units which either belong to arriving trains or have been stored in the shunting yards before the beginning of the planning period. For sake of simplicity, latter can be considered as composing arriving trains which arrive at the beginning of the planning period. As configurations of arriving trains do not necessarily correspond to departing trains ones, train units must be split and joined. Splitting two train units leads to an *uncoupling operation* and joining them leads to a *coupling operation*. It is possible that, at the end of the planning period, train units remain stored in the shunting yard.

The number of train units required for departing trains is usually consistent with the number of train units in arriving trains, especially if a rolling stock rotation is given. The initial rolling stock rotation can be changed by local planners through the TMP in order to perform shunting operations which may have been forgotten or neglected in rolling stock circulation planning or to respond to perturbations. Indeed, the time interval between an arrival and a departure from a station must take into account several operations. First, once a train arrives at a platform, passengers get off, then an uncoupling operation can be performed at the platform, if appropriate. Second, the train is moved to a shunting yard. Once in the yard, coupling, uncoupling or maintenance operations may be carried out on the train. Third, the train built must be moved to its departure platform. In principle, it must arrive at the platform soon enough for boarding passengers and a possible additional coupling operation. Many of the durations mentioned

Table 3.3: Solution for matching arrivals to departures

Arrivals			Departures		
Type	Id Number	Hour	Type	Id Number	Hour
ZGC3	01	D-1	ZGC3-ZGC3	834014	07h13
ZGC3	23722	05h36			
ZGC4	23734	07h56	ZGC4	837536	11h39
ZGC4	837511	08h21	ZGC4	23755	16h38
ZGC3	834356	08h36	ZGC3	701005	10h20
Z2	834013	08h44	Z2	837534	11h02
ZGC3	834013	08h46	ZGC3	837536	18h26
2N	837519	08h57	2N	88766	18h33
2N-2N	88719	09h14	2N	88730	10h33
			2N	88738	13h06
2N-2N	88721	09h27	2N	88506	11h33
			2N	88734	11h42
ZGC3-ZGC3	834358	12h36	ZGC3-ZGC3	837568	18h02
ZGC3-ZGC3	88727	12h59	ZGC3	88740	14h46
			ZGC3	837554	16h02
REG6	830304	13h15	REG6	830315	17h25
Z2-Z2	834023	13h58	Z2	837560	16h39
			Z2	837556	16h09
REG4	830306	14h15	REG4	830307	16h43

above are fixed or can be estimated, so that we can obtain the pairs of arriving and departing trains which may be matched, i.e., with arrival and departure times separated at least of the total resulting interval. For these pairs the *matching duration condition* is respected.

Local planners match arriving and departing trains in order to minimize the number of coupling or uncoupling operations.

If an arriving and a departing train are matched and do not respect the matching duration condition, then the departure will be delayed or an optional operation will be canceled. A solution of the TMP is a *matching plan*.

Definition 3.1. Let a timetable with arriving and departing trains and minimum matching durations. The *Train Matching Problem (TMP)* is to match arriving train units to every departing train units with a minimum number of coupling or uncoupling operations and matching duration condition violations.

If an initial rolling stock rotation is given its deviation with the matching plan must be minimized.

For readability, we do not mention the minimum matching durations for the example in Tables 3.1 and 3.2. Table 3.3 presents the solution of the TMP proposed by planners. In this solution, the arriving trains 01 and 23722 are combined. The arriving trains 88717, 88721 88727 and 834023 are split. Moreover, planners did not choose to match the arriving train 88727 with the departing train 837568 so that the arriving train 834358, which arrives sooner at 12:36, gets split to ensure departing trains 88740 at 14:46 and 837554 at 16:02. Planners considered the duration of maintenance operations to set the matching duration condition. We remark in Table 3.1 that the arriving train 834358 has five operations planned while the arriving train 88727 has only one operation planned. The matching condition states that the arriving train 834358 can ensure a departure after 16:50 while the arriving train 88727 can ensure a departure after 14:10.

In a TMP solution, arriving trains can be split. The group of train units that stay coupled, possibly made of a single train unit, are *intermediate trains*. Then, these intermediate trains can be combined to obtain departing trains. In the solution in Table 3.3 train units of the arriving train 88719 scheduled at

09h14 is split into two intermediate trains.

Shunting maintenance problem (SMP)

Cleaning and maintenance operations are scheduled once arriving and departing trains are matched. Then, the train units that remain coupled in the shunting yard are known. These are intermediate trains. In the SMP, operations are carried out on intermediate trains. The earliest possible start time of all operations carried out on an intermediate train is the moment at which it arrives in the shunting yard. This takes into account the time the passengers need to alight the corresponding arriving train, to move to the yard and possibly to be split. The due date of all operations carried out on an intermediate train is the time at which it must leave the yard, it depends on the departure time of its departing train. Before the departure time from a platform, intermediate trains may also have to be combined, they must be moved to the platform and allow passengers to board. If the operations cannot finish at the due date, then the departure is delayed.

The order of the operations to be performed is already fixed. Nevertheless any operation can be canceled in order to avoid departure delays. Each operation has a fixed duration and must be executed without interruption. Therefore, no preemption is allowed. An operation needs two resources: a *crew* and a shunting track. A crew is specialized (internal cleaning, engine maintenance...) and can carry out several operations in its specialization field. The composition of a crew is supposed to be fixed, in particular we cannot enlarge a crew to reduce the operation duration. Crews have fixed shift times and can perform one operation at a time. Moreover, as technical equipment on a shunting track can be used for only one operation, two operations cannot be performed at the same time on a shunting track. If an intermediate train needs to change track between two operations, the duration of the shunting movement needed is taken into account. In particular, a *standard shunting movement duration* must separate the start time of an operation from the end time of the preceding one. The standard shunting movement duration can be set as the time of a long movement according to operations history. In the SMP, we suppose that there is a path between every shunting track which has cleaning or maintenance equipment.

Therefore, the SMP can be modelled based on disjunctive resources. It is a job shop scheduling problem whose jobs are the intermediate trains. The shunting tracks are machines in this problem. This job shop scheduling is *flexible* (Brucker and Schlie [1990]) since the shunting track where an operation is carried out can be chosen among a set of compatible shunting tracks. There is also a *human resource constraint*, since a crew must be available to perform an operation. Thus we consider two types of decisions:

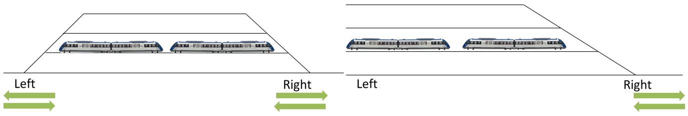
- a shunting track and a crew is assigned to each operation,
- operations are scheduled.

A solution of the SMP is a *maintenance schedule*.

Definition 3.2. The *Shunting Maintenance Problem (SMP)* is a flexible job shop scheduling with human resource constraints in which the total delay and the number of canceled operations is minimized.

Track assignment problem (TAP)

Parking train units in shunting yards is an issue if the available space on shunting tracks is scarce. This space must be used wisely, especially during the night when most of the train units have to be parked at shunting yards. The number of tracks is often lower than the number of train units parked at the same time. However, tracks are long enough for parking several train units. In many real life cases, tracks can contain up to 10 small train units which are 50 meters long. As several trains can be parked on a same track, planners must ensure that no train unit is blocked by another one placed in front of it if the former departs earlier than the latter. As mentioned in Section 3.1.2, there are shunting tracks in which the side on which train units enter or leave can be chosen. These are *regular tracks* (Figure 3.2a). Shunting track



(a) Regular tracks: train units can enter or leave at both sides

(b) LIFO tracks: train units can enter or leave at right side only

(c) FIFO tracks: train units can enter left side only and leave at right side only

Figure 3.2: Shunting tracks configurations

also contain dead-end tracks on which train units can enter or leave on one side (Figure 3.2b). We remark that the first train unit which enters in a such track is placed at the bottom of the track. Then, this first train unit is the last to leave the track. A dead-end track is a queue in which the last-in-first-out rule is respected. Dead-end tracks are called *LIFO tracks*. Tracks in which train units must enter on one side and leave on the other side are queues (Figure 3.2c). The first train unit which enters in a such track is the first one to leave. This is why such tracks are called *FIFO tracks*. FIFO tracks are not very common in railway, while LIFO tracks are usual. Regular tracks are also encountered in shunting yards. They are less restrictive and allow many parking combinations.

The management of parking space is different from usual infrastructure capacity management. Parking space on a track is a cumulative resource, contrary to a track circuit on main tracks which is a single-capacity resource. The use of parking space is based on two constraints. Firstly, the total length of train units which occupy a track at the same time cannot exceed the length of the track. This is the *length constraint*. Secondly, the order in which train units enter and leave the track must be such that every train unit can leave the track without moving another train unit placed in front of it. This is the *crossing constraint*. Indeed, it holds also for train units to be coupled to form departing trains, which must then be in coherent positions on the track and on the train. Some tracks can be dedicated to a type of rolling stock for technical reasons, such as electrification or gauge. If suitable, planners can decide that only train units of the same type can be parked on a same track. This can be helpful to deal with real-time disturbances. Indeed, if this is done, at any time a train unit of any type can leave the shunting yard without conflicts to ensure a rescheduled departure. Planners can also simplify the problem by splitting every track into positions with a fixed length. Then, each position is treated as a single-capacity resource. Doing so is much easier than dealing with length constraints, but there is a loss of parking capacity. This simplification may be particularly suitable if all the train units have the same length.

In the general case, the TAP must have a matching plan and a maintenance plan as input. Train units which belong to a same intermediate train are parked together, since they stay combined in the shunting yards. When a maintenance operation is carried out on an intermediate train, it must be parked on the track where the operation takes place. This is the *maintenance parking constraint*. Cleaning and maintenance operation crews must be protected from any train movement. Therefore, no train can enter or leave a track if a maintenance operation is in progress. This is the *protection constraint*. Moreover, coupling and uncoupling operations in the matching plan are performed on shunting tracks. We remind that maintenance operations are performed before uncoupling operations and after coupling operations on a same train unit. The track where coupling and uncoupling operations take place has to be decided by planners. Indeed, arriving trains are often split as soon as they arrive in the shunting yards. Moreover, once intermediate trains are coupled to become departing trains, they do not move to other shunting tracks. In between, intermediate trains may have to be parked at several shunting tracks. If an intermediate train needs to change track, the duration of the shunting movement needed is taken into account. A standard shunting movement duration must separate the arriving time at the destination track and the departing time at the origin track, similarly to what described for the SMP. Moreover, a path must exist between shunting tracks that are successively assigned to a train unit. A matching plan is essential to define a TAP since intermediate trains that stay in the shunting yards must be provided. Instead, a maintenance schedule is only useful for setting protection and maintenance parking constraints.

If train units can be parked on at most k tracks during their stay in shunting yards, we have a k *Multiple Track Assignment Problem (TAP^k)*. However, in the general case the number of tracks where a train unit has to be parked is not bounded: it is a *Multiple Track Assignment Problem (TAP*)*. A solution of TAP* is a *multiple track assignment*. This problem includes *matching plan constraints*:

- Train units which belong to a same arriving train must be parked at the same track just after they arrive in the shunting yards.
- Train units which belong to a same departing train must be parked at the same track just before they leave the shunting yards.
- A train unit must be parked on a track long enough to ensure drivers relief and coupling or uncoupling operations. That is why a minimum parking time is specified.

Definition 3.3. Considering a matching plan and a maintenance schedule, the *Multiple Track Assignment Problem (TAP*)* is to assign shunting tracks with entrance side and time, and leaving side and time, to intermediate trains such that crossing, length, matching plan, maintenance parking and protection constraints are respected, and the number of used tracks and shunting movements is minimized.

Another version of the TAP considers that intermediate trains are parked at only one track during the planning period: This is the 1-track assignment problem (TAP^1) also called *standard track assignment problem*. Considering this problem is relevant if train units do not have to change tracks for cleaning or maintenance operations. In the TAP^1, the time at which each train unit arrives in a shunting track is fixed since it is the time at which it arrives in the shunting yards. Similarly the time at which each train unit leaves a shunting track is known since it is the time at which it leaves the shunting yards. Therefore, there is no use to assign an entrance and a leaving time to each intermediate train that has to be parked in the shunting yard.

Definition 3.4. Considering a matching plan and a maintenance schedule, the *Standard Track Assignment Problem (TAP^1)* is to assign a parking track with entrance and leaving side to intermediate trains compatible with their maintenance operations, such that crossing and length constraints are respected and the number used tracks is minimized.

Shunting routing problem (SRP)

Shunting movement demands are the first input of the SRP. They can be deduced from a parking plan and a maintenance schedule. Arriving trains need to move from their platform to their first shunting track. Also, departing trains need to move from their last shunting track to their platform. Intermediate trains can have to move between shunting tracks. Hence, a shunting movement demand can be related to either an arriving, a departing or an intermediate train. The side from which trains must enter or leave a platform or a shunting track is provided by the parking plan. A shunting movement demand for a train indicates the window time in which it must start moving and a time window in which it must arrive at its destination. At shunting tracks, these time windows are set such that the maintenance schedule and the parking plan remain feasible.

The second input of the SRP is a station infrastructure. An infrastructure provides a set of possible paths, these paths may include turnarounds. Paths can be described with a microscopic approach such that each track circuit is considered. Other scales which aggregate track circuits may also be studied. Without loss of generality we focus on a track circuit scale representation. Signaling system characteristics are also precised. Minimum running times of different rolling stock types through parts of the infrastructure are given. Therefore, each shunting movement demand has a set of possible paths according to its origin and destination.

The SRP must also consider the timetable at main tracks as a third input. Since some shunting movements go through main tracks, they may affects passing trains. Passing trains have a path through

Table 3.4: Decision makers for shunting problems

Shunting problem	Resources	Decision maker at tactical level
Train Matching Problem (TMP)	Rolling stock	Ground operations planner
Shunting Maintenance Problem (SMP)	Crews, Shunting tracks	Maintenance and cleaning operations manager
Track Assignment Problem (TAP)	Shunting tracks	Ground operations planner
Shunting Routing Problem (SRP)	Station infrastructure	Station traffic scheduler

the station and schedules. These schedules indicate the time at which a passing train arrives at a waypoint along its path.

In the SRP, a path and schedules must be assigned to shunting movement demands. The schedules give the time at which the head of the train concerned by the shunting movement arrives at each track circuit. These schedules must be consistent with feasible running times. Schedules also have to respect the time needed for a turnaround, if a turnaround is planned. These two conditions are the *succession constraint*. With these schedules and signaling system features, the occupation time of each track circuit can be deduced. Then, the moment at which signals turn to "clear" is known. Every signal must be "clear" when crossed by a train, this is the *capacity constraint*. Shunting movements which end at a shunting track must arrive in the time window indicated in the demand, this is the *time window constraint*. If a shunting movement for a departing train arrives at its platform later than its time window, then there is a delay penalty.

Definition 3.5. Considering an infrastructure, a timetable and shunting movements demands, the *Shunting Routing Problem* is to assign a path and a schedule to each shunting movement demand such that succession, time window and capacity constraints are respected and delays penalties are minimized.

3.1.3 Integrating shunting problems

Shunting problems are usually solved separately because they involve different resources and decision makers (Table 3.4). Nevertheless, they can have significant dependencies, as Lentink [2006] underlines. Specifically, we can notice a partial hierarchy between shunting problems. Firstly, the main purpose of all these problems is the preparation of rolling stock to ensure on-time departures: a matching plan is essential to service. Secondly, the SMP, the TAP* and the SRP require a matching plan to be instantiated while the TMP can be instantiated and solved without solutions of other problems. Therefore, the TMP is a master problem. Moreover, the SRP is instantiated with shunting movement demands which are set with a parking plan and a maintenance schedule. Thus the SRP is a slave problem of the TAP* and the SMP. In addition, the TAP* needs a maintenance schedule. Then, the SMP is a master problem of the TAP*. The particular case of the TAP^1 can be instantiated and solved without any maintenance schedule, to solve the SMP with maintenance parking constraints and protection constraints.

With this partial hierarchy in mind, we can build the sequential planning approach showed in Figure 3.3. Here, some specific constraints are represented with white filled rectangles. Thick arrows represent the shunting problems solutions flow while thin arrows represent the information flow of data. If the destination of a thick arrow is a color filled rectangle, then the destination problem needs the corresponding input solution to be instantiated. If the destination of a thick arrow is a specific constraint, it is this constraint which directly depends on the corresponding input solution. Therefore, these arrows represent the hierarchical relation between problems. Indeed, if an arrow links two color filled rectangles, the origin is a master problem of the destination.

Indeed, the interaction between shunting problems must provide consistency. This also leads to questions about optimality. A slave problem includes constraints defined by a solution of the master problem, and we would like to ensure that a solution of the master problem makes the slave problem

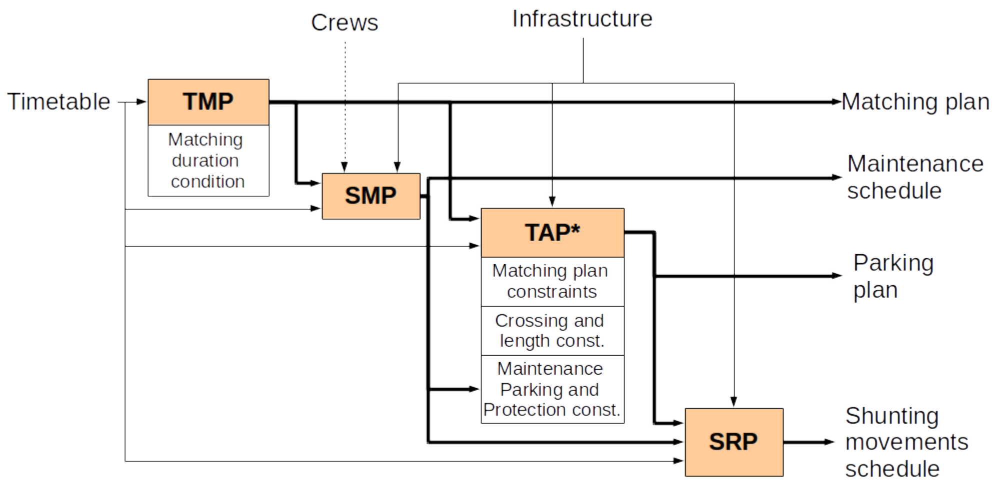


Figure 3.3: Sequential planning approach for shunting problems (const.: constraint(s))

instance at least feasible. This is why the master problem includes *aggregated constraints*, in which the detail of the resources used by the slave problems is not detailed.

The SMP is a slave problem of the TMP, because a matching plan defines the jobs instantiated in the former. In particular, the matching plan gives the time at which train units must be ready for departure. Hence, the TMP has aggregated constraints guaranteeing that maintenance and cleaning operations can be performed, with the minimum departure delays. This requirement is included in the matching duration condition.

The TAP is also a slave problem of the TMP, since a matching plan sets the intermediate trains which have to be stored in the shunting yards. The matching plan also specifies the arrival and departure time of each intermediate train. Indeed, coupling and uncoupling operations provided in a matching plan define matching plan constraints that the TAP must satisfy. TAP and TMP can have conflicting objective functions. In the TAP, we need to minimize the number of tracks used by train units. The more arriving trains are split, the smaller intermediate trains are. Then a more convenient way to fit these intermediate trains in fewer tracks can be found. However, the more arriving trains are split, the more coupling and uncoupling operations are needed, which is detrimental for the TMP.

The SMP and the TAP are linked by maintenance parking constraints and protection constraints. Train units must stay long enough at shunting tracks with cleaning or maintenance equipment so that operations can be performed. Because of protection constraint, performing maintenance operations on-time may spread train units on more shunting tracks than if such operations are canceled. Therefore more shunting tracks may be needed to fulfill the objective of the SMP, which is minimizing the total delay and the number of canceled operations. Thus, the objectives of SMP and TAP can be contradictory. Besides, as shunting tracks are assigned to operations in the SMP, a partial parking plan is already set in the maintenance schedule. A such parking plan is partial because the track assignment is known during maintenance only, while the TAP* deals with the whole time spent in the shunting yard. This partial parking plan is feasible for the TAP*. Indeed, length, protection and crossing constraints are respected because shunting tracks are treated as single-capacity resources in the SMP (one intermediate train per track). Matching plan constraints are also satisfied because in the SMP cleaning and maintenance operations can start once are uncoupling operations performed. Moreover, by definition, the SMP also considers a standard time needed to move between tracks.

Moving on, the SRP is a slave problem of the TAP: it is instantiated with a parking plan. Moreover, the TAP includes an objective function and constraints that, to some extent, facilitate the feasibility of the SRP. Firstly, in the TAP the number of shunting movements is minimized. This may prevent traffic conflicts due to a huge amount of simultaneous movements. Secondly, the TAP ensures that there exists a path between the shunting tracks assigned successively to an intermediate train. Thirdly, a standard shunting movement duration is taken into account.

The SRP is indirectly a slave problem of the TMP, through the mediation of the TAP. Thanks to the TAP, the intermediate trains that have to be moved are determined. The TMP takes into account the time needed for shunting movements with the matching duration condition.

The SRP and the SMP are also linked since time window constraints can be based on a maintenance schedule. Moreover, the SMP considers the time needed for shunting movements. If maintenance operations that have to be performed on a same intermediate train can be done on a same track, we only consider movements between platforms and shunting tracks. In this case, the SRP may have to include washing machine paths which might delay departures.

Remark that, in some cases the TAP^1 is to be solved, i.e., when train units do not need to move between shunting tracks in the planning period. This is the problem tackled for example by Lentink [2006]. In these cases, the TAP can be solved prior to the SMP, simply shifting to this latter problem the solution of maintenance parking and protection constraints.

3.2 Literature review

3.2.1 Classification of related works

The optimization approaches proposed in the literature either deal with one shunting problem in passenger transportation or combine a few. Literature also contains scheduling problems similar to shunting ones. We also review works that deals with the integration of shunting problems in rolling stock circulation planning as well as integrated problems in railway planning. In Figure 3.4, we propose a classification of these problems.

A first part of shunting literature tackles problems related to the TMP. In these problems, we assign vehicles (trains, buses, trams) to departures. Two main variants can be distinguished. In the first one, which is called *vehicle dispatching problem (VDP)*, vehicles are placed in lanes or tracks at a depot during the night. The assignment is set so that no crossing conflict occurs in the morning. The second variant is specific to trains and is called *train matching problem*. No parking issue is taken into account but train units can be uncoupled or coupled.

A second part of the literature deals with scheduling problem close to the SRP.

The third part of the literature tackles the TAP or similar problems with other vehicles (trams, buses) that have to be stored in a yard. Many contributions deal with the TAP^1 version (Definition 3.4), in which a vehicle matching is fixed. Then, arriving and departing time of each vehicle in the yard is known. TAP^1 often considers vehicles that are stored in a depot during the night or off-peak hours. In this case, every vehicle arrives before all departures, this is the *midnight condition*. If the midnight condition is respected, all train units occupy the depot at a same time. Other contributions combine the TAP with the TMP or the VDP in case of buses or trams. Such problems are called *vehicle positioning problems (VPP)*. Vehicles must be parked in a depot and assigned to a departure so that no crossing conflict occur. The *Train Unit Shunting Problem (TUSP)* is the VPP for railway transportation. The TUSP takes into account vehicle coupling and uncoupling. Some contributions also tackle both the SMP and the TAP: they deal with depots in which maintenance is performed. The maintenance schedule can be fixed or to decide. In the first case, we only have to chose tracks where maintenance operations occur, while in the second one the whole SRP must be solved. During maintenance operations, vehicles must be parked at locations with specific equipment. Moreover, during the studied period, a vehicle can be parked at several different locations to release critical equipment.

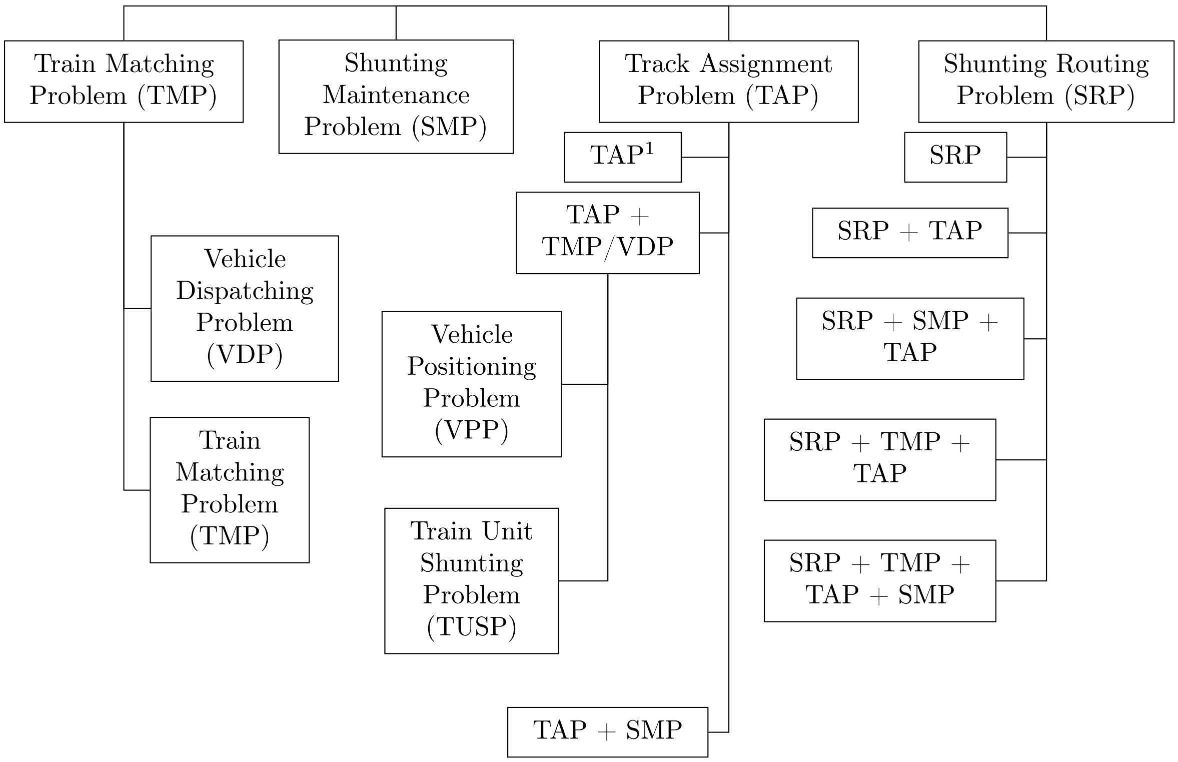


Figure 3.4: Passengers transportation shunting problems in the literature

The fourth part of the literature deals with the SRP and is specific to railway transportation. In the problems considered, shunting movements must be performed in order to change trains locations. Shunting movements can be taken into account with a more or less detailed model. The SRP can be solved separately or it can be merged with the TAP. The TMP and the SMP can also be added to a problem that models and solves the SRP and the TAP.

3.2.2 Literature review on the Train Matching Problem

Vehicle dispatching problem (VDP)

Blasum et al. [1999] introduce the vehicle dispatching problem (VDP) for trams. The authors consider trams that have been stored in a depot during the night. This depot contains LIFO tracks and the position of each tram is known. Therefore the track where a tram is parked is specified as well as its position on this track. These trams must be assigned to departures in order to avoid crossings. Departure times are an input. Each departure requires a specific type of tram, a dispatching solution assigns to it a tram of the desired type. Vehicle dispatching problem is of major interest at the operational level when it is necessary to modify the planning after an unforeseen event such as vehicle breakdown. Blasum et al. [1999] note that this decision problem is NP-complete thanks to a reduction from the 3 dimensional matching. The authors propose a dynamic programming search, in which the state of the depot is represented after each departure. They implement elimination techniques to cope with memory issues due to states representation. They solve in less than 1000 seconds most of a set of random instances which contain up to 150 trams of 46 types and 10 tracks. They also solve real word instances which contain up to 74 trams in less than 10 seconds.

Winter and Zimmermann [2000] consider two other approaches for solving the optimisation problem associated to the VDP. The number of crossings is minimized in a first version of the problem, while in a second version the number of departures performed with a tram of the desired type is maximized. An exact mathematical programming solution and a reactive tabu search initialized with a greedy algorithm are used on real and random instances. The reactive tabu search provides satisfying solutions in few seconds while the mathematical program solved with CPLEX can require several hours.

Eggermont et al. [2009] provide additional complexity results for the VDP. The authors prove that problems remains NP-complete despite a restriction on the number of vehicles and vehicle types a track contains. Nevertheless, the VDP can be solved in polynomial time if there are at most two vehicles of each type. Lübbecke and Zimmermann [2005] tackle a similar problem for freight transportation.

Train matching problem (TMP)

As by Definition 3.1, the TMP is linked to the rolling stock circulation problem in which train composition are planned (Fioole et al. [2006]; Alfieri et al. [2006]). However, in the TMP, there is no need to match train composition and demand: the given relation between departing trains and necessary train units takes care of this match. In the TMP, the focus is on train units coupling and uncoupling. Indeed in the TMP, the composition of arriving and departing trains in a yard is fixed. Arriving trains must be matched to departing trains in order to minimize the number of uncoupling operations. Lentink et al. [2006] describe the TMP thanks to the possible configurations which can be obtained after splitting arriving trains. These configurations are equivalent to paths in an oriented graph modeling of arriving trains.

Lentink [2006] proves that the TMP is NP-hard with a reduction from the 3 partition problem. The author also tackles cases in which the TMP is solvable in polynomial time and space. If the number of train units per arriving or departing train is at most three, then the TMP is polynomially solvable. The same holds if the number of train units per arriving or departing train is fixed as well as the number of train units stored in the yard.

Haahr and Bull [2015] propose an extension of the TMP which includes maintenance requirements. The set of operations that have to be performed on a train can depend of the matching plan. The authors prove that this problem is NP-hard and solve it with a column generation approach.

3.2.3 Literature review on the Shunting Maintenance Problem

In this section, we review contributions which tackle scheduling problems similar to the SMP. These contributions belong to the category of machine scheduling problems with availability constraints.

Lentink [2006] tackles cleaning operations scheduling. A parking plan is given, but human resources have to be assigned to these operations. The processing time of an operation depends of the number of crews assigned to perform it. The number of crews available is time dependant. This problems is equivalent to single machine scheduling problem at each shunting track. It is considered with a discrete time model and formulated with an integer linear program.

Flexible job-shop scheduling integrates two sub-problems: a *routing* in which a machine is chosen for each operation and a *scheduling*. It has been solved, for example, with genetic algorithms (Zhang et al. [2011]), local search (Dauzère-Pérès and Paulli [1997], Mastrolilli and Gambardella [2000]) and linear programming (Cemal et al. [2010]).

The SMP is also a multi-resource scheduling problem. Dauzère-Pérès et al. [1998] propose a local search algorithm based on disjunctive graphs. Guyon et al. [2014] specifically tackle human resource constraints issues. They have to assign operators to shifts and machines so that a feasible job-shop schedule can be found. An integer linear programming (ILP) formulation with a discrete time model is introduced by the authors. Artigues et al. [2009] solves this problem by combining ILP and constraint programming.

3.2.4 Literature review on the Track Assignment Problem

Combinatorial problems related to the Track Assignment Problem

The TAP suffers from major combinatorial issues due to both length and crossing constraints. Indeed, the TAP is linked to packing problems and require some graph theory notions to be formally handled.

First, we can notice that length constraints in the TAP^1 with midnight constraints bring to a bin-packing problem. Indeed as all the train arrivals occur before departures, there is a moment in which all train units must be parked in the shunting yard.

Definition 3.6 (Bin Packing Problem (Coffman Jr et al. [1996]))**.** Given a set of items $I = \{1, 2, .., n\}$ with respective sizes $\{\omega_1, \omega_2, .., \omega_n\}$ and an unlimited number of bins of capacity W, the bin-packing problem is to partition I into a minimum number subsets in which the total size of the items does not exceed W.

The problem is NP-hard and is widely tackled in literature. Martello and Toth [1990] propose the obvious lower bound $\left\lceil \frac{\sum_{i \in I} \omega_i}{W} \right\rceil$. More relevant lower bounds can be obtained with an accurate sorting of the items.

Crossing constraints in TAP are related to the position of train units on tracks. The same problem sometimes occurs in freight yards. In this context, the so-called *train sorting problem* pushes and pulls railcars on different tracks so as to achieve the desired sorting minimizing either the number of tracks used to store railcars or the number of movements. A first variant of the train sorting problem supposes that a railcar can be moved at most twice: after an uncoupling operation and before a coupling operation. This is the single stage variant (Dahlhaus et al. [2000]). The other variant of train sorting does not limit the number of movements and is called multistage (Jacob et al. [2011]). Gatto et al. [2009] provide a literature review which highlights algorithmic issues in shunting for freight transportation and reports some relations with passengers transportation.

Dahlhaus et al. [2000] and Jacob et al. [2011] remark that storing railcars in tracks is equivalent to providing a partition of a permutation. This holds for TAP^1 with midnight condition since a set of train units which can be described as a permutation is partitioned into subset of trains that are parked on a same track (Gatto et al. [2009], Di Stefano and Koči [2004]). The permutation represents the order in which train units enter and leave the shunting track. Let us consider a set of n train units. Every train unit is numbered according to its departure order, so that the first train unit which leaves the yard is labeled 1 and the last one is labeled n. For example, with $n = 4$, in $[3, 1, 4, 2]$ the first train which arrives is the third one to leave while the third train which arrives is the fourth one to leave. The decision space of TAP^1 is included in the set of partitions of a given permutation. Crossing constraints are respected if each partition respects monotony properties imposed by the type of track. For example, in the case of LIFO tracks each partition must include train units labeled in decreasing order (indeed the last train unit which arrives must be the first one to leave). Feasible partitions in this case are named decreasing sequences. For permutation $[3, 1, 4, 2]$, the partition $\{[3, 1], [4, 2]\}$ is feasible.

Bodlaender et al. [1995] remark that the problem of partitioning a permutation into increasing (or decreasing) sequences can be computed in polynomial time.

Standard Track Assignment Problem (TAP^1)

Di Stefano and Koči [2004] tackle crossing constraints in the TAP^1. They suppose that the length of parking tracks is at least equal to the sum of the vehicles length, therefore they do not have to deal with length constraints. Moreover, they consider that the midnight condition is respected. Then, arrival and departure order of trains can be described as permutation sequences. The authors try to minimize the number of tracks used according to four configurations. The first configuration is the one in which n trains enter and leave a track on only one side. This configuration gathers LIFO and FIFO cases. Then, the TAP with these conditions is equivalent to permutation graph coloring. For this case, the paper provides optimal solution with an algorithm which requires $O(n \ln n)$ time. In a second configuration, trains enter

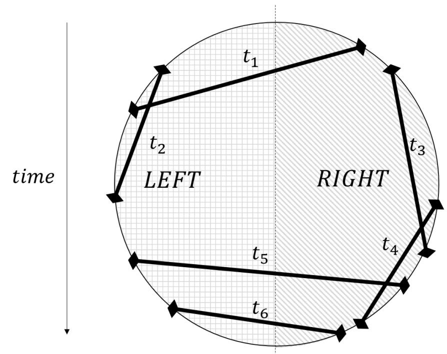


Figure 3.5: Graphical representation of crossing constraints on a regular track.Trains are represented by two points which indicate the side and the time at which they enter and leave the track. For each train, theses points are linked with a segment. If two segments cross, the corresponding trains can not be parked at the same track using the specific entrance and exit sides at the specified times.

a track through one side and can leave the track through both sides. Then the problem can be modeled as a 3-uniform hypergraph coloring. The authors find an upperbound for the number of tracks needed, whatever the order of trains. A polynomial algorithm which provides a solution below this bound is proposed. The third configuration tackled is the one in which trains can enter a track through both sides and leave the track through one side. The authors remark that this configuration can be solved with the algorithm used for the second configuration. The fourth configuration is the one in which there are regular tracks. In this case, the problem is equivalent to a 4-uniform hypergraph coloring. Moreover, the authors study the problem with FIFO or LIFO tracks and without the midnight condition. In this case, crossing constraints are handled thanks to a graphical representation that has been extended to regular tracks by Kroon et al. [2008]. Figure 3.5 shows an example of this representation for a specific track. A circle represents the track for a fixed time horizon. The circle is divided in two halves, indicating a side each. Time grows from top to bottom. Arrival and departure time of trains using the track are points on this circle. Trains are represented by segments which link their arrival and the departing points. If two segments cross, then the crossing constrains are violated. In Figure 3.5, the train t_1 enters on the right side before t_2 enters on the left side. Then, t_1 leaves on the left side before t_2 leaves on the same side. The segments representing t_1 and t_2 cross, then the crossing constraints are not respected for these trains. Indeed, the sequence of events is infeasible.

Cornelsen and Di Stefano [2007] deal with a version of the TAP[1], in which regular tracks have infinite lengths. The side on which a train can enter or leave a track is fixed. The authors provide complexity results based on graph coloring depending on whether the timetable is cyclic or acyclic.

Demange et al. [2012] tackle an online version of the TAP[1]. The authors consider LIFO, FIFO and regular tracks in which the number of trains that can be parked is limited. They consider several algorithms based on bounded graph coloring.

Gilg et al. [2018] deal with LIFO, FIFO, and regular tracks in the TAP[1]. They also consider length constraint. Trains can have different lengths. They propose two formulations based on mathematical programming. In a first formulation, arrival times are included in crossing constraints, while a second one simply represents conflicts. They note that the second formulation provides a better linear relaxation. A robust extension and a stochastic version are proposed to take into account possible delays.

Vehicle Positioning Problem (VPP)

Winter and Zimmermann [2000] introduce the VPP for the management of tram depots. Trams of different types have to be stored in LIFO tracks to ensure departures in the morning. The authors suppose that all trams have the same length. Then tracks are divided into positions which can contain a tram. Moreover, the midnight condition is respected. In a first version of the VPP, the number of crossings is minimized, while in a second version the number of departures performed with a tram of the desired type is maximized. The contribution notes that the VPP is NP-complete thanks to a reduction from the 3 dimensional matching problem. The authors propose mathematical programming formulations and three heuristics.

Winter [1999] also tackles an online version of the VPP, in which the midnight condition is not respected. The author considers trams with different lengths.

Gallo and Miele [2001] extend the work of Winter and Zimmermann [2000] to the management of bus depots. Buses, which can have different lengths, are stored in FIFO lanes. The contribution considers schedules in which the midnight condition is not respected. In this version of the VPP, the number of departures ensured by a bus is maximized. The authors tackle the problem with a two level approach. In the first level, they solve vehicle matching with a relaxation of crossing constraints. In the second level, they deal with these crossing constraints. In the first level, a Lagrangian relaxation is used to generate an upper bound. In a second level, they provide a feasible solution for the VPP thanks to a heuristic in which crossings are minimized. The solution approach is implemented for solving real instances of the Florence bus network in Italy. These instances, which contains up to 12 lanes and 77 buses of 4 different types, are also solved with CPLEX. Then, up to 13000 variables and 4 millions constraints are generated. CPLEX encounters many memory issues or is unable to provide a feasible solution within 10800 seconds for most of the instances. On the contrary, the two level approach gives satisfying solutions in up to 300 seconds.

Also Hamdouni et al. [2006] consider a version of the VPP applied to buses: lanes are divided into blocks and the midnight condition is respected. They seek for robust solutions in which buses of the same type are parked together. They handle the problem well with CLPEX. Hamdouni et al. [2007] propose then a Benders decomposition for the problem.

Cardonha and Borndörfer [2009] tackle the VPP for a depot with FIFO lanes with the same hypotheses as Winter and Zimmermann [2000]. They also consider a cyclic version. The number of crossings must be minimized in a 3 dimensional assignment problem. They propose three quadratic programming formulations. The authors propose a pseudo-polynomial time pricing based on the dynamic programming approach of Winter and Zimmermann [2000] for a column generation algorithm. Then, an integer solution is obtained thanks to a heuristic. They also deal with the robust configurations introduced by Hamdouni et al. [2006]. They consider fictive instances which contain up to 20 tracks and 200 vehicles of 40 types. They obtain integer solutions in few seconds and note significant gaps for some instances.

Train Unit Shunting Problem (TUSP)

The TUSP is an adaptation of the VPP to trains. It takes into accounts train units coupling and uncoupling. This problem is introduced by Freling et al. [2005]:

Definition 3.7 (Train unit shunting problem (TUSP) (Freling et al. [2005]))**.** Given a station layout and a timetable, the *train unit shunting problem* consists of (i) matching the arriving train units and departing train units and (ii) parking these on a shunting track.

Freling et al. [2005] solve the TMP in a first phase and the TAP in a second one, once the train matching is set. An ILP based on a graph associated to each train models the TMP. The TAP is tackled with a column generation approach which is based on dynamic programming as in Cardonha and Borndörfer [2009]. The authors solve 12 real-life instances based on Zwolle station data in the Netherlands, which contain about 80 train units, 15 regular tracks and 4 LIFO tracks. The TMP step is solved in less than 2 seconds while for busy day scenarios the TAP step can last up to 40 minutes without proof of optimality.

Haijema et al. [2006] solve the TUSP with a heuristic that tackles the problem split as function of sub-planning periods. Computation times do not exceed one second on Zwolle station instances.

Kroon et al. [2008] propose an ILP formulation which integrates TAP and TMP. The new issue is that the position of train units must be taken into account so that there is no crossing after uncoupling.

The authors also include valid inequalities based on the conflict graph cliques. The formulation is solved with CPLEX and tested on Zwolle and Enschede stations instances, in the Netherlands. Some solutions are obtained in few seconds while other instances can not be solved to optimality within 3 hours.

Haahr et al. [2017] provide a column generation method with LIFO tracks and allow trains to be parked at platforms. The pricing approach is based on a shortest path search. This method is compared with constraint programming formulations, a greedy algorithm, the one-step mathematical program proposed by Kroon et al. [2008] and the two-step approach by Freling et al. [2005].

Integrating the Track Assignment Problem and the Shunting Maintenance Problem

Usually, when the TAP and the SMP are combined, a train must be parked successively on different tracks to use various equipment necessary for its maintenance.

Jacobsen and Pisinger [2011] consider the TAP* and the SMP with a discrete time model. Shunting tracks are supposed to be LIFO. The paper uses three metaheuristics: guided fast local search, guided local search and simulated annealing. Experiments are run on virtual instances containing up to 10 trains. Guided local search provides results close to exact solutions found through an ILP on test instances. Solutions are obtained after a few seconds by guided local search while it takes several hours with CPLEX. The authors consider the use of simulated annealing to improve solutions obtained by local search.

A paper considers the TAP* and the SMP with a fixed maintenance schedule. Li et al. [2017] assume that arrival and departure time on shunting tracks are known and consider an homogeneous train fleet. In this case, the TAP* can be reduced to the TAP^1. Tracks can be LIFO or regular and are set to contain at most two train units. A mathematical program which models crossing conflicts and maximizes the number of parked trains is proposed.

3.2.5 Literature review on the Shunting Routing Problem

Shunting Routing Problem (SRP)

Zwaneveld et al. [1996] tackle the problem of routing trains through stations. Trains have an expected schedule on which variations can be applied. Incompatibilities between assignments are modeled using a conflict graph. This leads to the search for a maximum stable set in this graph. A branch-and-cut algorithm is proposed to solve instances based on the Zwolle station infrastructure.

Riezebos and Van Wezel [2009] study combinatorial aspects of shunting movements. They provide a two-step algorithm which searches for the shortest paths traversing a set of shunting tracks.

Lentink et al. [2006] consider a version for the SRP in which running times are fixed. The authors propose a heuristic based on the A* algorithm.

Van Den Broek and Kroon [2007] focus on the SRP. The parking tracks are fixed for every train while shunting movements between tracks have to be scheduled. They propose a first MILP formulation in which shunting routes are fixed. They also study a variant in which alternative shunting routes can be chosen. In both variants, a maximum number shunting movements must be scheduled in their planned time window. The MILP models are solved with CPLEX on three instances based on three Dutch stations. In most cases, the computation times do not exceed 5 minutes. The authors find that the multiplicity of alternative routes strongly increases computation time.

Combining the Shunting Routing Problem and the Track assignment Problem

Abbink [2006] tackles both TAP and SRP with a discrete time model. Tracks are divided in positions that can contain a train unit. A constraint programming algorithm is proposed to achieve conflict-free planning.

Combining the the Shunting Routing Problem, the Track assignment Problem and the Shunting Maintenance Problem

The contributions that combine TAP, SMP and SRP consider that each track can contain at most one train. Thus, they do not deal with crossing and length constraints. Shunting tracks are used as single-capacity resources: we have a *single-capacity track assignment problem* denoted TAP-SING.

Tomii and Zhou [2000] tackle SMP and TAP*-SING. Trains have to be shunted to appropriate track to be subject to operations. The authors consider the choice of shunting routes between sidings and the limited size of maintenance or driving crews. Scheduling is performed thanks to a PERT (Program Evaluation and Review Technique) network and resource assignments are changed with a genetic algorithm. Qi et al. [2017] propose a discrete time model which integrates SRP, TAP*-SING and SMP. The model is solved with a Lagrangian relaxation-based algorithm on Beijing South Railway Station instances. Ramond [2008] deals with a similar problem. Shunting tracks are also single-capacity resources. Scheduling and resources assignment are performed with local search.

Combining four shunting problems

Ramond and Nicolas [2014] propose an extension of the TUSP to incorporate the SMP. The maintenance operations that are carried out on train units can depend of the matching plan, though. The authors also integrate the SRP with a macroscopic model for traffic conflicts. This problem has been tackled by 36 teams in ROADEF/EURO challenge 2014 (Artigues et al. [2018]). Most of the teams solved the problem with two or three phases decomposition approaches and discrete time models.

Van Wezel and Riezebos [2010] introduce a software library for planning shunting operations based on an object oriented description. The authors study a prototype which integrates planning algorithms and an interactive user interface.

Van den Broek [2016] extends the TUSP to integrate both the SRP and the SMP. In this extension, the TAP* sub-problem is considered since trains may be parked at several shunting tracks. In the SRP, simultaneous movements are not treated and the duration of shunting movements is fixed. Therefore, microscopic scheduling is not needed. The author models the problem with a directed graph in which nodes correspond to activities such as parking, movement or cleaning, and arcs represent precedence constraints. Other arcs are set for single-capacity resources used for movements and cleaning. The problem is solved thanks to local search and simulated annealing. Local search operators are based on neighborhoods specific to each sub-problem. Suitable solutions are found in five minutes to artificial and real-world instances based on the Hague rolling stock maintenance plant, in the Netherlands. This plant contains 15 shunting tracks which are mostly regular tracks. The authors investigate then the exact position of trains that are parked on a track. It is a new issue in the TAP* specific to regular tracks. Indeed for LIFO tracks, train units are parked at the rear-end of the track. A dynamic programming approach is proposed.

3.2.6 Shunting issues in rolling stock circulation problem

Haahr and Lusby [2017] propose a model of the rolling stock circulation problem (RSCP) which integrates TAP decisions and constraints. A multi-commodity flow based model is combined with a track assignment at each station depot. They propose a compact ILP formulation and a path approach for column generation. They tackle real life instances based on Copenhagen commuter rail network. The results highlight the interest of integrating TUSP and RCSP.

Thorlacius et al. [2015] integrate shunting track assignment in the RSCP. Train units that are coupled can be parked on the same track. Then, position of train units must be taken into account. The authors add edges which represent parking in a multi-commodity network. Instances based on Copenhagen commuter rail network are solved with a hill-climbing heuristic.

Cadarso and Marín [2011] integrate some the TAP constraints in RCSP. The authors consider an aggregated capacity constraint for the TAP. They also take into account the cost of coupling and uncou-

Table 3.5: Summary of main contributions in shunting for passengers trains. The second column indicates if a papers tackle the TMP. The third column indicates the version of TAP considered. The fourth and the fifth columns respectively show if SMP and SRP are treated. In both columns, Cont. is for a continuous time model and Disc. for a discrete time model. The last column indicates if the paper proposes an integrated solution approach.

Contribution	TMP	TAP	SMP	SRP	Integrated
Tomii and Zhou [2000]	No	TAP^*-SING	Cont.	Cont.	Yes
Freling et al. [2005]	Yes	TAP^1	No	No	No
Abbink [2006]	No	TAP^1	No	Disc.	Yes
Haijema et al. [2006]	Yes	TAP^1	No	No	No
Kroon et al. [2008]	Yes	TAP^1	No	No	Yes
Lentink et al. [2006]	Yes	TAP^1	No	Cont.	No
Jacobsen and Pisinger [2011]	No	TAP^*	Disc.	No	Yes
Van den Broek [2016]	Yes	TAP^*	Cont.	Cont.	Yes
Haahr et al. [2017]	Yes	TAP^1	No	No	Yes
Li et al. [2017]	No	TAP^1	No	No	-
Qi et al. [2017]	No	TAP^*-SING	Disc.	Disc.	Yes
Gilg et al. [2018]	No	TAP^1	No	No	-
Chapter 4	Yes	TAP^*	Cont.	Cont.	Yes
Chapter 5	Yes	TAP^*	Cont.	Cont.	No

pling.

3.2.7 Other integrated problems in railway planning

As many shunting problems are often solved in an integrated way, the litterature also includes important contributions on the integration of other railway planning problems. In particular, Schöbel [2017] highlights optimality and consistency problems of sequential planning approaches in public transportation. In this section, we focus on solution approaches used for integrating rolling stock management and infrastructure capacity management. Schöbel [2017], Benhizia [2012] propose detailed surveys on this topic.

Schöbel [2017] tackles line planning, timetabling and vehicle scheduling problem (rolling stock circulation problem for railway transportation). She proposes an algorithmic framework for integrating the three aforementioned problems. This framework is based on iterations in which a single problem is solved, where the solution of the specific problem has to be consistent with possible solutions of the other problems found in previous steps.

Cadarso and Marín [2012] consider a discrete time model in order to integrate timetabling and rolling stock circulation problems for suburban services. They consider a MILP formulation which leads to better solutions than a rolling stock circulation model with fixed timetable. Liebchen and Möhring [2007] show that a periodic event based scheduling framework can be used to integrate cyclic timetabling and rolling stock management.

3.3 Conclusion

In this chapter we proposed formal presentations of four shunting problems which manage rolling stock, cleaning and maintenance crews, and infrastructure capacity. We specified constraints which highlight the dependencies between these problems.

In the analysis of state of the art, we brought out modeling approaches and algorithmic issues of each shunting problem. We also reviewed contributions which integrate some of the shunting problems thanks to sequential or integrated approaches. Table 3.5 reports a list of the main contributions. In this thesis, we aim to provide models and algorithms which integrate the four problems introduced in this chapter.

In Chapter 4, we will introduce a generalization of the TUSP called G-TUSP, in which train units can be parked successively on several shunting tracks. We will propose a MILP formulation for modeling and solving it. In Chapter 5 we will investigate sequential algorithms for solving the G-TUSP.

Remark that the approaches proposed in this thesis are the only ones dealing with all four problems considering TAP^* rather than TAP^1, and solving a precise routing and scheduling problem for the SRP.

Chapter 4

MILP formulation for integrating maintenance and routing

As discussed in Chapter 3, rolling stock planning must manage *train units* between an arriving trip and a departure trip in a station. This specific part of rolling-stock management is called *shunting.* Inside stations, train units are prepared for departure and possibly stored for several hours if they are not needed immediately. More precisely, they are cleaned and have maintenance checks. Moreover, train units can be coupled or uncoupled to match train configurations required for departure. This is done on shunting tracks located around platform tracks. Parallel shunting tracks form shunting *yards.* Some of these shunting tracks have specific amenities such as train-wash for external cleaning or pits for maintenance checks. To be stored in shunting yards, train units need first of all to be moved from their arrival platform. Then, they can possibly need to be moved there from one yard to another. Finally they need to be moved to their departure platform. Movements arriving or departing from a yard are called *shunting movements* and must respect traffic safety rules imposed by signalling system and by ground-agents instructions. Indeed, shunting movements must not create conflicts with the rest of train traffic in the station.

Shunting operations planning includes several decisions. First, arriving train units must be assigned to departures, which constitutes a matching decision. This matching must take into account rolling stock features required for departures. Another decision concerns train units location: they must be parked at one or several shunting tracks depending on amenities required by maintenance operations. Similarly, movements are set to achieve the parking locations. For these movements, route planning decisions are to be made, since paths are assigned to train units and movements are scheduled based on running times and potential conflicts. Finally, depending on maintenance crews availability, maintenance operations must be scheduled. Although all these decisions are often taken separately, they are usually strongly interdependent. For instance, some matching plans make train units parking or maintenance scheduling impossible.

The Generalized Train Unit Shunting problem (G-TUSP) is the problem of shunting operations planning. It integrates four sub-problems:

- The Train Matching Problem (TMP), the problem of matching arriving and departing train units.
- The Track Allocation Problem (TAP*), the problem of choosing train units location.
- The Shunting Routing Problem (SRP), the problem of determining train units routing during shunting movement.
- The Shunting Maintenance Problem (SMP), the problem of defining train units maintenance scheduling.

The G-TUSP considers a station and a timetable with arriving and departing trains that need to be shunted. It is a pre-operational problem, it is solved from 6 days to 4 hours before operations. The problem aims to minimize departure delays and cancellations if timetable perturbations are expected, as well as maintenance call off. Moreover, the minimization of the number of coupling and uncoupling operations is also sought.

The aim of this chapter is to provide a formal model of the G-TUSP. Specifically, the contribution consists in formulating it as a mixed-integer linear programming (MILP) formulation. The formulation is based on a microscopic representation of the infrastructure and on consideration of dummy train units in order to manage coupling and uncoupling. In particular, an accurate description of traffic conflicts due to shunting movements is proposed: shunting movements can be simultaneous and capacity utilization is wisely considered thanks to a model inspired by RECIFE-MILP Pellegrini et al. [2015]. This formulation has made the object of a paper published in the proceedings of the international conference RailNorrköping 2019 - 8th International Conference on Railway Operations Modelling and Analysis (ICROMA) Kamenga et al. [2019b] and has been presented in the 20th congress of the French Operations Research & Decision Support Society Kamenga et al. [2019c].

The rest of the chapter is organized as follows. The GTUSP is formally introduced in Section 4.1. Section 4.2 proposes a MILP formulation of the G-TUSP together with some proposals for making it stronger. Section 4.3 reports possible preprocessings for speeding up the solution of the formulation. Section 4.4 describes the experiments carried out as proof of concept of the applicability of the formulation. Finally, Section 4.5 concludes the chapter.

4.1 Formal description of G-TUSP

4.1.1 Modelling principles

In our formulation of the G-TUSP, we consider that train units can be coupled or uncoupled to form trains. Three formal sets of trains are introduced to model this: arriving, intermediate and departing trains. *Arriving trains* are moved from a platform track to the shunting yard. Once there, they are uncoupled if needed, and they become *intermediate trains*, which are moved in the yard and submitted to maintenance. Finally, intermediate trains are coupled if necessary and become *departing trains* to be moved to the suitable platform track. We suppose that arriving trains become intermediate trains at the first shunting track they use. Also once intermediate trains are parked at their last shunting track they become departing trains. Trains move on an infrastructure modeled microscopically through a *track-circuit* scale representation. A track-circuit is a portion of track on which the presence of a train unit is automatically detected. Thanks to this infrastructure model, detailed characteristics of *interlocking systems* are taken into account and train safety is ensured through suitable separation.

Figure 4.1 represents a simple example in which an orange, a green and a blue path are shown with their respective track-circuits named z followed by a number. Both the orange and blue paths use track-circuit z_{15}, therefore they cannot utilize it at the same time. The train with the orange path is an intermediate train whose path starts at shunting track 21. This train results from the arriving train using the green path and has to be cleaned. It is parked at shunting track 29 for cleaning. The train with the blue path is a departing train which uses platform A.

Some specific concepts with notations are introduced in Section 4.1.2, 4.1.3 and 4.1.4. Notations are reported in Table 4.1.

4.1.2 Trains

We denote T_T the set of arriving trains. Each arriving train can be split into several intermediate trains. For an arriving train t', $T_I(t')$ is the set of its possible intermediate trains. The set of departing trains is denoted T_S. For a departing train we denote $T_I(t)$ the set of intermediate trains which are compatible

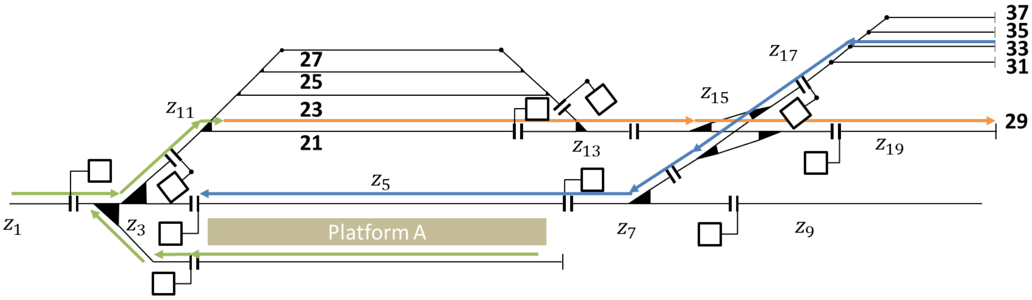


Figure 4.1: Simple example. Station layout with signals represented by squares. The green arriving train whose path is represented with a green line becomes the orange intermediate train at shunting track 21. The orange intermediate train's path is represented in orange. The blue departing train leaves the shunting track and is moved to platform A. This train uses the blue path.

with t. Those are intermediate trains which can be coupled to obtain t. In this definition intermediate trains in $T_I(t)$ must arrive before t's departure.

Remark 4.1. *An arriving train $t' \in T_T$ which contains k train units has $|T_I(t')| = \frac{(k+1)k}{2}$ possible intermediate trains.*

Proof. For an arriving train, we consider the sequence of its train units $\{1, ..., k\}$. Each intermediate train corresponds to a subsequence of consecutive terms. Then the number of intermediate trains is the number of these subsequences. There are $k - i + 1$ subsequences of consecutive terms which starts with element in position i. Then, there are $\sum_{i=1}^{k} k - i + 1 = \frac{(k+1)k}{2}$ subsequences of consecutive terms. □

Every train is composed of one or several train units. Train units are divided into types so that same type train units are interchangeable. Every arriving train entering the shunting yard disappears, and one or more intermediate trains appear. All intermediate trains do not disappear to become departing trains. Some intermediate trains may remain in the shunting yard at the end of the planning period. For trains that are stored in the station before the planning period, a dummy arriving train is introduced. This arriving train enters the station at the beginning of the planning period on the associated siding.

Besides, by definition, the sets $T_I(t')$ are disjoints. For readability, we introduce $T_I = \cup_{t' \in T_T} T_I(t')$ that is the set of intermediate trains. We can remark that a departing train t and an arriving train t' use the same set of train unit if and only if $T_I(t) = T_I(t')$. In Figure 4.2, three types of train units are considered: hashed ones, full colored ones and white ones. For each arriving train, the set of its intermediate trains is represented by a dark-blue lined dashed box. For each departing train, the set of compatible intermediate trains is represented with a light-green lined dashed box. Arrows represent a possible combination of coupling and uncoupling to use the train units available to compose the two departing trains. Here, the arriving train t_1 is uncoupled in order to obtain train t_A and two intermediate trains are coupled to obtain train t_B.

We also consider trains which stop or pass at the station without being shunted. Those are *passing trains*. The set of passing trains is denoted T_P.

4.1.3 Infrastructure

A *track-circuit scale* model is used in order to get a precise capacity occupation representation. In the station area, a train follows a path which is a track-circuits succession. As trains can turn around, a path may go twice through a track-circuit. Therefore, we introduce *formal track-circuits* to specify passing directions. For every real track-circuit, we consider a set of corresponding formal track-circuits. Each of these sets contain up to two formal track-circuits, since there is a formal track-circuit per direction.

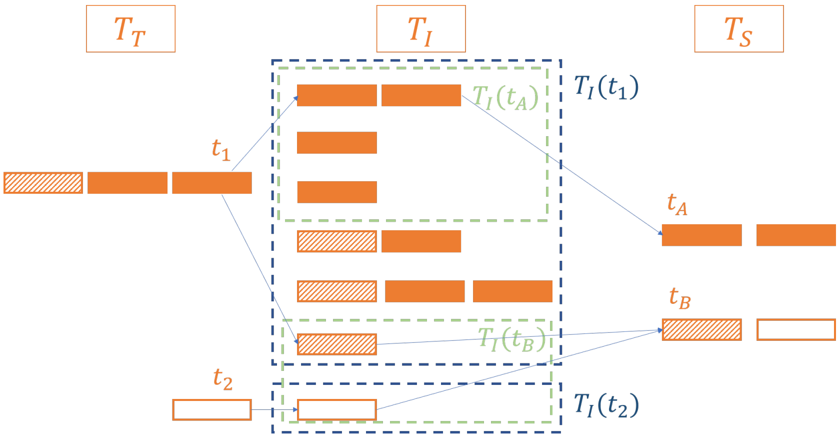


Figure 4.2: Train matching. Arriving trains T_T on the left are used for the departing trains T_S on the right thanks to intermediate trains T_I. A possible matching is represented with arrows.

We distinguish the notion of *path* from that of *route*. Routes are individually handled and defined by signalling control. A route starts at a starting signal and finishes at a finishing signal. A path is a concatenation of routes and may include turnarounds. In the turnarounds, a first route is defined up to the turnaround place where a second route starts. In a turnaround a train stops at a signal and then follows another signal at its back which leads to a reverse direction.

Capacity occupation is based on track-circuit reservation. When a train t needs to go through a track-circuit tc, the signal which allows t to move into the *block section* where tc is located must have a green aspect. A block section is a sequence of track-circuits which can be utilized by at most one train at a time. Thanks to the interlocking system, the green aspect can be obtained once the itinerary r that leads t to tc is formed. This is why we introduce *formation times*, which depend on block sections characteristics. However r can only be formed if all conflicting block sections are released. A block section locked by a train is released shortly after this train clears the last track-circuit it is using in the block section itself.

A path can imply parking on a shunting track. Paths are set such that shunting tracks are at the beginning or the end of the path. For a path r, we define Ps^r the set of shunting tracks where r starts and Pe^r the set of shunting tracks where r ends. Ps^r and Pe^r can contain one shunting track or be empty. For example, if Ps_r is empty, then path r does not start on a shunting track. Every train has a set of usable paths. Arriving trains' paths begin on a platform and terminate on a shunting track, while departing trains' paths begin on a shunting track and terminate on a platform. Exactly one path among usable paths is assigned to arriving, departing and passing trains. Intermediate trains' paths begin and terminate on shunting tracks. When an intermediate train needs to be parked successively at several tracks, a sequence of paths is assigned to it. Two fictive paths r_0 and r_∞ are assigned to intermediate trains. r_0 is at beginning of the sequence while r_∞ terminates it.

The set of exit points of a shunting track p is denoted $Ex(p)$. This set contains at most two elements which indicate a geographical location. We use the locations left and right, respectively denoted L and R. For a train which enters shunting track p with path r, we introduce the boolean indicator $Es(r, p)$ which is equal to 1 if the exit point where r ends in p is left and 0 if this exit point is right. Similarly, for a train which leaves shunting track p with path r, we introduce the boolean indicator $Ee(r, p) \in Ex(p)$ which is equal to 1 if the exit point where r begins in p is left and 0 if it is right. A path r_2 can only follow a path r_1 if r_1 ends at the shunting track where r_2 begins: $Ps^{r_2} \cap Pe^{r_1} \neq \emptyset$. In the example of Figure 4.1, the green path is denoted r_1 and the orange one is denoted r_2. r_2 follows r_1 at shunting track 21. Indeed $Ps^{r_2} = Pe^{r_1} = \{21\}$.

4.1.4 Maintenance operations

Cleaning or maintenance operations may be included in the rolling-stock plan. They are considered to be made on intermediate trains. The operations to be carried out on an intermediate train $t \in T_I$ form set O_t. An operation $o \in O_t$ can only be performed on shunting tracks with specific facilities. The sequence of operations is given. We introduce P^o set of shunting tracks where o can be carried out. In addition, an operation requires the use of specific human resources. We consider that an operation o requires a crew among the set HR^o of crews which can be assigned to o. Each crew is available from its shift start time to its shift end time.

We also note that when an operation is in progress, the shunting track where it is carried out must be protected to ensure staff safety. Thus, during this period, no other train can enter this shunting track or leave it.

4.1.5 Definition of the G-TUSP

Definition 4.2. Given the infrastructure of a station, a set of arriving, departing, intermediate and passing trains and a set of maintenance operations, the *Generalized Train Unit Shunting Problem (G-TUSP)* is defined by the following decisions, objective function to minimize and constraints.

- Decisions:
  - Matching arriving and intermediate trains, matching intermediate and departing trains;
  - Assigning a path that goes from a platform to a shunting track to each arriving trains, a path that goes from a shunting track to a platform to each departing train, a succession of paths between shunting tracks to each intermediate train, a path to each passing train;
  - Scheduling each shunting movement (i.e. determining the time at which trains pass at each track-circuit of a path);
  - Assigning a crew and a shunting track to each maintenance operation;
  - Scheduling each maintenance operation.
- Objective function: cost for departure cancellation, departure delays, assignments of intermediate trains to departing trains, coupling and uncoupling operations, maintenance operation cancellation, duration and number of shunting movements.
- Constraints:
  - Length constraint: the length of trains that occupy a shunting track at the same time can not exceed the length of the track;
  - Crossing constraint: when a train leaves a shunting track, it can not be blocked by another one;
  - Precedence constraint for maintenance operations and movements defined by sequences of formal track-circuits in a path and sequences of paths for an intermediate train;
  - No traffic conflict occurs, i.e trains can not use a same track-circuit at the same time during movements;
  - Maintenance parking constraint: when an operation is carried out on an intermediate train at a given shunting track, it must be parked at a this track;
  - Protection constraint: no train can enter or leave a track when a maintenance operation is in process;
  - Resource disjunction for maintenance operations: maintenance operations can not use a same crew or track at a same time.

Table 4.1: Notations for G-TUSP

Notation	**Description**
T_T,T_I,T_S,T_P	set of arriving trains, intermediate trains, departing trains, passing trains
$T = T_T \cup T_I \cup T_S \cup T_P$	set of trains
$T^* = T_T \cup T_I \cup T_S$	set of shunted trains
$T_I(t)$	set of intermediate trains compatible with the arriving or departing train $t \in T_T \cup T_S$
TU,$m_{t,tu}$	set of train unit types, number of train units of type $tu \in TU$ in train $t \in T^*$
index t	index of train $t \in T$
ty_t,l_t, d_t, ex_t	type of train $t \in T$, length of train $t \in T$, arrival time of train $t \in T_T \cup T_P$ or departure time of train $t \in T_S \cup T_P$, time at which $t \in T_S \cup T_P$ leaves the infrastructure
B_t,Q_t	cancellation cost of train $t \in T_S$, cost associated to the delay of train $t \in T_S \cup T_P$
A_t,Q_R	cost of one time unit duration of a shunting movement performed on intermediate train $t \in T_I$ and cost of the assignment of a route to intermediate train $t \in T_I$
$\omega S_{t,t'}$	weight associated to the assignment of intermediate train $t' \in T_I(t)$ to departing train $t \in T_S$
Q_C,Q_H	coupling cost, uncoupling cost
b_{t_1,t_2}	indicator function: is equal to 1 if $t_1 \in T_I(t)$ (with $t \in T_T$) is placed to the left of $t_2 \in T_I(t)$ with index $t_1 <$ index t_2 on their corresponding arriving train t and 0 otherwise
a_{t_1,t_2}	indicator function: is equal to 1 if $t_1 \in T_I(t)$ (with $t \in T_S$) must be placed to the left of $t_2 \in T_I(t)$ with index $t_1 <$ index t_2 and $ty_{t_1} \neq ty_{t_2}$ if they are assigned to the same departing train t and 0 otherwise
mp	minimum parking time
R_t,TC_t,Z_t	set of paths, formal track-circuits and real track-circuits which can be used by a train $t \in T$
$TC(z)$	set of formal track-circuits corresponding real track-circuit $z \in \cup_{t \in T} Z_t$
Z^r, TC^r	set of real and formal track-circuits for path $r \in \cup_{t \in T} R_t$
M_R	maximum number of paths which can be assigned to an intermediate train
$OTC_{ty,r,tc}$	set of consecutive formal track-circuits preceding $tc \in TC^r$ which are occupied by a train of type ty traveling along path $r \in \cup_{t \in T} R_t$ if its head is on tc, depending on train and track-circuit length
$pc_{r,t}$, $sc_{r,t}$	formal track-circuits preceding and following $tc \in TC^r$ along path $r \in R_t$
$rt_{ty,r,tc}$, $ct_{ty,r,tc}$	running and clearing time of $tc \in TC^r$ along $r \in \cup_{t \in T} R_t$ for a train of type ty
$ref_{r,tc}$	reference formal track-circuit for reservation of $tc \in TC^r$ along $r \in \cup_{t \in T} R_t$(depending on the interlocking system)
$bs_{r,tc}$	block section including formal track-circuit $tc \in TC^r$ along $r \in \cup_{t \in T} R_t$
for_{bs},rel_{bs}	formation and release time for block section bs

Ps^r,Pe^r, P^r	set of shunting tracks where $r \in \cup_{t\in T} R_t$ begins, set of shunting tracks where $r \in \cup_{t\in T} R_t$ ends, set of shunting tracks in r $P^r = Ps^r \cup Pe^r$. In particular Pe^{r_0} and Ps^{r_∞} contain all shunting tracks, while $Ps^{r_0} = Pe^{r_\infty} = \emptyset$
$Z(p)$,$Ex(p)$	set of real track-circuits and set of exit points composing a shunting track p
L_p	length of shunting track p
$tcp_{r,p}, tce_{r,p}$	reference formal track-circuit for parking at shunting track $p \in P^r$ along $r \in \cup_{t\in T} R_t$, first formal track after shunting track $p \in Ps^r$ along $r \in \cup_{t\in T} R_t$
$Es(r,p)$	boolean indicator equal to 1 if the exit point where $r \in \cup_{t\in T} R_t \setminus \{r_0\}$ ends in shunting track $p \in Pe^r$ is left and 0 if it is right
$Ee(r,p)$	boolean indicator equal to 1 if the exit point where $r \in \cup_{t\in T} R_t \setminus \{r_\infty\}$ begins in shunting track $p \in Ps^r$ is left and 0 if it is right
O_t	set of operations to carry out on $t \in T_I$
E_t	set of pairs of successive operations on $t \in T_I$. $(o, o') \in E_t$ if and only if operation o' follows operation o
pR^o,ω_o	duration and cancellation cost of operation $o \in \bigcup_{t\in T_I} O_t$
HR^o,P^o	set of crews and shunting tracks which can be assigned to operation $o \in \bigcup_{t\in T_I} O_t$
sR_{hr},eR_{hr}	shift start time and shift end time of crew hr
M,τ_M	large constant compared to event times, end of planning period

4.2 MILP formulation

4.2.1 Basic formulation

In the formulation, we introduce non-negative continuous variables:

- $oc_{t,r,tc}$, $\phi_{t,r,tc}$, $sU_{t,r,tc}$, $eU_{t,r,tc}$, with $t \in T, r \in R_t, tc \in TC^r$: time at which t starts the occupation of tc along r, additional running time of t on tc along r compared to a free-network situation, time at which tc starts being utilized by t along r, time at wich tc ends being utilized by t along r;

- $sO_{o,r,r',p,hr}$, with $t \in T_I, o \in O_t, r, r' \in R_t, p \in P^r \cap P^o, hr \in HR^o$: time at which o starts at shunting track p between paths r and r' with crew hr;

- D_t, with $t \in T_S \cup T_P$: delay suffered by train t when exiting the control area.

Moreover, we introduce binary variables:

- xT_t, with $t \in T_I$, is equal to 1 if t is created and 0 otherwise;

- $xS_{t,t'}$, with $t \in T_S, t' \in T_I(t)$, is equal to 1 if t' is assigned to t and 0 otherwise;

- $xR_{t,r}$, with $t \in T$, $r \in R_t$, is equal to 1 if t uses r and 0 otherwise;

- $xO_{o,r,r',p,hr}$, with $t \in T_I, o \in O_t, r, r' \in R_t, p \in Pe^r \cap Ps^{r'} \cap P^o, hr \in HR^o$, is equal to 1 if o is carried out at shunting track p between paths r and r' with crew hr and 0 otherwise;

- qS_t, with $t \in T_S$, is equal to 1 if t is cancelled and 0 otherwise;

- $yR_{t,t',r,r',tc,tc'}$ with $t, t' \in T$, $r \in R_t$, $r' \in R_{t'}$, $z \in Z^r \cap Z^{r'}$, $tc \in TC(z) \cap TC^r$, $tc' \in TC(z) \cap TC^{r'}$, index $t <$ index t', is equal to 1 if t uses tc along r before t' uses tc' along r' and 0 otherwise;

- $k_{t,r,r',p}$, with $t \in T_I, r, r' \in R_t, (p \in Ps^{r'} \cap Pe^r \neq \emptyset)$ (i.e. r' can follow r), is equal to 1 if t uses r followed by r' (also it is equal to 1 if t is parked at track p) and 0 otherwise;
- $y_{o,o',hr}$ with $t, t' \in T$, $o \in O_t$, $o' \in O_{t'}$, $hr \in HR^o \cap HR^{o'}$, index $t <$ index t', is equal to 1 if hr performs o before o' and 0 otherwise;
- $yO_{o,o',p}$ with $t, t' \in T$, $o \in O_t$, $o' \in O_{t'}$, $p \in P^o \cap P^{o'}$, index $t <$ index t', is equal to 1 if p is used by o before o' and 0 otherwise;
- $ysO_{o,t,r,p}$, with $t \in T_I$, $t' \in T_I$, $t \neq t'$, $o \in O_{t'}$, $r \in R_t$, $p \in P^o \cap Pe^{r'}$, is equal to 1 if o is carried out at shunting track p before t enters p through r and 0 otherwise;
- $yeO_{o,t,r,p}$, with $t \in T_I$, $t' \in T_I$, $t \neq t'$, $o \in O_{t'}$, $r \in R_t$, $p \in P^o \cap Ps^{r'}$, is equal to 1 if o is carried out at shunting track p before t leaves p through r and 0 otherwise;

We also introduce the following integer variables:

- u_t, with $t \in T_T$ gives the number of uncoupling operations on t;
- v_t, with $t \in T_S$ gives the number of coupling operations on t.

The objective function to minimize integrates several penalties (4.1). First, it takes into account the cost of departure cancellations and delays. The function includes uncoupling and coupling operations cost. Then, penalties for intermediate trains assignment to departing trains are added, to penalize the changes to the original plan. Moreover, maintenance operations cancellation costs are introduced. We note that we can have a penalty only if the intermediate train concerned by the operation is actually created. Finally, we minimize the number of shunting movements for an intermediate train and the duration of these movements.

$$\begin{aligned} \min \sum_{t \in T_S} B_t q S_t + \sum_{t \in T_S \cup T_P} Q_t D_t + \sum_{t \in T_T} Q_C u_t + \sum_{t \in T_S} Q_H v_t + \\ \sum_{t \in T_S} \sum_{t' \in T_I(t)} \omega S_{t,t'} x S_{t,t'} + \sum_{t \in T_I, o \in O_t} \omega_o \left(xT_t - \sum_{\substack{p \in P^o \cap Pe^r \cap Ps^{r'} \\ r,r' \in R_t, hr \in HR^o}} xO_{o,r,r',p,hr} \right) + \\ \sum_{t \in T_I} \sum_{r \in R_t, p \in Ps^r} Q_R x R_{t,r} + A_t(oc_{t,r,tc_\infty} - oc_{t,r,tce_{r,p}}) \end{aligned} \tag{4.1}$$

Matching constraints

The MILP formulation must consider TMP constraints. First, we need to check train compositions. We introduce constraints for the number of train units of a specific type in trains. For each type, each arriving train must have the same number of train units as intermediate trains created after uncoupling (4.2). Also, each departing train must have the same number of train units as the intermediate trains assigned to it for coupling (4.3). As intermediate trains cannot be split, each of them can be assigned at most to one departing train. If the intermediate train is not created, it cannot be assigned to a departing train (4.4). A departure train is cancelled if no intermediate train is assigned to it (4.5). Then, the number of uncoupling operations on an arriving train or coupling operations on a departing train is equal to the number of intermediates trains assigned minus one (4.6), (4.7).

$$m_{t,tu} = \sum_{t' \in T_I(t)} m_{t',tu} xT_{t'} \quad \forall t \in T_T, tu \in TU \tag{4.2}$$

$$m_{t,tu}(1 - qS_t) = \sum_{t' \in T_I(t)} m_{t',tu} xS_{t,t'} \quad \forall t \in T_S, tu \in TU \tag{4.3}$$

$$\sum_{t' \in T_S : t \in T_I(t')} xS_{t',t} \leq xT_t \quad \forall t \in T_I \tag{4.4}$$

$$1 - qS_t \geq xS_{t,t'} \quad \forall t \in T_S, t' \in T_I(t) \tag{4.5}$$

$$u_t \geq \sum_{t' \in T_I(t)} xT_{t'} - 1 \quad \forall t \in T_T \tag{4.6}$$

$$v_t \geq \sum_{t' \in T_I(t)} xS_{t,t'} - 1 \quad \forall t \in T_S \tag{4.7}$$

Routing constraints

A first set of constraints is based on the RECIFE-MILP model of Pellegrini et al. [2015]. An arriving or a passing train cannot be operated before its arrival time (4.8). The start time of track-circuit occupation by a train along a path is zero if the path itself is not used (4.9). A train starts occupying a track-circuit along a path after spending in the preceding track-circuit its running time and possibly an additional running time due to delay or dwelling, if the path is used (4.10). An arriving or a passing train uses exactly one path (4.11).

A second set of constraints is specific to the SRP in G-TUSP. A departing train uses exactly one path if it is created and zero otherwise (4.12). An intermediate train uses at most M_R paths if it is created and zero otherwise (4.13). If an intermediate train is created, it uses the dummy paths r_0 (4.14) and r_∞ (4.15).

$$oc_{t,r,tc} \geq d_t xR_{t,r} \quad \forall t \in T_T \cup T_P, r \in R_t, tc \in TC^r \tag{4.8}$$

$$oc_{t,r,tc} \leq M xR_{t,r} \quad \forall t \in T, r \in R_t, tc \in TC^r \tag{4.9}$$

$$oc_{t,r,tc} = oc_{t,r,pc_{r,tc}} + \phi_{t,r,pc_{r,tc}} + rt_{t,r,pc_{r,tc}} xR_{t,r} \quad \forall t \in T, r \in R_t, tc \in TC^r \tag{4.10}$$

$$\sum_{r \in R_t} xR_{t,r} = 1 \quad \forall t \in T_T \cup T_P \tag{4.11}$$

$$\sum_{r \in R_t} xR_{t,r} = 1 - qS_t \quad \forall t \in T_S \tag{4.12}$$

$$\sum_{r \in R_t} xR_{t,r} \leq M_R xT_t \quad \forall t \in T_I \tag{4.13}$$

$$xR_{t,r_0} = xT_t \quad \forall t \in T_I \tag{4.14}$$

$$xR_{t,r_\infty} = xT_t \quad \forall t \in T_I \tag{4.15}$$

Two constraints model the sequence of paths used by an intermediate train. If a path is used by an intermediate train:

- exactly one path follows it (4.16),

- exactly one path precedes it (4.17).

$$\sum_{r' \in R_t : Ps^{r'} = \{p\}} k_{t,r,r',p} = xR_{t,r} \quad \forall t \in T_I, r \in R_t, p \in Pe^r \tag{4.16}$$

$$\sum_{r' \in R_t : Pe^{r'} = \{p\}} k_{t,r',r,p} = xR_{t,r} \quad \forall t \in T_I, r \in R_t, p \in Ps^r \tag{4.17}$$

A train delay is at least equal to the difference between its actual exit time from the infrastructure and its scheduled scheduled one (4.18).

$$D_t \geq \sum_{r \in R_t} oc_{t,r,tc_\infty} - d_t \quad \forall t \in T_S \cup T_P \tag{4.18}$$

The formulation includes constraints that take into account train matching decisions and the sequence of paths used by intermediate trains. These constraints consider two trains t and t' which use the same rolling-stock. A minimum parking time must be ensured between t's arrival (at the end of t's path) and t''s departure on the shunting track. It happens when an arriving train t becomes an intermediate train t' (4.19), when an intermediate train uses two paths in a row (4.20) and when an intermediate train becomes a departing train (4.21).

$$\begin{aligned} oc_{t',r',tce_{r',p}} \geq & \sum_{r \in R_t : p \in Pe^r} [oc_{t,r,pc_{r,tc_\infty}} + (rt_{t,r,pc_{r,tc_\infty}} + mp)xR_{t,r}] \\ & -M(1 - k_{t,r_0,r',p}) \quad \forall t' \in T_T, t \in T_I(t), r' \in R_{t'}, p \in Ps^{r'} \end{aligned} \tag{4.19}$$

$$\begin{aligned} oc_{t,r',tce_{r',p}} \geq oc_{t,r,pc_{r,tc_\infty}} + rt_{t,r,pc_{r,tc_\infty}} + mp - M(1 - k_{t,r,r',p}) \\ \forall t \in T_I, r, r' \in R_t : p \in Pe^r \cap Ps^{r'} \end{aligned} \tag{4.20}$$

$$\begin{aligned} \sum_{r' \in R_{t'} : p \in Ps^{r'}} oc_{t',r',tce_{r',p}} \geq oc_{t,r,pc_{r,tc_\infty}} + (rt_{t,r,pc_{r,tc_\infty}} + mp)xR_{t,r} \\ -M(1 - xS_{t',t}) \quad \forall t' \in T_S, t \in T_I(t'), r \in R_t, p \in Pe^r \end{aligned} \tag{4.21}$$

Moreover, we need to ensure spatial coherence. It means that when an arriving train t becomes an intermediate train t', t uses a path which ends at the same shunting track where the path used by t' starts (4.22), (4.23). If the intermediate train t' is not created ($xT_{t'} = 0$), as $M_R \geq \sum_{r \in R_t} xR_{t,r}$, variables $k_{t',r_0,r,p}$ are not constrained. The same happens when an intermediate train t' becomes a departing train t (4.24), (4.25). If the intermediate train t' is not assigned to the departing train t ($xS_{t,t'} = 0$), as $M_R \geq \sum_{r \in R_t} xR_{t,r}$, variables $k_{t',r_\infty,r,p}$ are not constrained.

$$\begin{aligned} \sum_{r \in R_t : p \in Pe^r} xR_{t,r} \leq \sum_{r \in R_{t'} : p \in Ps^r} k_{t',r_0,r,p} + M_R(1 - xT_{t'}) \\ \forall t \in T_T, t' \in T_I(t), p \in \bigcup_{r \in R_t, r' \in R_{t'}} (Pe^r \cup Ps^{r'}) \end{aligned} \tag{4.22}$$

$$\begin{aligned} \sum_{r \in R_{t'} : p \in Ps^r} k_{t',r_0,r,p} \leq \sum_{r \in R_t : p \in Pe^r} xR_{t,r} + M_R(1 - xT_{t'}) \\ \forall t \in T_T, t' \in T_I(t), p \in \bigcup_{r \in R_t, r' \in R_{t'}} (Pe^r \cup Ps^{r'}) \end{aligned} \tag{4.23}$$

$$\sum_{r\in R_{t'}:p\in Pe^r} k_{t',r,r_\infty,p} \leq \sum_{r\in R_t:p\in Ps^r} xR_{t,r} + M_R(1-xS_{t,t'}) \quad \forall t\in T_S, t'\in T_I(t), p\in \bigcup_{r\in R_t, r'\in R_{t'}} (Ps^r \cup Pe^{r'}) \tag{4.24}$$

$$\sum_{r\in R_t:p\in Ps^r} xR_{t,r} \leq \sum_{r\in R_{t'}:p\in Pe^r} k_{t',r,r_\infty,p} + M_R(1-xS_{t,t'}) \quad \forall t\in T_S, t'\in T_I(t), p\in \bigcup_{r\in R_t, r'\in R_{t'}} (Ps^r \cup Pe^{r'}) \tag{4.25}$$

Track-circuits of shunting tracks must remain in use when a train is parked there. Thus, when an arriving train t becomes an intermediate train t', t' starts using the first track-circuit of its path before t finishes using the last track-circuit of its path (4.26). The same happens when an intermediate train uses two paths in a row (4.27) and when an intermediate train becomes a departing train (4.28).

$$sU_{t',r',sc_{r',tc_0}} \leq eU_{t,r,pc_{r,tc_\infty}} - M(2-k_{t',r_0,r',p} - xR_{t,r}) \quad \forall t\in T_T, t'\in T_I(t), r\in R_t, r'\in R_{t'}, p\in Pe^r\cap Ps^{r'} \tag{4.26}$$

$$sU_{t,r',sc_{r',tc_0}} \leq eU_{t,r,pc_{r,tc_\infty}} - M(1-k_{t,r,r',p}) \quad \forall t\in T_I, r, r'\in R_t, p\in Pe^r\cap Ps^{r'} \tag{4.27}$$

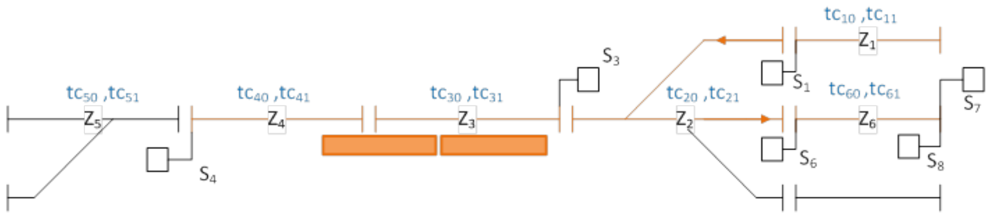
$$sU_{t,r,sc_{r,tc_0}} \leq eU_{t',r',pc_{r',tc_\infty}} - M(2-k_{t',r',r_\infty,p} - xR_{t,r}) \quad \forall t\in T_S, t'\in T_I(t), r\in R_t, r'\in R_{t'}, p\in Ps^r\cap Pe^{r'} \tag{4.28}$$

An additional set of constraints deals with formal track-circuit tc reservation. Their role is illustrated in Figure 4.3. Path r of train t represented in orange in Figure 4.3a contains formal track-circuits $TC^r = \{tc_0, tc_{10}, tc_{20}, tc_{30}, tc_{40}, tc_{31}, tc_{21}, tc_{60}, tc_\infty\}$. In Figure 4.3b, values of variables $sU_{t,r,tc}$, $eU_{t,r,tc}$, $oc_{t,r,tc}$ and $\phi_{t,r,tc}$ are represented for each formal track-circuit tc along r. Formal track-circuit tc_{10} is for example utilized by t between instants $sU_{t,r,tc_{10}}$ and $eU_{t,r,tc_{10}}$. The block section between signals S_1 and S_4 along r is reserved once t reaches formal track circuit tc_{20}. At the same moment, tc_{30} and tc_{40} are also reserved by t.

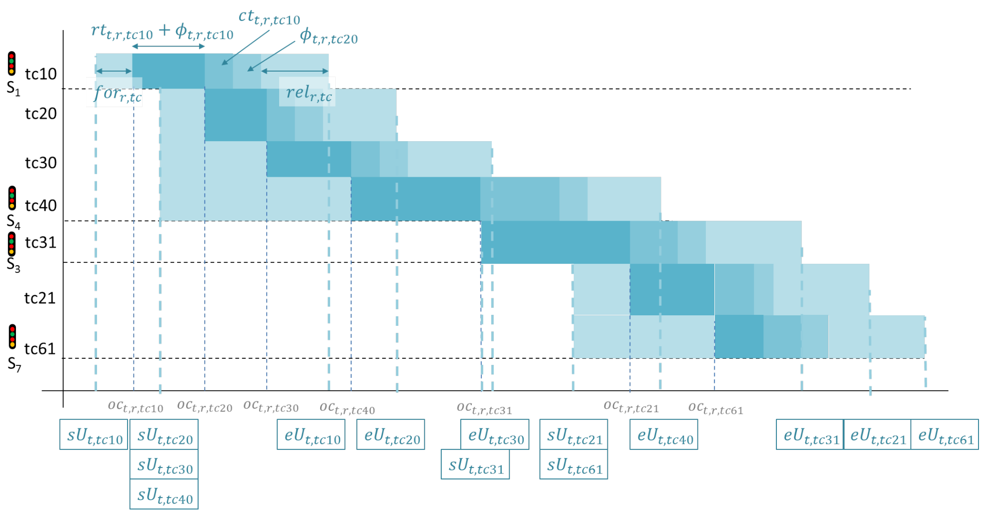
A train's utilization of a track-circuit along a route starts as soon as the train starts occupying the reference formal track-circuit $ref_{r,tc}$ for the reservation of tc minus the formation time (4.29). A train's utilization of a track-circuit along a route ends when the track-circuit has been physically cleared plus the release time (4.30). Thus, the equality considers running time, additional running time and clearing time on track-circuit tc along path r. Finally, it incorporates possible additional running time on following track-circuits if train t is long enough to occupy more than one track-circuit at a time. In this case, tc' exists such that tc is physically occupied by t while the head of t reaches the end of track-circuit tc', i.e. $tc \in OTC(t,r,tc')$. There are also disjunctive constraints (4.31)(4.32) so that two trains cannot utilize a track-circuit at the same time. These constraints are not applied for track-circuits of common shunting tracks to be coherent with track-circuit utilization constraints during parking (4.26)-(4.28).

$$sU_{t,r,tc} = oc_{t,r,ref_{r,tc}} - for_{bs_{r,tc}} xR_{t,r} \quad \forall t\in T, r\in R_t, tc\in TC^r \tag{4.29}$$

$$eU_{t,r,tc} = oc_{t,r,tc} + ((rt_{t,r,tc} + ct_{t,r,tc} + rel_{bs_{r,tc}})xR_{t,r} + \phi_{t,r,tc}) + \sum_{tc'\in TC:tc\in OTC(t,r,tc')} \phi_{t,r,tc'} \quad \forall t\in T, r\in R_t, tc\in TC^r \tag{4.30}$$



(a) Path r of the train t. A turnaround is necessary between signals S_3 and S_4.



(b) utilization of formal track circuits by t. Variables for occupation and utilization of track circuits represented in a timeline

Figure 4.3: Reservation of track circuits along a path

$$\begin{aligned} eU_{t,r,tc} - M(1 - yR_{t,t',r,r',tc,tc'}) \leq sU_{t',r',tc'} \\ \forall t,t' \in T, \text{index } t < \text{index } t', r \in R_t, r' \in R_{t'}, z \in Z^r \cap Z^{r'} \setminus \bigcup_{p \in P^r \cap P^{r'}} Z(p), \\ tc \in TC(z) \cap TC^r, tc' \in TC(z) \cap TC^{r'} \end{aligned} \tag{4.31}$$

$$\begin{aligned} eU_{t',r',tc'} - MyR_{t,t',r,r',tc,tc'} \leq sU_{t,r,tc} \\ \forall t,t' \in T, \text{index } t < \text{index } t', r \in R_t, r' \in R_{t'}, z \in Z^r \cap Z^{r'} \setminus \bigcup_{p \in P^r \cap P^{r'}} Z(p), \\ tc \in TC(z) \cap TC^r, tc' \in TC(z) \cap TC^{r'} \end{aligned} \tag{4.32}$$

Maintenance scheduling constraints

For maintenance operations, we specify the inequalities that must be verified at the beginning of the tasks. This must take into account the availability of crews and shunting tracks.

If an intermediate train t is created, any operation carried out on t can use only one crew and one shunting track along a given path (4.33). An operation of train t can not be performed on a shunting track p if t is not parked at p between two paths r and r' (4.34). The starting time of an operation is set to 0 if it is not assigned to a shunting track (4.35). An operation performed by crew hr must start after the shift start time of hr (4.36) and end before its shift end time (4.37). In particular with (4.37) a maintenance operation starts at 0 when it is not assigned to a crew or a track. For readability, we introduce the following expressions:

- $sP_{t,r,p}$ is t's arrival time on p through r
- $eP_{t,r',p}$ is t's departure time from p through r'.

An operation carried out on train t at shunting track p between paths r and r' needs to start after t's arrival on p through r and finish before t's departure from p through r'. t's arrival time on p through r is given by the expression $sP_{t,r,p}$ (4.38). If $r \neq \{r_0\}$, $sP_{t,r,p}$ is the moment when t starts utilizing the reference track-circuit for parking at p (4.39). Otherwise, $r = r_0$ and we need to consider the arriving train which uses the same rolling-stock. Then an intermediate train arrives at its first shunting track when its corresponding arriving train arrives (4.40). t's departure time from p through r' is given by the expression $eP_{t,r',p}$ (4.41). If $r \neq \{r_\infty\}$, $eP_{t,r,p}$ is the moment when t ends utilizing the reference track-circuit for parking at p (4.42). Otherwise, $r = r_\infty$ and we need to consider the departing train which uses the same rolling-stock. Then an intermediate train leaves its first shunting track when its corresponding departing train leaves (4.43),(4.44). If no departing train is assigned to t, then t stays at its last shunting track until the end of the planning period (4.45). Finally, if an operation o' follows an operation o, then o' starts after the end of o (4.46).

$$\sum_{hr \in HR^o, r, r' \in R_t, p \in Pe^r \cap Ps^{r'} \cap P^o} xO_{o,r,r',p,hr} \leq xT_t \quad \forall t \in T_I, o \in O_t \tag{4.33}$$

$$\sum_{hr \in HR^o} xO_{o,r,r',p,hr} \leq k_{t,r,r',p} \quad \forall t \in T_I, o \in O_t, r, r' \in R_t, p \in Pe^r \cap Ps^{r'} \cap P^o \tag{4.34}$$

$$\sum_{hr \in HR^o} sO_{o,r,r',p,hr} \leq M \sum_{hr \in HR^o} xO_{o,r,r',p,hr} \quad \forall t \in T_I, o \in O_t, r, r' \in R_t, p \in Pe^r \cap Ps^{r'} \cap P^o \tag{4.35}$$

$$\begin{gathered} sO_{o,r,r',p,hr} \geq sR_{hr} xO_{o,r,r',p,hr} \\ \forall t \in T_I, o \in O_t, r, r', \in R_t, p \in Pe^r \cap Ps^{r'} \cap P^o, hr \in HR^o \end{gathered} \tag{4.36}$$

$$\begin{gathered} sO_{o,r,r',p,hr} \leq (eR_{hr} - pR^o) xO_{o,r,r',p,hr} \\ \forall t \in T_I, o \in O_t, r, r', \in R_t, p \in Pe^r \cap Ps^{r'} \cap P^o, hr \in HR^o \end{gathered} \tag{4.37}$$

$$\begin{gathered} sO_{o,r,r',p,hr} \geq sP_{t,r,p} - M(1 - xO_{o,r,r',p,hr}) \\ \forall t \in T_I, o \in O_t, r, r' \in R_t, p \in Pe^r \cap Ps^{r'} \cap P^o, hr \in HR^o \end{gathered} \tag{4.38}$$

$$sP_{t,r,p} = sU_{t,r,tcp_{r,p}} \quad \forall t \in T_I, r \in R_t \setminus \{r_0, r_\infty\}, p \in Pe^r \tag{4.39}$$

$$sP_{t,r_0,p} = \sum_{r' \in R_{t'} : p \in Pe^{r'}} sU_{t',r',tcp_{r',p}} \quad \forall t' \in T_T, t \in T_I(t'), p \in \bigcup_{r \in R_t} Ps^r \tag{4.40}$$

$$\begin{gathered} sO_{o,p,r,r',hr} + pR^o xO_{o,p,r,hr} \leq eP_{t,r',p} + M(1 - xO_{o,p,r,hr}) \\ \forall t \in T_I, o \in O_t, r, r' \in R_t, p \in Pe^r \cap Ps^{r'} \cap P^o, hr \in HR^o \end{gathered} \tag{4.41}$$

$$eP_{t,r,p} = eU_{t,r,tcp_{r,p}} \quad \forall t \in T_I, r \in R_t \setminus \{r_0, r_\infty\}, p \in Ps^r \tag{4.42}$$

$$\begin{aligned} eP_{t,r_\infty,p} \geq \sum_{r' \in R_{t'}: p \in Pe^{r'}} eU_{t',r',tcp_{r',p}} - M(1 - xS_{t',t}) \\ \forall t' \in T_S, t \in T_I(t'), p \in \bigcup_{r \in R_t} Ps^r \end{aligned} \tag{4.43}$$

$$\begin{aligned} eP_{t,r_\infty,p} \leq \sum_{r' \in R_{t'}: p \in Pe^{r'}} eU_{t',r',tcp_{r',p}} + M(1 - xS_{t',t}) \\ \forall t' \in T_S, t \in T_I(t'), p \in \bigcup_{r \in R_t} Ps^r \end{aligned} \tag{4.44}$$

$$\begin{aligned} eP_{t,r_\infty,p} \geq \tau_M - M\left(1 - \sum_{t' \in T_S: t \in T_I(t')} xS_{t',t}\right) \\ \forall t \in T_I, p \in \bigcup_{r \in R_t} Ps^r \end{aligned} \tag{4.45}$$

$$\begin{aligned} sO_{o',r_1',r_2',p',hr'} \geq sO_{o,r_1,r_2,p,hr} + pR^o xO_{o,r_1,r_2,p,hr} - M(1 - xO_{o',r_1',r_2',p',hr'}) \\ \forall t \in T_I, \forall (o,o') \in E_t, r_1, r_1', r_2, r_2' \in R_t, p \in P^o \cap Pe^{r_1} \cap Ps^{r_2}, \\ p' \in P^{o'} \cap Pe^{r_1'} \cap Ps^{r_2'}, hr \in HR^o, hr' \in HR^{o'} \end{aligned} \tag{4.46}$$

As two operations cannot be performed by a crew at the same time, there are disjunctive constraints (4.47), (4.48).

$$\begin{aligned} sO_{o',r_1',r_2',p',hr} \geq sO_{o,r_1,r_2,p,hr} + pR^o xO_{o,r_1,r_2,p,hr} - M(1 - y_{o,o',hr}) \\ \forall t, t' \in T, o \in O_t, o' \in O_{t'}, hr \in HR^o \cap HR^{o'}, r_1, r_2 \in R_t, r_1', r_2' \in R_{t'}, \\ p \in P^o \cap Pe^{r_1} \cap Ps^{r_2}, p' \in P^{o'} \cap Pe^{r_1'} \cap Ps^{r_2'}, \text{index } t < \text{index } t' \end{aligned} \tag{4.47}$$

$$\begin{aligned} sO_{o,r_1,r_2,p,hr} \geq sO_{o',r_1',r_2',p',hr} + pR^{o'} xO_{o',r_1',r_2',p',hr} - My_{o,o',hr} \\ \forall t, t' \in T, o \in O_t, o' \in O_{t'}, hr \in HR^o \cap HR^{o'}, r_1, r_2 \in R_t, r_1', r_2' \in R_{t'}, \\ p \in P^o \cap Pe^{r_1} \cap Ps^{r_2}, p' \in P^{o'} \cap Pe^{r_1'} \cap Ps^{r_2'}, \text{index } t < \text{index } t' \end{aligned} \tag{4.48}$$

Besides, two operations cannot use a shunting track at the same time, and there are disjunctive constraints (4.49), (4.50).

$$\begin{aligned} sO_{o',r_1',r_2',p,hr'} \geq sO_{o,r_1,r_2,p,hr} + pR^o xO_{o,r_1,r_2,p,hr} - M(1 - yO_{o,o',p}) \\ \forall t, t' \in T, o \in O_t, o' \in O_{t'}, p \in P^o \cap P^{o'}, r_1, r_2 \in R_t, r_1', r_2' \in R_{t'}, \\ p \in P^o \cap Pe^{r_1} \cap Ps^{r_2} \cap P^{o'} \cap Pe^{r_1'} \cap Ps^{r_2'}, hr \in HR^o, hr' \in HR^{o'}, \text{index } t < \text{index } t' \end{aligned} \tag{4.49}$$

$$sO_{o,r_1,r_2,p,hr} \geq sO_{o',r'_1,r'_2,p',hr} + pR^{o'} xO_{o',r'_1,r'_2,p',hr} - MyO_{o,o',p}$$
$$\forall t, t' \in T, o \in O_t, o' \in O_{t'}, p \in P^o \cap P^{o'}, r_1, r_2 \in R_t, r'_1, r'_2 \in R_{t'}, \tag{4.50}$$
$$p \in P^o \cap Pe^{r_1} \cap Ps^{r_2} \cap P^{o'} \cap Pe^{r'_1} \cap Ps^{r'_2}, hr \in HR^o, hr' \in HR^{o'}, \text{index } t < \text{index } t'$$

Finally, we need to impose the protection of a shunting track during an operation. A disjunction sets that trains must enter a shunting track before the beginning (4.52) or after the end (4.51) of an operation. Another disjunction sets that trains must leave a shunting track before the beginning (4.54) or after the end (4.53) of an operation.

$$sP_{t,r,p} \geq \sum_{r_1,r_2 \in R_{t'}, hr \in HR^o : p \in Pe^{r_1} \cap Ps^{r_2}} sO_{o,r_1,r_2,p,hr} + pR^o xO_{o,r_1,r_2,p,hr} - M(1 - ysO_{o,t,r,p})$$
$$\forall t \in T_I, t' \in T_I, t \neq t', o \in O_{t'}, r \in R_t, p \in P^o \cap Pe^r \tag{4.51}$$

$$\sum_{r_1,r_2 \in R_{t'}, hr \in HR^o : p \in Pe^{r_1} \cap Ps^{r_2}} sO_{o,r_1,r_2,p,hr} \geq sP_{t,r,p} - MysO_{o,t,r,p}$$
$$\forall t \in T_I, t' \in T_I, t \neq t', o \in O_{t'}, r \in R_t, p \in P^o \cap Pe^r \tag{4.52}$$

$$eP_{t,r,p} \geq \sum_{r_1,r_2 \in R_{t'}, hr \in HR^o : p \in Pe^{r_1} \cap Ps^{r_2}} sO_{o,r_1,r_2,p,hr} + pR^o xO_{o,r_1,r_2,p,hr} - M(1 - yeO_{o,t,r,p})$$
$$\forall t \in T_I, t' \in T_I, t \neq t', o \in O_{t'}, r \in R_t, p \in P^o \cap Ps^r \tag{4.53}$$

$$\sum_{r_1,r_2 \in R_{t'}, hr \in HR^o : p \in Pe^{r_1} \cap Ps^{r_2}} sO_{o,r_1,r_2,p,hr} \geq eP_{t,r,p} - MyeO_{o,t,r,p}$$
$$\forall t \in T_I, t' \in T_I, t \neq t', o \in O_{t'}, r \in R_t, p \in P^o \cap Ps^r \tag{4.54}$$

Parking constraints

Parking constraints are based on constraints which involve precedence between events.

These precedences are based on the following additional variables:

- $xP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p}$ where $t_1, t_2 \in T_I$, $r_1 \in R_{t_1}$, $r'_1 \in R_{t_1}$, $r_2 \in R_{t_2}$, $r'_2 \in R_{t_2}$, $p \in Pe^{r_1} \cap Pe^{r_2} \cap Ps^{r'_1} \cap Ps^{r'_2}$, index $t_1 <$ index t_2 is equal to 1 if t_1 that enters track p by path r_1 and leaves by r'_1 and t_2 that enters track p by path r_2 and leaves by r'_2 use track p at same time and 0 otherwise,
- $yP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p}$ where $t_1, t_2 \in T_I$, $r_1 \in R_{t_1}$, $r'_1 \in R_{t_1}$, $r_2 \in R_{t_2}$, $r'_2 \in R_{t_2}$, $p \in Pe^{r_1} \cap Pe^{r_2} \cap Ps^{r'_1} \cap Ps^{r'_2}$, index $t_1 <$ index t_2 is equal to 1 if t_1 that enters track p by path r_1 and leaves by r'_1 uses track p before t_2 that enters track p by path r_2 and leaves by r'_2 and 0 otherwise,
- $sE_{t,r,p}$, with $t \in T_I$, $r \in R_t$, $p \in Pe^r$, is equal to 1 if t enters shunting track p through path r by side L and 0 if t enters p through r by side R,
- $eE_{t,r,p}$, with $t \in T_I$, $r \in R_t$, $p \in Ps^r$, is equal to 1 if t leaves shunting track p through path r by side L and 0 if t enters p through r by side R,
- $ysE_{t,t',r,r',p}$ with $t, t' \in T_I$, $r \in R_t$, $r' \in R_{t'}$, $p \in Ps^r \cap Ps^{r'}$, index $t <$ index t' is equal to 1 if t has to be placed on the left of t' to enter track p through path r and 0 if t has to be placed on the right of t' to enter track p through path r,

- $yeE_{t,t',r,r',p}$ where $t, t' \in T_I$, $r \in R_t$, $r' \in R_{t'}$, $p \in Pe^r \cap Pe^{r'}$, index $t <$ index t' is equal to 1 if t must be placed on the left of t' to leave track p through path r without crossing and 0 if t must be placed on the right of t' to leave track p through path r without crossing.

With these variables we can set length constraints: for every intermediate train t_1 which is parked at track p, the length of intermediate trains that use p at the same time as t_1 must be less or equal than the length of the track.

$$\begin{gathered} l_{t_1} k_{t_1,r_1,r'_1,p} + \sum_{t_2 \in T_I, r_2, r'_2 \in R_{t_2} : p \in Pe^{r_2} \cap Ps^{r'_2}} l_{t_2} xP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p} \leq L_p \\ \forall t_1 \in T_I, r_1, r'_1 \in R_{t_1}, p \in Pe^{r_1} \cap Ps^{r'_1} \end{gathered} \tag{4.55}$$

In a first disjunction, we impose that a shunting track p can be occupied by at most one intermediate train t_1, t_2 at any time, unless one of the variables $xP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p}$ is equal to 1. Either t_1 enters p after t_2 leaves it (4.56) or t_2 enters p after t_1 leaves it (4.57).

$$\begin{gathered} sP_{t_1,r_1,p} \geq eP_{t_2,r'_2,p} - M(yP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p} + xP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p}) \\ \forall t_1, t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, r'_2 \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2} \cap Ps^{r'_1} \cap Ps^{r'_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.56}$$

$$\begin{gathered} sP_{t_2,r_2,p} \geq eP_{t_1,r'_1,p} - M(1 - yP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p} + xP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p}) \\ \forall t_1, t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, r'_2 \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2} \cap Ps^{r'_1} \cap Ps^{r'_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.57}$$

Crossing constraints state that if two intermediate trains t_1, t_2 use a same shunting track p at the same time (then a variable $xP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p}$ is equal to 1), relative positions of t_2 and t_1 when they enter the track and when they leave the track must be equal. If t_1 has to placed on the left of t_2 to leave the track without crossing, t_1 cannot be placed on the right of t_2 when it enters the track (4.58). If t_1 has to placed on the right of t_2 to leave the track without crossing, t_1 cannot be placed on the left of t_2 when it enters the track (4.59).

$$\begin{gathered} yeE_{t_1,t_2,r'_1,r'_2,p} + (1 - ysE_{t_1,t_2,r_1,r_2,p}) \leq 2 - xP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p} \\ \forall t_1, t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, r'_2 \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2} \cap Ps^{r'_1} \cap Ps^{r'_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.58}$$

$$\begin{gathered} (1 - yeE_{t_1,t_2,r'_1,r'_2,p}) + ysE_{t_1,t_2,r_1,r_2,p} \leq 2 - xP_{t_1,t_2,r_1,r'_1,r_2,r'_2,p} \\ \forall t_1, t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, r'_2 \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2} \cap Ps^{r'_1} \cap Ps^{r'_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.59}$$

Positioning variables $ysE_{t_1,t_2,r_1,r_2,p}$ and $yeE_{t_1,t_2,r'_1,r'_2,p}$ are deduced thanks to disjunctions. These disjunctions are based on two assertions:

- if t_1 enters shunting track p by the left (or right) side after t_2 enters p, then t_1 is placed on the left (or right) side of t_2 (Figure 4.4) (4.60)-(4.63)

- if t_1 leaves shunting track p by the left (or right) side before t_2 leaves p, t_1 has to be placed on the left (or right) side of t_2 so that no crossing occurs (Figure 4.5) (4.64)-(4.67)

The first assertion leads to Table 4.2 which presents a disjunction for variables $ysE_{t_1,t_2,r_1,r_2,p}$. An example can illustrate (4.60). Suppose that t_1 enters p through r_1 by the left side after t_2 enters p through r_2. Then $sP_{t_1,r_1,p} > sP_{t_2,r_2,p}$. As t_1 enters through the left side, $sE(r_1, p) = 1$. Therefore we must have $ysE_{t_1,t_2,r_1,r_2,p} = 1$.

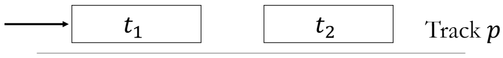


Figure 4.4: An intermediate train t_1 enters a track p through the left side after an intermediate train t_2. Then t_1 is placed on the left side of t_2.

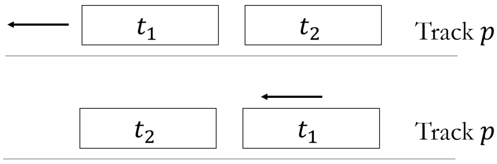


Figure 4.5: An intermediate train t_1 leaves a track p through the left side before an intermediate train t_2. Then t_1 must be placed on the left side of t_2 in order to leave without crossing.

$$\begin{gathered} M(ysE_{t_1,t_2,r_1,r_2,p} + 1 - sE_{t_1,r_1,p}) \geq sP_{t_1,r_1,p} - sP_{t_2,r_2,p} \\ \forall t_1, t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.60}$$

$$\begin{gathered} M(1 - ysE_{t_1,t_2,r_1,r_2,p} + sE_{t_1,r_1,p}) \geq sP_{t_1,r_1,p} - sP_{t_2,r_2,p} \\ \forall t_1, t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.61}$$

$$\begin{gathered} M(ysE_{t_1,t_2,r_1,r_2,p} + sE_{t_2,r_2,p}) \geq sP_{t_2,r_2,p} - sP_{t_1,r_1,p} \\ \forall t_1, t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.62}$$

$$\begin{gathered} M(1 - ysE_{t_1,t_2,r_1,r_2,p} + 1 - sE_{t_2,r_2,p}) \geq sP_{t_2,r_2,p} - sP_{t_1,r_1,p} \\ \forall t_1, t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.63}$$

The second assertion leads to Table 4.3 which presents a disjunction for variables $yeE_{t_1,t_2,r'_1,r'_2,p}$.

$$\begin{gathered} M(yeE_{t_1,t_2,r'_1,r'_2,p} + 1 - Ee(r'_1,p)) \geq eP_{t_2,r'_2,p} - eP_{t_1,r'_1,p} \\ \forall t_1, t_2 \in T_I, r'_1 \in R_{t_1}, r'_2 \in R_{t_2}, p \in Pe^{r'_1} \cap Pe^{r'_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.64}$$

Table 4.2: Values of variable $ysE_{t_1,t_2,r_1,r_2,p}$, with $t_1, t_2 \in T_I$, index $t_1 <$ index t_2, $r_1 \in R_{t_1}, r_2 \in R_{t_2}$, $p \in Pe^{r_1} \cap Pe^{r_2}$. $ysE_{t_1,t_2,r_1,r_2,p}$ is equal to 1 if t_1 is placed on the left side of t_2 and 0 if t_1 is placed on the right side of t_2.

		$ysE_{t_1,t_2,r_1,r_2,p}$	
Entrance side of t_1 on p $sE(r_1,p)$	Entrance side of t_2 on p $sE(r_2,p)$	t_1 enters before t_2	t_2 enters before t_1
1 (Left)	1 (Left)	0	1
1 (Left)	0 (Right)	1	1
0 (Right)	1 (Left)	0	0
0 (Right)	0 (Right)	1	0

Table 4.3: Values of variable $yeE_{t_1,t_2,r'_1,r'_2,p}$, with $t_1, t_2 \in T_I$, index $t_1 <$ index t_2, $r'_1 \in R_{t_1}, r'_2 \in R_{t_2}$, $p \in Pe^{r'_1} \cap Pe^{r'_2}$. $ysE_{t_1,t_2,r'_1,r'_2,p}$ is equal to 1 if t_1 is placed on the left side of t_2 and 0 if t_1 is placed on the right side of t_2.

		$yeE_{t_1,t_2,r'_1,r'_2,p}$	
Exit side of t_1 on p $eE_{t_1,r'_1,p}$	Exit side of t_2 on p $eE_{t_2,r'_2,p}$	t_1 leaves before t_2	t_2 leaves before t_1
1 (left)	1 (left)	1	0
1 (Left)	0 (Right)	1	1
0 (Right)	1 (Left)	0	0
0 (Right)	0 (Right)	0	1

$$\begin{aligned} M(1 - yeE_{t_1,t_2,r'_1,r'_2,p} + Ee(r'_1, p)) \geq eP_{t_2,r'_2,p} - eP_{t_1,r'_1,p} \\ \forall t_1, t_2 \in T_I, r'_1 \in R_{t_1}, r'_2 \in R_{t_2}, p \in Pe^{r'_1} \cap Pe^{r'_2}, \text{index } t_1 < \text{index } t_2 \end{aligned} \tag{4.65}$$

$$\begin{aligned} M(yeE_{t_1,t_2,r'_1,r'_2,p} + Ee(r'_2, p)) \geq eP_{t_1,r'_1,p} - eP_{t_2,r'_2,p} \\ \forall t_1, t_2 \in T_I, r'_1 \in R_{t_1}, r'_2 \in R_{t_2}, p \in Pe^{r'_1} \cap Pe^{r'_2}, \text{index } t_1 < \text{index } t_2 \end{aligned} \tag{4.66}$$

$$\begin{aligned} M(1 - yeE_{t_1,t_2,r'_1,r'_2,p} + 1 - Ee(r'_2, p)) \geq eP_{t_1,r'_1,p} - eP_{t_2,r'_2,p} \\ \forall t_1, t_2 \in T_I, r'_1 \in R_{t_1}, r'_2 \in R_{t_2}, p \in Pe^{r'_1} \cap Pe^{r'_2}, \text{index } t_1 < \text{index } t_2 \end{aligned} \tag{4.67}$$

We also need to consider the special case in which two intermediate trains t_1 and t_2 enter or leave a track p coupled, and hence at a same time. Their relative positions on p is given by either a_{t_1,t_2}, which is their positions once coupled in a same departing train, or b_{t_1,t_2}, which is their positions when they are coupled in a same arriving train.

If two intermediate trains t_1 and t_2 arrive coupled in a same track, their relative position given by variables $ysE_{t_1,t_2,r_1,r_2,p}$ is determined by b_{t_1,t_2} (4.68)(4.69). If an intermediate train t_1 is placed on the left of an another one t_1 in the same arriving train, t_1 obviously remains on the left of t_2 once they are parked at a shunting track. In this example $b_{t_1,t_2} = 1$. If p is the first track that t_1 and t_2 use ($k_{t_1,r_0,r_1,p} = 1$ and $k_{t_2,r_0,r_2,p} = 1$), then $3 - k_{t_1,r_0,r_1,p} - k_{t_2,r_0,r_2,p} = 1$. In this case, we must have $ysE_{t_1,t_2,r_1,r_2,p} = b_{t_1,t_2} = 1$, t_1 is located left of t_2 on the track p.

If two intermediate trains t_1 and t_2 which do not belong to a same type are coupled to be assigned to a same departing train t_0, the relative position of t_1 and t_2 given by variables $yeE_{t_1,t_2,r_1,r_2,p}$ must be equal to a_{t_1,t_2} (4.70),(4.71). We remark that we only consider intermediate trains that do not belong to a same type because trains of a same type are interchangeable.

$$\begin{aligned} b_{t_1,t_2} + (1 - ysE_{t_1,t_2,r_1,r_2,p}) \leq 3 - k_{t_1,r_0,r_1,p} - k_{t_2,r_0,r_2,p} \\ \forall t_0 \in T_T, t_1, t_2 \in T_I(t_0), \text{index } t_1 < \text{index } t_2, r_1 \in T_{t_1}, r_2 \in T_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2} \end{aligned} \tag{4.68}$$

$$\begin{aligned} (1 - b_{t_1,t_2}) + ysE_{t_1,t_2,r_1,r_2,p} \leq 3 - k_{t_1,r_0,r_1,p} - k_{t_2,r_0,r_2,p} \\ \forall t_0 \in T_T, t_1, t_2 \in T_I(t_0), \text{index } t_1 < \text{index } t_2, r_1 \in R_{t_1}, r_2 \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2} \end{aligned} \tag{4.69}$$

$$\begin{aligned} a_{t_1,t_2} + (1 - yeE_{t_1,t_2,r'_1,r'_2,p}) \leq 5 - k_{t_1,r_1,r_\infty,p} - xS_{t_0,t_1} - k_{t_2,r_2,r_\infty,,p} - xS_{t_0,t_2} \\ \forall t_0 \in T_S, t_1, t_2 \in T_I(t_0), \text{index } t_1 < \text{index } t_2, r'_1 \in R_{t_1}, r'_2 \in R_{t_2}, p \in Ps^{r_1} \cap Ps^{r_2}, ty_{t_1} \neq ty_{t_2} \end{aligned} \tag{4.70}$$

$$(1 - a_{t_1,t_2}) + yeE_{t_1,t_2,r'_1,r'_2,p} \leq 5 - k_{t_1,r_1,r_\infty,p} - xS_{t_0,t_1} - k_{t_2,r_2,r_\infty,p} - xS_{t_0,t_2}$$
$$\forall t_0 \in T_S, t_1, t_2 \in T_I(t_0), \text{index } t_1 < \text{index } t_2, r'_1 \in R_{t_1}, r'_2 \in R_{t_2}, p \in Ps^{r_1} \cap Ps^{r_2}, ty_{t_1} \neq ty_{t_2} \tag{4.71}$$

Variables $sE_{t,r,p}$ are set in case r is not a dummy path (4.72) and in case $r = r_0$ (4.73). Variables $eE_{t,r,p}$ are set in case r is not a dummy path (4.74) and in case $r = r_\infty$ (4.75),(4.76).

$$sE_{t,r,p} = Es(r,p) \quad \forall t \in T_I, r \in R_t \setminus \{r_0\}, p \in Pe^r \tag{4.72}$$

$$sE_{t,r_0,p} = \sum_{r \in R_{t_0}} Es(r,p) xR_{t_0,r} \quad \forall t_0 \in T_T, t \in T_I(t_0) \tag{4.73}$$

$$eE_{t,r,p} = Ee(r,p) \quad \forall t \in T_I, r \in R_t \setminus \{r_\infty\}, p \in Ps^r \tag{4.74}$$

$$sE_{t,r_\infty,p} \geq \sum_{t_0 \in T_S, r \in R_{t_0} : t \in T_I(t_0)} Es(r,p)(xR_{t_0,r} + xS_{t_0,t} - 2) \quad \forall t_0 \in T_T, t \in T_I(t_0) \tag{4.75}$$

$$1 - sE_{t,r_\infty,p} \geq \sum_{t_0 \in T_S, r \in R_{t_0} : t \in T_I(t_0)} (1 - Es(r,p))(xR_{t_0,r} + xS_{t_0,t} - 2) \quad \forall t_0 \in T_T, t \in T_I(t_0) \tag{4.76}$$

4.2.2 Model strengthening

In this section, we propose ways for strengthening the formulation described in Section 4.2.1. They are based on the relaxation of integrality constraints, on the reduction of the number of binary variables and on the removal of symmetries in the branching tree.

Removing integrality constraints for specific variables

Integrality constraint on variables qS_t which indicate if departing train t is cancelled or not can be removed. Then we only need to specify:

$$0 \leq qS_t \leq 1 \ \forall t \in T_S.$$

Variable qS_t appears in (4.3) and (4.5). If at least one intermediate train t' is assigned to t ($xS_{t,t'} = 1$), then the right hand side of (4.5) is greater than or equal to 1. Therefore, $1 - qS_t$ must be greater than or equal to 1, thus $qS_t = 0$. On contrary, if no intermediate train is assigned to t, then the right hand side of (4.3) is equal to 0 for any type of train units $tu \in TU$. In particular for $tu \in TU$ such that $m_{t,tu} > 0$ we have $m_{t,tu}(1 - qS_t) = 0$. Therefore $1 - qS_t$ must be equal to 0, thus $qS_t = 1$. This integrality removal speeds up the preprocessing operated by MILP solvers.

Integrality constraints on variables u_t, v_t which respectively represent the number of uncoupling operations on an arriving train or coupling operations on a departing train can be released, too. u_t and v_t only appear respectively in (4.6) and (4.7). In these inequalities the right hand sides are integers. u_t and v_t also appear in the objective function which is minimized. Therefore, u_t and v_t will automatically set equal to the right hand sides of the constraints.

Reducing the number of disjunction variables for routing

In this section, as Pellegrini et al. [2019], we propose to reduce the number of binary variables used for scheduling shunting movements as well as the rest of a station traffic. More precisely, for two formal track-circuits tc, tc' which correspond to the same real track-circuit, we consider variables $yR_{t,t',r,r',tc,tc'}$: they indicate if train t uses tc along path r before t' uses tc' along r'. Some of these variables may be redundant depending of the path sequences. For example, consider the case in which the sequences of real track-circuits of r and r' are respectively $Z^r = \{z_0, z_1, z_2, z_3, z_4, z_5, z_\infty\}$ and $Z^{r'} = \{z_0, z_9, z_8, z_4, z_3, z_2, z_\infty\}$. Then, the sequences of formal track-circuits of these paths are respectively $TC^r = \{tc_0, tc_{10}, tc_{20}, tc_{30}, tc_{40}, tc_{50}, tc_\infty\}$ and $TC^{r'} = \{tc_0, tc_{91}, tc_{81}, tc_{41}, tc_{31}, tc_{21}, tc_\infty\}$. Here, for $i = 1, ..., 9$ formal track-circuits corresponding to real track-circuit z_i are $TC(z_i) = \{tc_{i0}, tc_{i1}\}$. In this example, both r and r' contain track-circuits z_2, z_3 and z_4 and they use them in opposite directions. Then, constraints (4.31),(4.32) have to be defined at each of these track-circuits. However, as t and t' move in opposite directions, t' cannot start using z_4 when t uses either z_2, z_3 or z_4. Similarly, t cannot start using z_2 when t' uses either z_2, z_3 or z_4. Therefore $yR_{t,t',r,r',tc_{20},tc_{21}} = yR_{t,t',r,r',tc_{30},tc_{31}} = yR_{t,t',r,r',tc_{40},tc_{41}}$. It is possible to consider a single disjunction variable for track-circuits z_2,z_3 and z_4. This variable indicates which train enters first in the section composed of z_2,z_3 and z_4. Likewise, when two paths follow a same sequence of track-circuits in the same direction the values of variables $yR_{t,t',r,r',tc,tc'}$ for tc, tc' in this sequence do not change.

Following this reasoning, for two paths r,r' we identify the sequences of consecutive common track-circuits. We name these sequences sections. A section contains couples of formal track-circuits (tc, tc') such that $tc \in TC^r$ and $tc' \in TC^{r'}$ correspond to a same real track-circuit. In the example of the previous paragraph, paths r and r' share the section composed by $\{z_2, z_3, z_4\}$. This section contains the couples of formal track-circuits $\{(tc_{20}, tc_{21}), (tc_{30}, tc_{31}), (tc_{40}, tc_{41})\}$. In the general case, we denote $S_{r,r'}$ the set of sections that share two paths r and r'. These sections do not include track-circuits of common shunting tracks of r and r'.

In the formulation, sets $S_{r,r'}$ are input data. In Section 4.3 we propose an algorithm to compute these sets. We consider binary variables:

- $yR_{t,t',r,r',s}$ where $t, t' \in T$, index $t <$ index t', $r \in R_t$, $r' \in R_{t'}$,$s \in S_{r,r'}$ is equal to 1 if t uses the section s before t' and 0 otherwise

Inequalities (4.31), (4.32) can be replaced by (4.77),(4.78). In (4.77),(4.78) the variables $yR_{t,t',r,r',s}$ which represent precedence at each section are used. These variables are less numerous than variables $yR_{t,t',r,r',tc,tc'}$. In Section 4.4, we evaluate the benefit of this alternative formulation in a case study.

$$\begin{gathered} eU_{t,r,tc} - M(1 - yR_{t,t',r,r',s}) \leq sU_{t',r',tc'} \\ \forall t, t' \in T, \text{index } t < \text{index } t', r \in R_t, r' \in R_{t'}, s \in S_{r,r'}, (tc, tc') \in s \end{gathered} \tag{4.77}$$

$$\begin{gathered} eU_{t',r',tc'} - MyR_{t,t',r,r',s} \leq sU_{t,r,tc} \\ \forall t, t' \in T, \text{index } t < \text{index } t', r \in R_t, r' \in R_{t'}, s \in S_{r,r'}, (tc, tc') \in s \end{gathered} \tag{4.78}$$

Breaking Symmetries for disjunction variables

In the formulation, many binary variables are involved in disjunctive scheduling constraints. These are disjunctive variables. These variables are defined for two shunting operations (maintenance, routing or parking) and a common resource. They indicate how two shunting operations use a common resource. If none of the operations use the possibly common resource, the disjunctive variable can be equal to either 1 or 0. Therefore, the MILP formulation has symmetries and a MILP solver may have to explore equivalent nodes in a branching tree. In order to break these symmetries, we impose that disjunctive variables are equal to 0 if the corresponding operations do not use the common resource.

We impose this constraint for variables $yR_{t,t',r,r',tc,tc'}$ which represent a disjunction for the use of a track-circuit by two trains t and t'. $yR_{t,t',r,r',tc,tc'}$ is equal to 0 if neither r nor r' are assigned respectively to t and t' (4.79). Variables $y_{o,o',hr}$ which represent a disjunction for the use of a crew hr by two maintenance operations o and o' are equal to 0 if neither o nor o' uses crew hr (4.80). The same holds for variables $yP_{o,o',p}$ which indicate the order in which o and o' use shunting track p (4.81). Variable $ysO_{o,t,r,p}$, which indicates if intermediate train t enters in shunting track p trough path r before maintenance operation o starts or after o finishes, is equal to 0 if o is not performed on p and t does not use r (4.82).

An equivalent constraint (4.83) is added for variable $yeO_{o,t,r,p}$ which indicates if intermediate train t leaves shunting track p trough path r before maintenance operation o starts or after o finishes. Variables $yP_{t_1,t_2,r_1,r_1',r_2,r_2',p}$ and $xP_{t_1,t_2,r_1,r_1',r_2,r_2',p}$ which represent the order in which two intermediate trains t_1, t_2 use shunting track p for parking are equal to 0 if neither t_1 or t_2 is parked on p (4.84), (4.85). When $xP_{t_1,t_2,r_1,r_1',r_2,r_2',p} = 1$, we impose that $yP_{t_1,t_2,r_1,r_1',r_2,r_2',p} = 0$ (4.86). Similarly, variable $ysE_{t_1,t_2,r_1,r_2,p}$ gives the relative position of t_1 and t_2 on p if t_1 uses path r_1 and t_2 uses path r_2. $ysE_{t_1,t_2,r_1,r_2,p}$ is equal to 0 if neither t_1 uses r_1 nor t_2 uses r_1 (4.87). An equivalent inequality is set for variable $yeE_{t_1,t_2,r_1,r_2,p}$ which indicates crossing free relative positions (4.88).

$$\begin{gathered} yR_{t,t',r,r',tc,tc'} \leq xR_{t,r} + xR_{t',r'} \\ \forall t,t' \in T, r \in R_t, r' \in R_{t'}, z \in Z^r \cap Z^{r'}, tc \in TC(z) \cap TC^r, tc' \in TC(z) \cap TC^{r'}, \text{index } t < \text{index } t' \end{gathered} \tag{4.79}$$

$$\begin{gathered} y_{o,o',hr} \leq \sum_{r,r' \in R_t, p \in Pe^r \cap Ps^{r'} \cap P^o} xO_{o,r,r',p,hr} + \sum_{r,r' \in R_{t'}, p \in Pe^r \cap Ps^{r'} \cap P^o} xO_{o',r,r',p,hr} \\ \forall t,t' \in T_I, \text{index } t < \text{index } t', o \in O_t, o' \in O_{t'}, hr \in HR^o \cap HR^{o'} \end{gathered} \tag{4.80}$$

$$\begin{gathered} yO_{o,o',p} \leq \sum_{hr \in HR^o, r,r' \in R_t : p \in Pe^r \cap Ps^r} xO_{o,r,r',p,hr} + \sum_{hr \in HR^o, r,r' \in R_{t'} : p \in Pe^r \cap Ps^r} xO_{o',r,r',p,hr} \\ \forall t,t' \in T_I, \text{index } t < \text{index } t', o \in O_t, o' \in O_{t'}, p \in P^o \cap P^{o'} \end{gathered} \tag{4.81}$$

$$\begin{gathered} ysO_{o,t,r,p} \leq \sum_{hr \in HR^o, r_1,r_2 \in R_{t'} : p \in Pe^{r_1} \cap Ps^{r_2}} xO_{o,r_1,r_2,p,hr} + xR_{t,r} \\ \forall t \in T_I, t' \in T_I, t \neq t', o \in O_{t'}, r \in R_t, p \in P^o \cap Pe^r \end{gathered} \tag{4.82}$$

$$\begin{gathered} yeO_{o,t,r,p} \leq \sum_{hr \in HR^o, r_1,r_2 \in R_{t'} : p \in Pe^{r_1} \cap Ps^{r_2}} xO_{o,r_1,r_2,p,hr} + xR_{t,r} \\ \forall t \in T_I, t' \in T_I, t \neq t', o \in O_{t'}, r \in R_t, p \in P^o \cap Ps^r \end{gathered} \tag{4.83}$$

$$\begin{gathered} yP_{t_1,t_2,r_1,r_1',r_2,r_2',p} \leq k_{t_1,r_1,r_1',p} + k_{t_2,r_2,r_2',p} \\ \forall t_1,t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, r_2' \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2} \cap Ps^{r_1'} \cap Ps^{r_2'}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.84}$$

$$\begin{gathered} xP_{t_1,t_2,r_1,r_1',r_2,r_2',p} \leq k_{t_1,r_1,r_1',p} + k_{t_2,r_2,r_2',p} \\ \forall t_1,t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, r_2' \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2} \cap Ps^{r_1'} \cap Ps^{r_2'}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.85}$$

$$\begin{gathered} yP_{t_1,t_2,r_1,r_1',r_2,r_2',p} \leq 1 - xP_{t_1,t_2,r_1,r_1',r_2,r_2',p} \\ \forall t_1,t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, r_2' \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2} \cap Ps^{r_1'} \cap Ps^{r_2'}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.86}$$

$$\begin{gathered} ysE_{t_1,t_2,r_1,r_2,p} \leq xR_{t_1,r_1} + xR_{t_2,r_2} \\ \forall t_1,t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, p \in Pe^{r_1} \cap Pe^{r_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.87}$$

$$\begin{gathered} yeE_{t_1,t_2,r_1,r_2,p} \leq xR_{t_1,r_1} + xR_{t_2,r_2} \\ \forall t_1,t_2 \in T_I, r_1 \in R_{t_1}, r_2 \in R_{t_2}, p \in Ps^{r_1} \cap Ps^{r_2}, \text{index } t_1 < \text{index } t_2 \end{gathered} \tag{4.88}$$

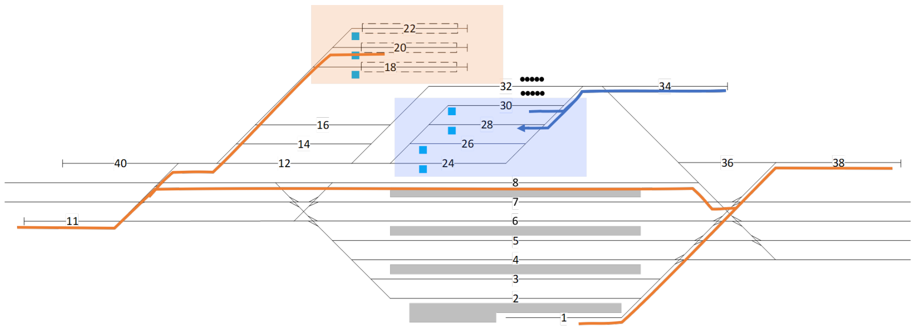


Figure 4.6: Simple example of station layout. The blue path links to shunting tracks that belong to the same group which is blue squared. The orange path links the bottom platform with a shunting track in the orange filled rectangle.

4.3 Instances generation and checking

In this section, we consider the preprocessing phase in which a G-TUSP instance is set. We focus on instance generation aspects that typically affect solution quality and computation time.

We first look at on the generation of the set of alternative paths for trains. Then, we report a method to check a necessary condition for instance feasibility before starting the MILP solution.

4.3.1 Path generation

In section 4.1.3, we noticed that routes are defined by signalling system specifications. Basically an infrastructure contains a set of routes. In order to have a G-TUSP instance, a set of paths is generated thanks to the existing set of routes. A second input is a set of trains, for which a starting and finishing points in the infrastructure are specified. Such points are signals at which train paths have to start or finish. A passing train has one starting point and one finishing point. It may also have to pass through a specific platform. An arriving train has one starting point and several possible finishing points. The number of finishing points increases with the number of shunting tracks it can access. A departing has one finishing point and several possible starting points. The number of starting points increases with the number of shunting tracks it can access. An intermediate train t has several possible starting and finishing points which are located on tracks where t can be parked. In this section, we report how we generate the set of alternative paths to obtain G-TUSP instances. Moreover, we show how to derive common sections to these paths, needed to strengthen the model as discussed in Section 4.2.2.

Generating paths with a set of routes

Set R_t is computed by looking for all paths that link starting and finishing points of t.

When we do so, we consider a maximum number of turnaround for each path, and we set it as small as possible depending on the infrastructure layout. In particular, consider the example in Figure 4.6. Here, one turnaround is allowed to move between tracks in the same group (blue path). Instead, two turnarounds must be allowed to move between some tracks and some platforms (orange path). Considering maximum number of turnarounds significantly reduces the number of possible paths. For the infracture of Metz-Ville station (Section 4.4) the algorithm generates 8567 paths while 56949 paths are obtained with a simple breadth first search algorithm.

Computing the set of sections for track-circuit utilization disjunction

In Section 4.2.2, we introduced set $S_{r,r'}$ of sections shared by paths r and r'. $S_{r,r'}$ is made of subsequences of formal track-circuits in TC^r and $TC^{r'}$. These subsequences can consider TC^r and $TC^{r'}$ in the same order or in the opposite order as proposed in Algorithm 1. This algorithm requires $|TC^r||TC^{r'}|$ iterations.

Algorithm 1: SECTIONS GENERATION

```
Data: two paths r and r'
tc ← first formal track-circuit in TC^r
while tc is unmarked do
    mark tc, unmark all elements of TC^{r'}
    tc' ← first formal track-circuit in TC^{r'}
    while tc' is unmarked do
        if tc' corresponds to the same real track-circuit of tc and tc' is unmarked then
            mark tc'
            create a section s with (tc, tc')
            tc ← formal track-circuit which follows tc in TC^r
            tcf' ← formal track-circuit which follows tc' in TC^{r'}
            tcp' ← formal track-circuit which precedes tc' in TC^{r'}
            while tcf' corresponds to the same track circuit of tc do
                mark tc and tcf'
                add (tc, tcf') to s
                tc ← formal track circuit which follows tc in TC^r
                tcf' ← formal track circuit which follows tcf' in TC^{r'}
            while tcp' corresponds to the same track circuit of tc do
                mark tc and tcp'
                add (tc, tcp') to s
                tc ← formal track circuit which follows tc in TC^r
                tcp' ← formal track circuit which precedes tcp' in TC^{r'}
            add s to S_{r,r'}
        tc' ← first unmarked element of TC^{r'} which follows tc'
```

4.3.2 Alternative paths number reduction

The size of the set of paths R_t that can be assigned to a train t depends of two factors:

- the multiplicity of paths between a starting point and a finishing point
- the multiplicity of starting points and finishing points

The first factor concerns all trains, while the second one is specific to shunted trains. Indeed, the multiplicity of starting or finishing points is due to the multiplicity of shunting tracks. In this section, we provide techniques to limit the size of R_t considering these two factors.

Reducing the number of alternative paths between two points

We implement a selection of alternative paths based on a simplification of the approach given by Riezebos and Van Wezel [2009]. As for shunted trains paths are defined between shunting tracks or between shunting tracks and platforms, there is no use to consider the sequence of tracks a route goes through, as

Riezebos and Van Wezel [2009] do. In addition some paths that are deemed to be useless are removed. More precisely, the paths that link two shunting tracks passing through platforms are not taken into account.

Then for a starting point and a finishing point we consider the routes between two points with the lowest nominal running time. For an integer k, we limit the number of alternative routes between two points by searching for the k-shortest paths.

Reducing the number of shunting tracks available

If the number of alternative routes between two points can be reduced to an integer k, the number of routes available for a train may remain high. Indeed, there are at least k routes for every couple of starting and finishing points. But for passing trains, in principle this corresponds to the number of reachable shunting tracks. Nevertheless, there are some strategies to reduce the number of shunting tracks available to each train. The tracks that are not long enough to host a train are obviously removed from the set of available tracks. Moreover, dispatchers tend to dedicate shunting tracks to particular rolling stock types. We assume that trains only use their dedicated tracks. This issue is investigated by Haahr et al. [2017], who study several ways to split a yard to specific rolling stock types. We consider that this splitting is an input data.

4.3.3 Aggregated capacity feasibility check

In G-TUSP, a cause of instance infeasiblity may be the fact that train units cannot be parked or routed. A feasible solution must satisfy a necessary condition based on an aggregate capacity criterion (Haahr et al. [2017]). We need to check that the total length of trains that occupy the shunting yard at the same time does not exceed the total tracks length. The trains length can be computed each time a train leaves or enter the shunting yard. These times are given by the timetable. Then, this aggregated capacity criterion can be checked in $O(|T_T| + |T_S|)$ time.

4.4 Experiments

In this section, we report on experiments that test the model on a panel of instances. The model is coded in Java and solved exactly using the commercial solver CPLEX considering at most one-hour computation. Beyond this time there is no practical interest for pre-operational planning. The experiment are executed on a 64 bit operation system equipped with a 2.2 GHz Intel®Core™i5-8400 processor and 16GB RAM. We study the Metz-Ville station infrastructure. It is a major hub for Eastern France railway traffic. We tackle real scenarios which include disturbances such as arrival delay or track closure.

4.4.1 Instances description

We consider traffic in the Metz-Ville infrastructure and its passengers shunting yards represented in Figure 4.7. It is a major junction where the Nancy-Luxembourg and Metz-Strasbourg lines intersect. The station mainly hosts regional trains. Many of these trains start or end their service in Metz-Ville. The area is 3.8 km long and has 10 platforms including a dead-end one. Yards F1 and F2 are controlled from the signal box, while switches are directly handled by a ground-agent in yards F3 and F4. Yard F3 contains a track with a washing machine. Two tracks in yard F3 and one track in yard F4 have equipment for technical inspection. A predefined coupling exists between the other shunting track and rolling stock types. Hence, each train can be parked only on a subset of these tracks. The infrastructure is composed of 138 track-circuits, 68 signals, 421 block sections and 405 routes.

We consider two regular week days and two disturbed week days in 2018. One disturbed day includes several delays from Luxembourg between 16:30 and 19:40. These delays were due to urgent infrastructure maintenance works during the day. Such delays are known by dispatchers at 15:15. During the other disturbed day, one of the two north side shunting necks is closed. This shunted neck is circled in the red (Figure 4.7) and the available one is circled in green. During the other perturbed day, one of the two north side shunting necks is closed because of a major track failure. A shunting neck is a track used for train turnarounds during shunting movements. The closed shunted neck is circled in red in Figure 4.7 (north side, up). Another neck remains available on the north side of the station, and it is circled in green in the figure. The track closure scenario reduces the set of possible routes and increase the likelihood of the occurrence of conflicts. For example, if a train has to be moved from yard F2 to yard F4, in this scenario it has to use the green shunting neck and traverse the main station tracks on which passing trains travel too. For each of the four days, we consider a day and a night time scenario. The former includes traffic between between the morning (7:00) and the evening peak hour (19:00). In this interval, between 14 and 18 train units have to be shunted and there are between 241 and 243 passing trains. The latter includes trains between evening (18:30) and the next morning peak hour (7:30). In this interval, between 19 and 26 train units have to be shunted and there are between 158 and 165 passing trains.

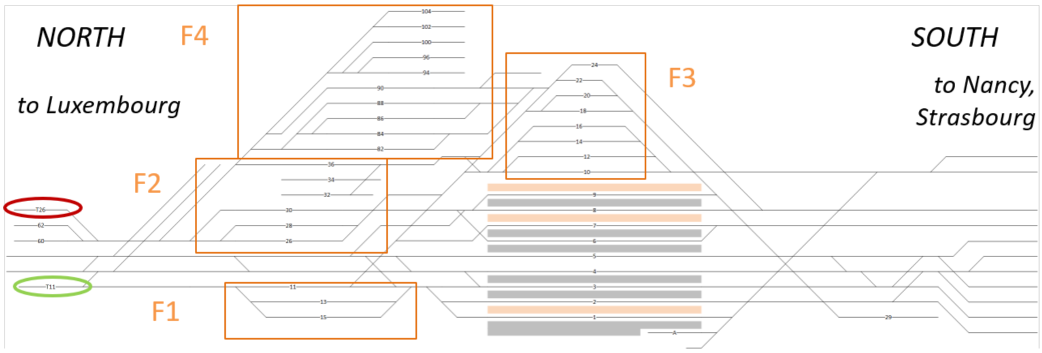


Figure 4.7: Layout of Metz-Ville station

Table 4.4: Details on the instances tackled in the experimental analysis ($|T_P|$: number of passing trains, $|T_T|$:number of arriving trains, cont.: continuous, α: percentage of yR variables removed with the variable reduction procedure)

Name	Day/ Night	Disturbance	$\|T_P\|$	$\|T_T\|$	# of train units	# of cont. variables	# of binary variables	# of constraints	α (%)
D1-0	Day	None	243	6	14	87 768	8 400 706	9 150 824	75.8
D2-0	Day	Track closure	241	7	17	91 207	11 871 303	13 066 517	71.3
D3-0	Day	Arrival delays	243	6	15	82 960	10 491 601	13 102 017	75.2
D4-0	Day	None	242	7	18	94 437	13 406 847	14 782 921	71.4
N1-0	Night	None	162	9	23	131 084	14 803 127	17 541 805	66.0
N2-0	Night	Track closure	165	8	19	119 640	10 790 572	13 654 132	65.2
N3-0	Night	Arrival delays	161	10	26	153 742	18 263 150	22 184 403	72.6
N4-0	Night	None	158	10	24	146 055	14 812 464	18 676 208	71.5

Table 4.4 reports the details on the eight real life instances tackled. In each of them there are 7 types of trains on which 4 different operations can be performed: arrival check, internal cleaning, WC cleaning and external cleaning. The track closure scenario reduces the set of possible shunting paths and implies the occurrence of conflicts. Indeed, if a train has to be moved from yard F2 to yard F4, it has to cross main tracks. In instance N3, as trains arrive late in the evening peak hour, their operation cannot start on time. In this scenario, in reality as cleaning crews shift ended too early, some cleaning operations were actually postponed to the morning or cancelled. The number of alternative paths mentioned in Section

Table 4.5: Coefficient of penalties in the G-TUSP objective function

Type of cost	unit	range
departure cancellation	per departing train	135 K - 0.5 M
maintenance operation cancellation	per operation	10K - 13.5K
delay	per second of delay	900 - 1800
number of shunting movements	per movement	15
duration of shunting movements	per second of movement	1
modification to the initial matching	per modification	7
coupling or uncoupling operation	per operation	900

4.3.2 does not exceed 5. In Table 4.4, we report the number of passing trains $|T_P|$ (Column 4), arriving trains $|T_T|$ (Column 5) and train units (Column 6) as well as the number of continuous (Column 7) and binary variables (Column 8) created in the model by reducing the number of disjunction variables yR for routing. Despite the limited set of trains, we get large numbers of variables. In last column, we detail the percentage α of variables yR removed thanks to the reduction procedure presented in Section 4.2.2. For each instance, the number of yR variables is reduced of more than 65%.

As mentioned in Definition 4.2, the G-TUSP objective function includes penalties due to delays, maintenance operations cancelling, coupling and uncoupling operations, modifications of the initial rolling stock rotation, as well as the number and the duration of shunting movements. Coefficients in the objective function reported in Table 4.5 have been obtained with operations experts of Metz-Ville station.

4.4.2 Results

Table 4.6 reports the results obtained on the 8 instances described in Section 4.4.1. It shows the number of coupling and uncoupling required on shunted trains as well as the number of modifications to the planned train matching. It also reports the average number of routes allocated to an intermediate train by our solution and the average number of routes actually allocated by dispatchers. Moreover, we indicate delays taken by departures performed by trains coming from shunting yards. However passing trains departures can also be delayed in addition to shunted trains delays, then the total delay reported in Table 4.6 comes from these two sources. The table also shows the actual total delay recorded on the traffic database. We remark that the solver does not reach an optimal solution or a proof of optimality in the allocated time for any of the instances. The percentage optimality gap exceeds 10 % in both arrival delay scenarios D3-0 and N3-0.

The solution of the MILP has no delay for disturbance free scenarios D1-0, D4-0, N1-0 and N4-0. However, more shunting movements are performed in our solution than in the one implemented by dispatchers. The solution of D2-0 brings a departure delayed as in the actual traffic data. It is in both cases the same departing train, nevertheless it suffers from a 500 seconds delay in our solution while it is 600 seconds in reality. In the total delay, we consider delays of departing trains and passing trains. In the solutions of D2-0, some passing trains are also delayed. This is why we have more than 700 seconds of delay. In instance D3-0, despite a significant gap, the solution obtained reduces the total delay. The solution of N2-0 switches two trains in order to reduce the delay. For N3-0, the result is notably different from the actual decisions. The solution changes three assignments from the planned train matching. Then fewer operations are canceled than in the solution implemented by dispatchers, in which the planned train matching is not changed. In summary, the implementation of our MILP formulation calls the attention on relevant alternatives to the choices of dispatchers in the tested instances. In particular, they highlight the significant effect of changes in the train matching for G-TUSP.

Table 4.6: Experimental results (act. # cancel. op.: number of cancelled operations by rolling-stock managers, modif. match.: modifications to the planned train matching, av. # shunt. paths: average number of shunting paths allocated to intermediate trains by our solution, act. av. # shunt. paths: average number of shunting paths allocated to intermediate train by dispatchers, shunt. dep. del.: shunted departure trains delayed, act. # shunt. dep. del.: number of shunted departure trains delayed by dispatchers solution, tot. shunt. mov. time: total shunting movement time, act. total delay: total delay in dispatchers solution).

Instance	D1-0	D2-0	D3-0	D4-0	N1-0	N2-0	N3-0	N4-0
running time (sec)	3 600	3 600	3 600	3 600	3 600	3 600	3 600	3 600
# cancelled operations	0	0	0	0	0	0	1	0
act. # cancel. op.	**0**	**0**	**0**	**0**	**0**	**0**	**2**	**0**
# coupling	1	2	0	2	1	2	2	2
# uncoupling	2	1	1	2	2	0	3	3
# modif. match.	0	0	0	0	0	2	3	0
av. # shunt. paths	2.5	3.09	2.67	2.7	2.89	3.38	3.10	2.25
act. av. # shunt. paths	**2.17**	**2.43**	**2.33**	**2.60**	**2.56**	**2.75**	**2.40**	**2.25**
# shunt. dep. del.	0	1	0	0	0	0	1	0
act. # shunt. dep. del.	**0**	**1**	**0**	**0**	**0**	**1**	**1**	**0**
tot. shunt. mov. time (s)	9 092	11 032	7 703	12 268	20 867	24 076	19 510	16 389
total delay (s)	0	712	296	0	0	206	871	0
act. total delay (s)	**0**	**750**	**1 167**	**0**	**0**	**480**	**1 462**	**0**
integer solution value	21692	37912	24038	31168	44267	46590	62642	33489
gap (%)	9.7	7.4	11.3	10.3	8.2	12.5	18.6	13.5

4.5 Conclusion

In this chapter, we provided a formal description for the G-TUSP, which is the integrated problem of managing shunting operations planning in passenger trains. It is a large decision problem that includes many specific operational constraints. We presented a MILP formulation for allocations and continuous time scheduling.

The formulation copes with both rolling-stock management and capacity management. It extends some literature approaches which combine the TAP with the TMP. Moreover, we introduced microscopic-scale routing features based on a MILP formulation for real-time traffic management, and maintenance scheduling aspects. Maintenance aspects lead us to consider that the trains can be successively parked on several tracks which is typically not considered in most of the TUSP literature. The proof of concept carried out on the Metz-Ville case-study validates the formulation relevance. Indeed, it confirms the interest of implementing an integrated approach for improving the operating performance of a station. Even if we cannot prove the optimality of the solutions, they are very satisfying compared to the decisions made by dispatchers.

Nonetheless, before being able to deploy an optimization tool for this problem, we need to reduce computational times. A heuristic phase may provide a first integer solution to the MILP solver, which typically has a major impact on performance. We may also reduce the number of variables by reducing further the number of alternative paths to consider. The choice of the remaining paths is in this case critical, and a suitable approach must be found. Other solution techniques such as decomposition or sequential algorithms can be applied. In the next chapter, we will propose such algorithms and assess their performance.

Chapter 5

Sequential algorithms for the generalized train unit shunting problem

As mentioned in Chapter 3 shunting operations planning includes several decisions. First, arriving train units must be assigned to departures, which constitutes a matching decision. This matching must take into account rolling stock features required for departures. Another decision concerns train units location: they must be parked at one or several shunting tracks depending on amenities required by maintenance operations. Similarly, movements are set to achieve the parking locations. For these movements, route planning decisions are to be made, since paths are assigned to train units and movements are scheduled based on running times and potential conflicts. Finally, respecting maintenance crews availability, maintenance operations must be scheduled. Although all these four types of decisions are often made separately, they are usually strongly interdependent. For instance, some matching plans make train units parking or maintenance scheduling impossible.

In Chapter 4, we introduced the G-TUSP, a problem which integrates four shunting problems: the Train Matching Problem (TMP), the Shunting Maintenance Problem (SMP), the Track Assignment Problem (TAP) in its multiple track version (TAP*) and the Shunting Routing Problem (SRP). The mixed integer linear programming formulation gave very satisfying solutions compared to decisions made by dispatchers. Nevertheless, calculation times are quite high. In particular, they are unpractical to allow dispatchers to study different traffic scenarios.

In Chapter 3, we presented a sequential planning approach based on the hierarchy between the TMP, the SMP, the TAP* and the SRP. To obtain good solutions to the G-TUSP in shorter time, processing sequentially the different sub-problems is a natural strategy that we investigate in this chapter. However, as above mentioned, intuitively at least some of these sub-problems would rather be solved in an integrated way. In this study, we assess the importance of the interdependence between sub-problems considering their sequential or integrated solution.

In this chapter, we propose algorithms in which a group of sub-problems is solved exactly while the decision variables related to the other sub-problems are set. Then, we select the best algorithm to solve the overall G-TUSP considering the most appropriate sub-problems integration. We model the G-TUSP on an infrastructure microscopically represented. All the algorithms proposed are tested on several instances which cover different types of disturbances. Some of these instances replicate actually occurred situations, other are artificially generated starting from them. These solution approaches have been presented in the 12th World Congress in Railway Research (WCRR) Kamenga et al. [2019a] and in the 21th congress of the French Operations Research & Decision Support Society Kamenga et al. [2020].

The rest of the chapter is organized as follows. Solutions algorithms are presented in Section 5.1. Section 5.2 reports experiments and Section 5.3 concludes the chapter.

Table 5.1: Schematic representation of the structure of the algorithms proposed. In the first step a solution of one or more sub-problems is found. This solution is passed to step 2, where other sub-problems are solved, including the SRP, and where only one path is considered available for each movement. In step 3, alternative pathss are also considered. When a problem is mentioned in two subsequent steps, the solution of the former is used to initialize the search in the latter.

Name	**Step 1** **Partial solution**		**Step 2** **FIX-Solution**		**Step 3** **VAR-Solution**
S-FM	MILP_{TMP}	$\rightarrow$	$\text{MILP}_{SMP,TAP^*,SRP^1}$	$\rightarrow$	$\text{MILP}_{SMP,TAP^*,SRP}$
S-FMM	$\text{MILP}_{TMP,SMP}$	$\rightarrow$	$\text{MILP}_{TAP^*,SRP^1}$	$\rightarrow$	$\text{MILP}_{TAP^*,SRP}$
S-FMMP	$\left.\begin{array}{c}\text{MILP}_{TMP,SMP}\\ \downarrow \\ \text{Heuristic}_{TAP^*}\end{array}\right\}$	$\rightarrow$	MILP_{SRP^1}	$\rightarrow$	MILP_{SRP}
S-FP	$\text{MILP}_{TMP,SMP}$ $\downarrow$ Heuristic_{TAP^*}	$\rightarrow$	$\text{MILP}_{TMP,SMP,SRP^1}$	$\rightarrow$	$\text{MILP}_{TMP,SMP,SRP}$

5.1 Sequential algorithms for G-TUSP

In this section, we introduce different algorithms based on various combinations of integrated and sequential solution of the G-TUSP sub-problems.

In chapter 4, we proposed a comprehensive MILP formulation which suffers from computational time issues when its exact solution is attempted. In this section, we investigate different possibilities to reduce the computational time necessary to achieve a good quality solution. Remark that, on one hand, the literature shows that many sub-problems are already difficult to solve when they are tackled independently from one another. On the other hand, the interaction between the sub-problems significantly complicates the solution task. The possibilities we propose concretize in sequential algorithms, in which different phases integrate different sub-problems.

In the sequential algorithms, we independently tackle a sub-problem or a group of sub-problems and we use the so obtained solutions as input to the following ones. As discussed in Chapter 3, in sequential order, the TMP is the first problem to be solved in the G-TUSP. Indeed, it is thanks to a TMP solution that the time at which a train must be ready for departure is known. Then, the time spent by every train in the shunting yard can be determined. Moreover, the TMP can be instantiated without the need for a solution to other sub-problems. After the TMP, the SMP can be solved. In particular, compatible crews and shunting tracks are allocated to maintenance operations. In the next step, the TAP* assigns shunting tracks for trains to be parked on when they are not undergoing maintenance operations. Finally, the SRP is solved.

Table 5.1 reports the structure of different sequential algorithms. We consider four algorithms in which we progressively find exact or heuristic solutions to different sub-problems. All algorithms are in three steps. In the first step, a first part of the G-TUSP sub-problems are solved. Therefore a partial solution is found, since a subset of sub-problems is solved. This step is named Partial solution. In this step, depending on the algorithm, the TMP and the SMP are solved with the MILP formulations detailed in Sections 5.1.1 and 5.1.2. The TAP* is solved heuristically, using the algorithm described in Section 5.1.3. In the second step, the missing sub-problems are integrated, but only one path is considered in the SRP to link each origin-destination pair. This step is hence named FIX-Solution. In this step, we use the formulation of Chapter 4 in which we fix the value of the decision variables corresponding to some of the problems solved in the first step. Finally, in the third step, all sub-problems are solved considering all available alternative paths. This step is named VAR-Solution. Also this step uses the formulation of

Chapter 4 in which we fix the same variables as in the second step. The search is initialized with the FIX-solution.

The first algorithm reported in Table 5.1 is named S-FM: shunting with fixed matching. Here, first, the TMP is exactly solved. Then, the three remaining sub-problems are exactly solved altogether considering the matching as given and non-modifiable: the TMP is not solved again in the second and third step. Instead, all other sub-problems are solved in both these steps: the solution found in the former is used as initial solution in the search of the latter.

The second algorithm is S-FMM: shunting with fixed matching and maintenance. Here, the integrated problem composed of the TMP and the SMP is exactly solved, and its solution is used as input for the integrated exact solution of the TAP* and the SRP.

In the third algorithm, S-FMMP for shunting with fixed matching, maintenance and parking, the first step starts as in S-FMM. However, here the solution is used as an input of both a heuristic algorithm for the TAP* and the exact solution of the SRP.

Finally, the fourth algorithm is the S-FP: shunting with fixed parking. Here, as in S-FMMP, we have the exact integrated solution of the TMP and the SMP, and the heuristic one of the TAP*. However, while the TAP* cannot be modified in the last two steps, the TMP and the SMP ones can.

The reason for using sub-problem solutions as non-modifiable inputs for subsequent steps of the algorithms is the attempt of limiting the size of their search space, so that they can be explored more efficiently. Indeed, this efficiency comes at the cost of possibly excluding the overall optimal solution of the G-TUSP from the explored space. Instead, when solutions are only used to initialize the search but can be modified, it is a different approach to try to increase efficiency, this time without excluding the optimum *a priori*. As the number of alternative paths appears to be an element strongly increasing the difficulty of the instances, we limit it thanks to the application of a pre-processing technique presented in Section 4.3.2.

5.1.1 MILP formulation for TMP

We consider the notation of Table 4.1. For a departing train $t \in T_S$ and an intermediate train $t' \in T_I(t)$, the weight $\omega S_{t,t'}$ penalizes changes to the original rolling-stock rotation. As the TMP is solved separately from other G-TUSP subpoblems, we penalize solutions that do not let enough time between the arrival and departure of a train to carry out maintenance operations. Indeed, as operations can be performed on intermediate trains, they may not be ready to depart as soon as they arrive at the shunting yard. Then for a departing train t, $T_I(t)$ may contain some trains that can not be ready at t's departure time, unless an operation is cancelled or the departure delayed. Let $T_I^N(t) \subseteq T_I(t)$, be the subset of intermediate trains that cannot be ready on time for t's departure if all planned operations are executed. $T_I^N(t)$ can be computed by considering the total duration of the operations planned to be carried out on each intermediate train and then deducing its earliest exit time from the shunting yard. The definition of the cost $\omega S_{t,t'}$ for $t \in T_S$, $t \in T_I(t)$ is updated to consider this issue. $\omega S_{t,t'}$ is positive if the assignment does not belong to the initial matching or if $t' \in T_I^N(t)$ and is equal to 0 otherwise. This cost represents the fact that it is preferable to keep the initial assignment if possible: if a precise assignment has been made in the tactical G-TUSP, then we may avoid the violation of macroscopic constraints by keeping it. Indeed, if by doing so major infeasibilities occur in the station under consideration, then changes are allowed, which motivates the relevance of the TMP.

We consider the following MILP formulation for the TMP. We first consider the following binary variables:

- xT_t, with $t \in T_I$, is equal to 1 if t is created and 0 otherwise
- $xS_{t,t'}$, with $t \in T_S, t' \in T_I(t)$, is equal to 1 if t' is assigned to t and 0 otherwise
- qS_t, with $t \in T_S$, is equal to 1 if t is cancelled and 0 otherwise

We also introduce the following integer variables:

- u_t, with $t \in T_T$ gives the number of uncoupling operations on t
- v_t, with $t \in T_S$ gives the number of coupling operations on t

The objective function to minimize integrates several penalties (5.1). First, it takes into account the cost of departure cancellations. The function includes uncoupling and coupling operations cost. Then, penalties for intermediate trains assignment to departing trains are added. The formulation includes inequalities (4.2)-(4.7). As discussed in Section 4.2.2, integrality constraints can be removed on variables qS_t, u_t and v_t.

$$\min \sum_{t \in T_S} B_t \cdot qS_t + \sum_{t \in T_T} Q_C \cdot u_t + \sum_{t \in T_S} Q_H \cdot v_t + \sum_{t \in T_S} \sum_{t' \in T_I(t)} \omega S_{t,t'} xS_{t,t'} \tag{5.1}$$

5.1.2 MILP formulation for SMP and TMP

In this section we present a MILP formulation which integrates SMP and TMP. It uses notation of Table 4.1 as well as variables used in Section 5.1.1. We also provide additional notation:

- P_t is the set of shunting tracks compatible with t, $P_t = \bigcup_{r \in R_t} P^r$.
- mr is the maximum duration of a shunting movement.
- bt is the minimum time that must separate the arrival of a train on a track and the departure of another train which used the same track before.

We introduce non-negative continuous variables:

- $sO_{o,p,hr}$, with $o \in O_t\,(t \in T_I), p \in P^o, hr \in HR^o$, time at which operation o starts on shunting track p with the crew hr;
- dO_t, with $t \in T_I$, time at which intermediate train t ends all its operations;
- D_t, with $t \in T_S$, delay suffered by departing train t when exiting the control area.

Moreover, we introduce binary variables:

- $xO_{o,p,hr}$, with $o \in O_t\,(t \in T_I), p \in P^o, hr \in HR^o$, is equal to 1 if o is carried out on shunting track p by crew hr and 0 otherwise;
- $y_{o,o',hr}$ with $o \in O_t$, $o' \in O_{t'}$ $(t, t' \in T_I,\, t < t')$, $hr \in HR^o \cap HR^{o'}$, is equal to 1 if crew hr performs operation o before operation o' and 0 otherwise;
- $yP_{o,o',p}$, with $o \in O_t$, $o' \in O_{t'}$ $(t, t' \in T_I,\, t < t')$, $p \in P^o \cap P^{o'}$, is equal to 1 if operation o is carried out before o', and they are both carried out on shunting track p, 0 otherwise.

The objective function to minimize includes function (5.1) and adds penalties for departures delay and operations cancellation (5.2). We note that we can have an operation cancellation penalty only if the intermediate train concerned by the operation is actually created.

$$\begin{aligned} \min \sum_{t \in T_S} Q_t D_t + \sum_{t \in T_I, o \in O_t} \omega_o \left(xT_t - \sum_{p \in P^o, hr \in HR^o} xO_{o,p,hr} \right) + \\ \sum_{t \in T_S} B_t qS_t + \sum_{t \in T_T} Q_C u_t + \sum_{t \in T_S} Q_H v_t + \sum_{t \in T_S} \sum_{t' \in T_I(t)} \omega_{t,t'} xS_{t,t'} \end{aligned} \tag{5.2}$$

First of all, all constraints in the TMP formulation (4.2)-(4.7) must be respected. Moreover, any operation carried out on t must use exactly one crew and one shunting track (5.3). The starting time of an operation is set to 0 if it is not assigned to a shunting track (5.4). Operations must start after the concerned train arrives at a shunting track (5.5), if they are performed. Remark that, when solving this problem, we have no information on the precise time that will see the train entering a shunting track. Hence, to be conservative, we consider this time equal to the train's arrival time plus (a_t) the maximum duration of a shunting movement (mr). If uncoupling operations take place on the train, for each of them we also add the minimum uncoupling time (mc) and the time needed for another shunting movement. An operation performed by crew hr must start after the shift start time of hr (5.6) and end before the shift end time (5.7). Note that (5.6) imposes that the starting time of an operation is 0 if it is not assigned to a crew. If operation o' follows operation o, then o' starts after the end of o. We consider the case in which the successive operations are performed on the same track (5.8) and the one in which they are performed on different tracks and a shunting movement is necessary (5.10). Constraints (5.11) specify when intermediate trains end all their performed operations. If these operations end after the departure time of the associated departing train, then the latter is delayed (5.12).

$$\sum_{hr \in HR^o, p \in P^o} xO_{o,p,hr} \leq xT_t \quad \forall t \in T_I, o \in O_t \tag{5.3}$$

$$\sum_{hr \in HR^o} sO_{o,p,hr} \leq M \sum_{hr \in HR^o} xO_{o,p,hr} \quad \forall t \in T_I, o \in O_t, p \in P^o \tag{5.4}$$

$$sO_{o,p,hr} \geq a_t + mr + (mc + mr)u_t - M(1 - xO_{o,p,hr}) \quad \forall t \in T_I, t' \in T_I(t), o \in O_{t'}, p \in P^o, hr \in HR^o \tag{5.5}$$

$$sO_{o,p,hr} \geq sR_{hr} xO_{o,p,hr} \quad \forall t \in T_I, o \in O_t, p \in P^o, hr \in HR^o \tag{5.6}$$

$$sO_{o,p,hr} \leq (eR_{hr} - pR^o) xO_{o,p,hr} \quad \forall t \in T_I, o \in O_t, p \in P^o, hr \in HR^o \tag{5.7}$$

$$sO_{o',p,hr'} \geq sO_{o,p,hr} + pR^o xO_{o,p,hr} - M(1 - xO_{o',p,hr'}) \quad \forall t \in T_I, (o,o') \in E_t, p \in P^o \cap P^{o'}, \tag{5.8}$$

$$hr \in HR^o, hr' \in HR^{o'} \tag{5.9}$$

$$sO_{o',p',hr'} \geq sO_{o,p,hr} + (pR^o + mr) xO_{o,p,hr} - M(1 - xO_{o',p',hr'}) \quad \forall t \in T_I, (o,o') \in E_t, p \in P^o, p' \in P^{o'}, p \neq p', hr \in HR^o, \quad hr' \in HR^{o'} \tag{5.10}$$

$$dO_t \geq sO_{o,p,hr} + pR^o xO_{o,p,hr} \quad \forall t \in T_I, o \in O_t, p \in P^o, hr \in HR^o \tag{5.11}$$

$$D_t \geq dO_{t'} - d_t - M(1 - xS_{t,t'}) \quad \forall t \in T_S, t' \in T_I(t) \tag{5.12}$$

$$\tag{5.13}$$

As two operations can not use a crew at the same time, we set disjunctive constraints (5.14), (5.15).

$$\begin{gathered} sO_{o',p',hr} \geq sO_{o,p,hr} + pR^o xO_{o,p,hr} - M(1 - y_{o,o',hr}) \\ \forall t, t' \in T_I, o \in O_t, o' \in O_{t'}, hr \in HR^o \cap HR^{o'}, p \in P^o, p' \in P^{o'}, t < t' \end{gathered} \tag{5.14}$$

$$\begin{gathered} sO_{o,p,hr} \geq sO_{o',p',hr} + pR^{o'} xO_{o',p',hr} - M y_{o,o',hr} \\ \forall t, t' \in T_I, o \in O_t, o' \in O_{t'}, hr \in HR^o \cap HR^{o'}, p \in P^o, p' \in P^{o'}, t < t' \end{gathered} \tag{5.15}$$

Also, as two operations cannot be performed on the same shunting track concurrently, we set disjunctive constraints (5.16), (5.17).

$$\begin{gathered} sO_{o',p,hr'} \geq sO_{o,p,hr} + (pR^o + bt) \cdot xO_{o,p,hr} - M(1 - yP_{o,o',p}) \\ \forall t, t' \in T_I, o \in O_t, o' \in O_{t'}, p \in P^o \cap P^{o'}, hr \in HR^o, hr \in HR^{o'}, t < t' \end{gathered} \tag{5.16}$$

$$\begin{gathered} sO_{o,p,hr} \geq sO_{o',p,hr'} + (pR^{o'} + bt) \cdot xO_{o',p,hr} - MyP_{o,o',p} \\ \forall t, t' \in T_I, o \in O_t, o' \in O_{t'}, p \in P^o \cap P^{o'}, hr \in HR^o, hr \in HR^{o'}, t < t' \end{gathered} \tag{5.17}$$

5.1.3 Heuristic for the TAP

In two of the algorithms we propose, the TAP* is solved heuristically, taking in input a solution of the integration of TMP and SMP. This solution can be obtained with the MILP formulation in Section 5.1.2, for example.

In the heuristic, initially, a routine generates parking time slots according to the maintenance schedule as well as the coupling and uncoupling operations planned. Many of them can be directly associated to a shunting track, i.e., all those related to an operation for which the track has already been chosen. Others are not associated to any specific shunting track, being them before, between or after maintenance operations. Then a greedy algorithm is applied to assign a shunting track to these latter parking time slots.

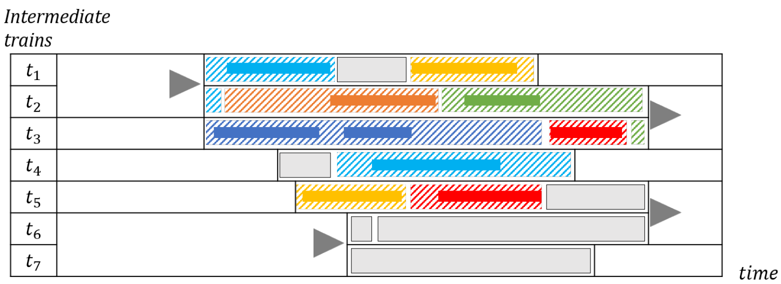


Figure 5.1: Example of parking time slots for a set of intermediate trains. Each row represents an intermediate train. Arrows indicate when intermediate trains enter or leave the shunting yard being coupled. A color corresponds to a shunting track. Full-colored rectangles represent operations scheduled. Dash-coloured rectangles represent the parking time slots created with operations. Gray rectangles represent the unassociated slots created.

Figure 5.1 is used to illustrate the algorithm functioning. Some parking time slots are set thanks to the SMP solution. In the SMP solution, shunting tracks are assigned to maintenance operations. In Figure 5.1, operations schedule appears with full-coloured rectangles. Different colours correspond to different shunting tracks. We assume that an intermediate train arrives to the track where its operations have to be performed as soon as possible. For example, in Figure 5.1, t_2 can move to the green track right after the end of its orange operation, since the track is available. A different situation holds for t_3: it is first on the blue track, where two operations are carried out, and then it moves to the red track as soon as the operation performed there on t_5 is finished. Since no other train needs the blue track we can extend t_3's parking time slot until this movement. It is not always possible for an intermediate train which is maintained to remain on the track assigned for its last operation. For example, t_1 has a first operation on the light blue track and a second one on the yellow track. Before t_1, the yellow track is used for t_5's maintenance. After t_1, the light blue track is used for t_4's maintenance. Then, a new parking time slot is created. This new slot is represented with a gray rectangle, since it concerns a generic shunting track different from those on which maintenance operations take place: it is an *unassociated slot*. It starts as soon as t_1 finishes its operation at the light blue track and ends just before t_1 starts its operation at the yellow track. Another unassociated slot is created for t_4 when it arrives in the shunting yard, since the light blue track is used for t_1's maintenance before being available for t_4 itself.

This algorithm also takes into account coupling and uncoupling operations. Intermediate trains that enter or leave the shunting yard coupled are represented with a gray arrow in Figure 5.1. t_1 and t_2 arrive coupled. As t_1 starts its maintenance before t_2, we suppose that the coupled intermediate trains move directly to the light blue track. Hence, a parking time slot is created for t_2 there. This parking time slot is long enough to allow the uncoupling operation and it ends as soon as possible, in any case before the operation on t_1 starts. As explained in Section 5.1.2, Constraints (5.5) ensure that a maintenance operation starts at the earliest once trains have had time to be moved from platforms to a shunting track, to get uncoupled and to be moved to another shunting track. Similarly, as t_2 has to be coupled to t_3, t_3 joins t_2 on the green track before leaving the shunting yard. Here we follow the same principle as in the rest of the algorithm: t_3 moves as soon as possible. It is the same for t_6 and t_5. Intermediate trains that are not maintained have common slots if they are from a same arrival train like t_6 and t_7. These are unassociated slots.

Intermediate trains are treated in their arrival order in the shunting yard. If many intermediate trains arrive at the same time at the shunting yard, the intermediate train whose first maintenance operation starts the earliest is treated first. If a group of coupled intermediate trains arrive, uncoupling parking time slot are assigned to all intermediate trains but the first one treated.

After computing all parking time slots, we focus on unassociated slots. We consider them following their starting time chronological order. In the example in Figure 5.1, the unassociated slot of t_4 is considered first. For each of these slots, we identify the set of shunting tracks that satisfy three criteria:

1. They must be compatible with the intermediate train, e.g., they must be electrified if the train is electric;

2. They must respect the length constraint, i.e., have a sufficiently long available portion for the whole slot duration;

3. They must respect the crossing constraint, i.e., allow the association of the slot without the occurrence of crossing issues.

Then, a slot is assigned to the the track with the shortest remaining length. The yard layout and the latter criterion may impose an entrance side to the shunting track. When the entrance side is not constrained, we set an entrance side defined in input for each shunting track. We never experienced an infeasibility, a situation in which some slots remain unassociated.

This algorithm is formalized in Algorithm 2.

5.2 Experimental analysis

In this section, we present the results of the assessment of the four proposed algorithms (Table 5.1). The algorithms are coded in Java, and MILP models are solved with the commercial solver CPLEX. The experiments are run with an Intel®Xeon™CPU X5650 2.67 GHz, 12 cores, 24 GB RAM. We study traffic in Metz-Ville station. It is a major hub for Eastern France railway traffic. We tackle real scenarios which include perturbations such as arrival delay or track closure. We also consider scenarios in which we insert fictive perturbations in order to enrich the experimental analysis.

We set running time limits for each algorithm's phase that are mentioned in Table 5.1:

- In the first step, that gives a partial solution (partial solution), the running time is limited to 30 seconds. In S-FM and S-FMM, this is the running time for solving a MILP formulation, while in S-FMMP and S-FP the running time includes both a MILP solution and a heuristic run.

Algorithm 2: Greedy algorithm for the TAP*

```
for t – intermediate train do
    if t needs uncoupling and it is not the first treated in its group then
        create an uncoupling slot for t
    for o – operations of t do
        if track of o is free then
            move t to track of o
        else
            if o is not the first operation and track of o′ predecessor of o is free then
                extend slot of t on track of o′
            else
                create an unassociated slot for t
    if t needs coupling then
        if t is the first treated in its group then
            if track of last operation ô of t is free until departure time then
                extend slot of t on track of ô
                set coupling track of the group to track of ô
            else
                create an unassociated slot for t
                set coupling track of the group to dummy track
        if coupling track of the group of t is free until departure time no intermediate train of the
          group occupies the track then
            move t to coupling track
        else
            if track of last operation ô of t is free until departure time minus coupling time then
                extend slot of t on track of ô
                create a later slot at coupling track
            else
                create an unassociated slot for t until departure time minus coupling time
                create a later slot at coupling track
for unassociated slot do
    determine set of feasible tracks
    select a track with the shortest remaining length in the set
```

- In the second step, where a complete solution without alternative paths (FIX-solution) is sought, the running time is limited to 600 seconds.
- In the last step, where a complete solution considering alternative paths (VAR-solution) is sought, the running time is limited to 900 seconds.

Coefficients in the objective function are reported in Table 4.5.

5.2.1 Case study

We consider traffic in Metz-Ville station and its passengers shunting yards described in Section 4.4.1 and represented in Figure 5.2.

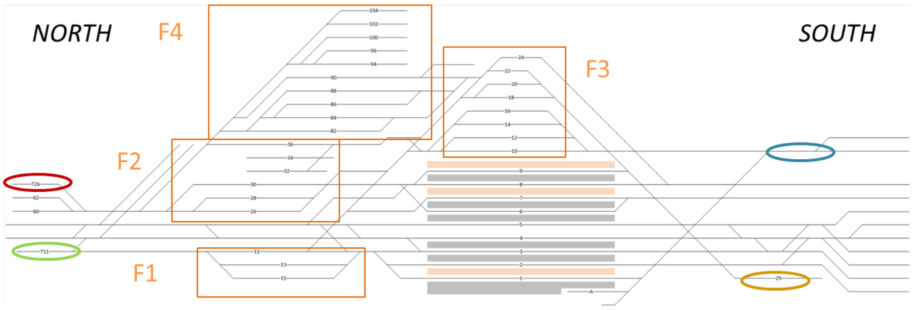


Figure 5.2: Layout of Metz-Ville station: Filled rectangles represent platforms. Yards are orange squared and shunting necks are circled.

We study real-life scenarios defined in Section 4.4.1. Additional scenarios are created by adding fictive perturbation in the real-life ones. First, we generate scenarios in which two trains suffer a two hours arrival delay. These trains are randomly chosen with uniform distribution among all trains requiring shunting. Second, we increase the number of delayed trains to four. Third, we consider a new perturbation, in which track 74 is closed. This is one of the two south side shunting necks, it is circled in blue in Figure 5.2 (south side, up). Therefore, trains have to use shunting neck 29 circled in yellow in Figure 5.2 to perform a turnaround on the south side. When this happens, a high number of shunting movements encounter station traffic.

Table 5.2: Details on the instances tackled in the experimental analysis ($|T_P|$: number of passing trains)

Name	Day /Night	Number of train units	$\|T_P\|$	Original disturbance	Infrastructure disturbance added	Number of delays added
D1-0	Day	14	243	None	None	0
D2-0	Day	17	241	T26 Closure	None	0
D3-0	Day	15	243	2 arrival delays	None	0
D4-0	Day	18	242	None	None	0
N1-0	Night	23	162	None	None	0
N2-0	Night	19	165	T26 Closure	None	0
N3-0	Night	26	161	2 arrival delays	None	0
N4-0	Night	24	158	None	None	0
D1-1	Day	14	243	None	None	2
D2-1	Day	17	241	T26 Closure	None	2
D3-1	Day	15	243	2 arrival delays	None	2
D4-1	Day	18	242	None	None	2
N1-1	Night	23	162	None	None	2

N2-1	Night	19	165	T26 Closure	None	2
N3-1	Night	26	161	2 arrival delays	None	2
N4-1	Night	24	158	None	None	2
D1-2	Day	14	243	None	None	4
D2-2	Day	17	241	T26 Closure	None	4
D3-2	Day	15	243	2 arrival delays	None	4
D4-2	Day	18	242	None	None	4
N1-2	Night	23	162	None	None	4
N2-2	Night	19	165	T26 Closure	None	4
N3-2	Night	26	161	2 arrival delays	None	4
N4-2	Night	24	158	None	None	4
D1-3	Day	14	243	None	Track 74 closure	0
D2-3	Day	17	241	T26 Closure	Track 74 closure	0
D3-3	Day	15	243	2 arrival delays	Track 74 closure	0
D4-3	Day	18	242	None	Track 74 closure	0
N1-3	Night	23	162	None	Track 74 closure	0
N2-3	Night	19	165	T26 Closure	Track 74 closure	0
N3-3	Night	26	161	2 arrival delays	Track 74 closure	0
N4-3	Night	24	158	None	Track 74 closure	0

In summary, 32 instances including 8 real-life scenarios are tackled and the Table 5.2 reports their details:

- 4 instances do not contain perturbations (all real-life scenarios),
- 14 instances only contain delays (2 real-life and 12 fictive scenarios),
- 8 instances contain only track closures (2 real-life and 6 fictive scenarios),
- 6 instances contain both delays and track closures (all fictive scenarios).

Before running the algorithms on each of these instances, we execute the feasibility test described in Section 4.3.3.

5.2.2 Experimental results

In this section, we focus first on computation times to compare our algorithms, then we tackle objective function values.

Table 5.3: Computational times of the algorithms proposed

	Step 1 partial solution (sec)			Step 2 FIX - solution (sec)			Step 3 VAR - solution (sec)		
algorithm	min	mean	max	min	mean	max	min	mean	max
S-FM	0.2	0.6	0.8	110	192.4	362	900	900	900
S-FMM	2.2	3.5	5.1	115	177.1	373	572	870.6	900
S-FMMP	3.1	4.7	6.6	12	23.9	51	86	257.3	715
S-FP	3.1	4.7	6.6	12	25.6	56	153	428.8	900

The first analysis shows the impact of the integration of different sub-problems on the difficulty of the G-TUSP. Computational times are summarized in Table 5.3 for each step in the algorithms. Let us

Table 5.4: Mean values of objective function components (# modif. matching: number of modifications to the planned train matching, # oper. cancel.: number of maintenance operations cancelled, # coupling or uncoupling: sum of the number of coupling and uncoupling operations, # shunt mov.: number of shunting movements, total shunt. mov. time: total shunting movements duration)

algorithm	objective	# modif. matching	total delay (sec)	# oper. cancel.	# coupling or uncoupling	# shunt mov.	total shunt. mov. time (sec)
S-FM	48093	1.4	604	0.22	2.9	19.7	16078
S-FMM	48502	1.7	549	0.19	2.9	20.3	16830
S-FMMP	49559	1.7	707	0.19	2.9	18.8	16972
S-FP	47012	2.2	527	0.19	3	18.8	17008

remark that the maximum computational time plays a role only if CPLEX does not prove the optimality of a solution earlier. If this happens, the corresponding algorithm step is stopped.

The main observation concerns the solution of the TAP*. This problem has a great impact on the difficulty of the G-TUSP. Computational times increase significantly if the TAP* is integrated with other G-TUSP sub-problems and an exact solution for this problem is searched. Recall that S-FM and S-FMM neglect the TAP* in Step 1 and solve it exactly in Steps 2 and 3, integrating it with the SRP and both the SMP and the SRP, respectively. Differently, in S-FMMP and S-FP, Step 1 is the same and it solves the TAP* with a heuristic. The TAP* solution is considered fixed from there on. If we focus on Step 1, the computational time of the last two algorithms does not exceed 6.6 seconds with an average only slightly higher than the one of the first two algorithms (4.1 seconds more than S-FM and 1.2 seconds more than S-FMM). In Step 2, when the TAP* is considered the computational time reaches 362 (S-FM) and 373 (S-FMM) seconds, while it remains lower than one minute otherwise. In Step 3, only S-FMMP always achieves optimality: the computational time is at most 715 seconds. S-FMM, which in this step only differs for the integration of TAP*, only solves 12.5% of the instances to optimality within 900 seconds. The latter algorithm exits the search with an average optimality gap of 4.5% and a maximum of 7.3%. Although reaching in some cases the time limit, S-FP solves 93.9% of the instances to optimality in step 3 with an average gap of 0.3% and a maximum gap of 0.4%. Finally, S-FM never manages to close the gap, exiting with a value of 4.5% in average, and getting to a maximum of 9.5%. Indeed, this algorithm re-assess the highest number of sub-problems in Step 3, including the TAP*. A quantitative measure of the impact of the integration of this problem is also given by the number of binary variables included in the MILP formulations of Step 3 of the different algorithms: they are 1838K and 1831K for S-FM and S-FMM, respectively, while only 49K and 162K for S-FMMP and S-FP.

Focusing on the other sub-problems, we observe that the computational time to solve Step 1 in S-FM, and hence to solve the TMP, does not exceed 0.8 seconds. Indeed, As mentioned by Freling et al. [2005], the TMP alone is rather easy to handle. The same holds when this problem is integrated with the SMP: solving the MILP that constitutes Step 1 of S-FMM takes between 2.2 and 5.1 seconds. Apparently, integrating the SMP has a minor impact on the difficulty of the problem. Indeed, in Step 2, when solving to optimality the TAP and the SRP (S-FMM) or the TAP, the SRP and the SMP (S-FM) the average computation times differ by less than 10%. In Step 3, where alternative paths are allowed, S-FM has an average computation only slightly higher than S-FMM.

After this analysis on difficulty, we focus on the impact of integrating G-TUSP sub-problems on solution quality. We discuss the quality of solutions according to objective function components, first, and scenario types, second.

Table 5.4 reports the mean values of the G-TUSP objective function and some of its components returned by the four algorithms as their final solution, i.e., after Step 3. Departure cancellations are not mentioned in Table 5.4 since no departure is cancelled in the solutions. S-FP gives solutions with the best average objective function, while S-FMMP gives the worst. This is not surprising since, in Steps 2 and 3, the latter considers fixed the solutions of all sub-problems but the SRP, while S-FP only does so for the TAP*. Indeed, here the specific solution of the TAP* does not really make a difference, as S-FM and S-FMM solutions are in average worse than S-FP and better than S-FMMP.

If we look at the number of modifications to the planned train matching (third column of Table 5.4), we see that S-FM is the algorithm that modifies the planned train matching the least, while S-FP modifies it the most. Indeed, S-FM, S-FMM and S-FMMP only modify the train matching in Step 1, integrating the TMP at most with the SMP. As it can be expected, the higher the number of sub-problems the TMP is integrated with, the higher the number of modifications.

The average total delay and number of cancelled maintenance operations are reported in columns four and five of Table 5.4. S-FMMP gives solutions with the longest delay, and in particular with significantly longer delay than S-FMM. In both algorithms, a MILP integrates the TMP and the SMP in Step 1, which then returns the same solution for both algorithms. This solution is not re-assessed in the following steps. Hence, maintenance operations have the same schedule in the final solution of both algorithms. The higher delay in S-FMMP is due to shunting movements between the shunting tracks and the platforms. Indeed, in S-FMMP, as the TAP* is heuristically solved in Step 1, the SRP solved in Steps 2 and 3 has a limited alternatives to find good solutions: the shunting track where the trains' paths begin and finish have to be consistent with the TAP* solution. In S-FMM, the TAP* and the SRP are integrated in Steps 2 and 3: a larger set of possible paths is available and traffic conflicts can be avoided. Although the TAP* is heuristically solved in Step 1 of S-FP, too, this algorithm gives the solutions with the shortest delay in average. Indeed, this algorithm profits from the solution of the TMP in Steps 2 and 3, together with the SMP and the SRP. Therefore, it solves a trade-off between total delay, number of modifications to the initial matching and number of coupling or uncoupling operations, which allows the reduction of delay. S-FM achieves the worst results in terms of number of maintenance operations cancelled and quite bad results in terms of total delay. It is the only algorithm in which only the TMP is solved in Step 1. The train matching is then considered fixed in the following. As mentioned in Section 5.1.1, the MILP formulation for the TMP used by S-FM includes a penalty if the solution does not let enough time for maintenance operations. However, the penalty simply considers operations duration and does not take into account crews or tracks availability, which instead may have an impact on solutions. The average number of operations that need to be canceled to fit the train matching is then slightly higher than for the other algorithms. The solutions found by S-FMM, S-FMMP and S-FP have the same number of maintenance operations cancellations for all instances. While for most of the instances S-FM finds the same numbers, it cancels an additional operation for one instance. It is a daytime instance (D3-2) in which many arrivals are delayed. In this case, S-FM gives a solution which follows the initial matching, but there are not enough crews to carry out internal cleaning and ensure the on-time departures. Therefore, one internal cleaning operation is cancelled to avoid a long delay. Differently, S-FMM, S-FMMP and S-FP change the initial matching so that no operation is cancelled.

Looking at the shunting movements, whose number and duration are reported in the last two columns of Table 5.4, S-FMMP and S-FP provide the solutions with fewer shunting movements. These algorithms return solutions with the same number of shunting movements. Indeed, the number of shunting movements is set with a TAP* solution: the more shunting tracks an intermediate train is parked on the more shunting movements are performed. In S-FMMP and S-FP, the TAP* solution is obtained with a heuristic in Step 1. The greedy algorithm for the TAP* favors the minimization of the number of shunting movements. On the contrary, once the SRP is solved taking as input a TAP* solution, the total duration of shunting movements is higher than in algorithms where the TAP* and the SRP are integrated (S-FM and S-FMM).

Table 5.5 contains mean objective function values for different types of scenario.

In the perturbation-free instances, S-FM and S-FMM provide slightly better solutions than S-FMMP and S-FP. Nevertheless, all the algorithms provide solutions without delays. The difference is due to a higher duration of shunting movements, that increases the cost of S-FMMP and S-FP solutions.

In the scenarios with arrival train delays, S-FP gives particularly good results compared to the other algorithms. Its final delays are almost twice as low as the other on average. Moreover, S-FM gives the worst solutions in average. This attests the benefit of integrating the TMP and the SMP in these

Table 5.5: Mean values of objective function for different types of scenarios.

algorithm	no perturbation scenarios	delay scenarios	track closure scenarios	delay and track closure scenarios
S-FM	32654	45333	49924	62385
S-FMM	32654	43499	54515	62722
S-FMMP	33202	43986	57227	63243
S-FP	33202	41413	56984	55990

Table 5.6: Shunting movements in track closure scenarios

algorithm	mean number of shunting movements	mean total duration of shunting movements (sec)
S-FM	21.07	17887
S-FMM	22.36	19649
S-FMMP	19.14	19395
S-FP	19.14	19668

scenarios. If a train matching is set, then SMP solutions are often of low quality in case of delays. Indeed, the train matching set in S-FM is different from the one in S-FMM. The latter algorithm provides better solutions than S-FMMP in delay scenarios. As the two algorithms use the same solution for the TMP and the SMP, found by a MILP in Step 1, this observation highlights the benefit of integrating the TAP* and the SRP.

Conversely, in track closure scenarios, S-FM provides the best solutions on average. The reason why S-FMM and S-FMMP are less successful may be linked to the SMP solution. Indeed, setting the maintenance schedule might be an issue, since it limits the set of alternative paths for a train. During track-closure periods, alternative paths are crucial to avoid potential traffic conflicts. Solving the TAP* in Step 1 can be detrimental for the same reason. Indeed, S-FMMP gives worse solutions than S-FMM. S-FP manages to partially compensate the early solution of the TAP* by re-assessing the solutions of the other problems. However, this does not allow the complete recovery of solution quality. In Table 5.6, we propose additional details for these scenarios, concerning the number and the duration of shunting movements. We can observe that S-FMM generates the highest number of shunting movements. Through them, the algorithm avoids conflicts with passing trains, which otherwise would cause departure delays. Conversely S-FP and S-FMMP give solutions that have fewer shunting movements and longer delays.

In scenarios that combine delays and track closures, S-FP provides more better solutions than other algorithms. In particular, they have shortest total delay.

For the eight real-life instances available, we can compare the solutions of our algorithms with the one implemented by yard planners. The available traffic data contains delays, operations cancelled and number of modifications to the initial matching. All our algorithms obtain better results in terms of delay except for a specific instance in which S-FM and S-FMM have a total delay 10 seconds longer than the realized one, which is 200 seconds. However in this same instance, no operation is cancelled by our algorithms. Our solutions have been considered efficient by experts of Metz-Ville station.

In summary, we can conclude that there is no algorithm that always outperforms the others, although integrating the TAP* to other problems significantly increases the problem difficulty. However, this increase does not always imply solution quality worsening in the computational time limit considered. Globally, we think that S-FP can be considered the best algorithm, as it achieves the best objective function values overall (Table 5.4) and in two out of four types of scenarios. In the two remaining types, its average objective function value is 2% and 14% worse than its best competitor S-FM. Instead, when S-FM is not the best, the difference with respect to S-FP is 9% and 10% in two types of scenarios (Table 5.5).

5.3 Conclusion

In this chapter, we proposed four solution algorithms for the G-TUSP, based on the sequential or integrated solution of different groups of sub-problems. We assess their performance on a real case study, observing their computational times and the quality of their solutions. Our experiments show that the TAP* is sub-problem that mostly complicates the G-TUSP. Once this sub-problem tackled, the others can be solved quite easily. However, to successfully solve the TAP*, appropriate solutions of the TMP and the SMP must be provided. Indeed, different instance characteristics may imply different relative performance of the algorithms proposed. In the chapter, we assessed these performance when various types of perturbations occur.

Chapter 6

Conclusion

This thesis proposes combinatorial optimization models and algorithms to solve the problem of shunting passenger trains in stations. Integrated approaches are considered to plan coupling, uncoupling, parking, maintenance and movements of train units. As mentioned in Chapter 1, we tackle both service quality and capacity consumption issues. We also highlight the practical relevance of our approaches. In this chapter we overview the main issues tackled in the thesis. In Section 6.1, we summarize the content of this thesis and we draw conclusions. Then, in Section 6.2, we report some directions for future research that emerge from the analysis of our work.

6.1 Summary and conclusions

We organized this manuscript is six chapters, including this concluding one.

In Chapter 1, we introduced our research work. In particular, we described its context and motivation, and we set its objectives. The main objective (O) is the proposal of a model and algorithm for scheduling shunting operations. It is declined in two sub-objectives: (O1) the algorithms must supply solutions at least as good as the ones currently implemented, and (O2) they must do so in reasonable time for real life instances.

In Chapter 2, we focused on the description of the railway system, considering its aspect relevant for this thesis. Namely, we first took into account the main resources that allow the execution of rail services: infrastructure capacity consumption and rolling stock management, including aspects as maintenance needs, and coupling and uncoupling of train units. Then, we moved to the passenger railway service planning process. This process is mostly resource centered and made at different time horizons. When narrowing the perspective to shunting problems, we observed how their definition is centered on train units preparation for train operating companies, and on station capacity for the infrastructure manager. However, the common need for efficiency increase is currently pushing all actors to aim to common objectives including the minimization of: departure cancellations, delays, shunting operations as train unit uncoupling and coupling, maintenance operations cancellation, . . . In France, shunting is planned over a short-term horizon which does not exceed one month. It is considered after the rolling stock circulation and is strongly related to platforming. In shunting, the use of rolling stock, the use of infrastructure and the use of crews are strictly interlinked.

In Chapter 3, we formally described the shunting problems we consider in the thesis. In particular, we identified and discussed four problems: the Train Matching Problem (TMP), the Track Assignment Problem (TAP), the Shunting Maintenance Problem (SMP) and the Shunting Routing Problem (SRP). First, we discussed these problems as well as their interdependence. Then, we analyzed the literature on transportation problems related to one or more of them. Indeed, the literature deals with a specific sub-problem or few of them, solved either sequentially or with integrated algorithms.

In Chapter 4, we formalized the integration of the four problems identified above, and proposed the

Generalized Train Unit Scheduling Problem (G-TUSP). In addition, this problem includes a large set of features that are present in real life but so far neglected in the literature. For example, it includes maintenance crews scheduling and protection constraints, that guarantee the safety of these crews when working on a train unit. No existing approach in the literature considers such an comprehensive problem, that must however be daily solved by planners.

For the G-TUSP, we designed a MILP formulation, considering an extremely detailed representation of all problem aspects. For example, a microscopic representation of the infrastructure is modeled to define train routing. Moreover, we proposed techniques to reduce the size of instances such as the removal of redundant binary variables used for ensuring train unit's separation during movements. We tested the solution of this formulation on real instances representing traffic in Metz-Ville station, in France, and we obtained satisfying solutions within one hour of computation. We compared these solutions with the ones implemented by planners, and we showed that our optimization outperforms the current practice under several criteria. In case of planned disturbances, such as works or delays, the use of the formulation was clearly relevant. Indeed, service quality indicators such as delays and canceled operations were improved. Nevertheless, operating costs we obtained were sometimes higher than in the implemented solutions. However, a service quality is considered the priority in this process, our solutions were appreciated by planners. Indeed, thanks to our comprehensive model, different solutions of the trade-off between costs and service quality can be considered.

This analysis allows concluding the fulfillment of objective (O1) as well as in general objective (O). However, as for the computation time necessary to find good solutions to real life instances, the results presented in this chapter were not completely satisfactory. Hence, we subsequently focused on the second objective of the thesis.

In Chapter 5, we introduced algorithms to solve the G-TUSP considering its sub-problems either sequentially or in an integrated way. In particular, we designed and developed four algorithms, different in the subsets of sub-problems considered in their subsequent steps. We assessed the performance of these algorithms in a thorough experimental analysis, including real and artificial instances representing traffic in Metz-Ville station. This analysis showed the achievement of objective (O2): all algorithms are suitable to tackled the G-TUSP in the preoperational planning phase which we consider in this thesis, going from six days to few hours before operations. In particular, even in the difficult artificial instances designed to challenge the algorithms, good solutions were returned in about 20 minutes in the very worst cases.

The thorough experimental analysis performed also allows the analysis of the impact of the approach used for different sub-problems and of their integration. In particular, the TAP stands out for its complicating role. If it is included in a formulation aimed to be solved exactly, the computation times grow significantly. However, in terms of final objective function value, and hence in terms of costs and service quality, all the combinations of sub-problem integrations we tested were successful. Although the performance of the proposed algorithms were not impressively different, the best algorithm appears to be S-FP: shunting with fixed parking. Here, first a MILP formulation integrating the TMP and the SMP is solved. Then, a heuristic for the TAP finds a track assignment to park train units consistently with the matching and maintenance schedule previously found. Finally, keeping this track assignment unchanged, matching and maintenance schedule are re-optimized through a MILP solution first with pre-affected routes for all train units, and then allowing route flexibility.

6.2 Future research

The studies carried out during this thesis raise new questions which may be the subject of future research directions. In this section we mention directions that cover three issues: performance improvement, quality of solutions, extension to real time management.

In this thesis, several approaches are proposed to achieved good performance in the G-TUSP solution: reducing instances sizes, boosting a MILP formulation and considering sequential algorithms. Possibly, the performance can be further improved by exploring three research directions: further reduction of instance size, alternative solution approaches, alternative modeling approaches.

We remark that the number of paths per train strongly affects the size of instances that solve the TAP and the SRP. A reduction of the number of paths per train can reduce significantly the size of instances. In the thesis, we propose a selection of shunting tracks and a basic approach based on shortest path search in order to reduce the number of alternative paths between two points. However, such reductions can remove paths that belong to an optimal solution. The selection of alternative routes between two points without eliminating optimal solutions is the problem formulated and tacked with a metaheuristic by Sama et al. [2016]. The approach is implemented for the problem of minimizing total delay propagation when optimizing train routing and scheduling in real-time in case of perturbation. A such approach may have to be adapted to the shunting problem and to consider an objective function that including several very different aspects. From a similar perspective, we also think that an algorithm that pre-selects shunting tracks where train units can be parked may be pertinent.

Other algorithms based on exact approaches, heuristics or metaheuristics may be investigated. Decomposition methods in mathematical programming may deal with the overall G-TUSP. For example, Lagrangian relaxation may be studied in order to tackle difficult constraints in the TAP. Benders decomposition and column generation may also be considered to tackle large size instances. A local search based metaheuristic may also be considered to solve the G-TUSP. As mentioned in our literature review, thanks to local search and simulated annealing, Van den Broek [2016] obtains convincing results for a similar problem in which no simultaneous shunting movements is allowed. Indeed, a such approach will have to be extended to assess its suitability to the microscopic scale routing and scheduling of shunting movements, when simultaneous movements are possible.

A further possibility is the design of a meta-algorithm, including different algorithms to be executed in different situations. The main issue with the design of a such an approach is the definition of the triggers associated to each algorithm. A deep understanding of the problem is necessary to do so. Alternatively, a machine learning mechanism may be implemented for the meta-algorithm to learn these triggers as it is deployed. Indeed, our experimental analysis may be considered as a sort of first training set. In particular, it allowed identifying some relations between the type of perturbation occurring and the most suitable sub-problem integration. For example, we observed that integrating the TMP and the SMP is very beneficial in case of delays and that integrating the TAP and the SRP is relevant in case of track closures. Of course, a very large experimental testbed should be obtained to set up a reasonable machine learning approach. However, we think that it may be possible to acquire it while, for example, deploying a single algorithm and running in the background the alternative ones for some time.

Significantly changing perspective, discrete time models may be studied. A discrete time model can be particularly relevant to model shunting yard operations, where a rather large time step, probably in the order of the minute, can be set to handle crews and track utilization. Instead, on main tracks a lower time step or a continuous model must be considered to deal with conflicts between shunting and regular movements. Then methods that ensure consistency between scheduling at shunting tracks and scheduling at main tracks should be found.

From a different point of view, the quality of solutions of integrated shunting problems involves two issues that may be investigated in future research: multiobjective optimization and robustness.

In the thesis, the tradeoff between service quality or operational cost criteria is considered with a weighted sum. Nevertheless weighted sum can only represent preferences based on compensation and do not consider all efficient solutions Ehrgott and Gandibleux [2000]. Alternative approaches in multiobjective optimization can be studied. Preference modelling and elicitation should be considered.

Moreover, solutions that have the same objective function value may not be equivalent in practice. This is due to uncertainty of parameters in our models. for planners. Indeed, an additional issue that may be considered is robustness: solutions capable to cope with uncertainty undoubtedly preferable. In, a first approach we require that solutions are feasible in all scenarios of a certain restricted disturbances scenario set. In a second approach, recovery robustness (Liebchen et al. [2009], Cicerone et al. [2009])

may be considered. In recovery robustness,we consider also means or recovery possibilities to modify the plan which can be made feasible in the limited set of scenarios. These means or recovery effort is limited. In this approach, a solution is feasible if a recovery possibilities exist for every scenario studied.

We can investigate extensions of our results to real time shunting operations management. In case of disturbances, rerouting or rescheduling may be necessary. The G-TUSP solution can be used in a loop control of shunting operations with a rolling horizon. The horizon duration should be chosen wisely. Indeed, this horizon determines the set of possible modifications. It may start few minutes after a disturbance and finish once traffic is recovered. As decisions must be taken in a very short time, heuristics and metaheuristics can be suitable solution approaches.

Bibliography

Abbink, E. (2006). *Intelligent Shunting: Dealing with Constraints (Satisfaction)*, pages 391–413. John Wiley & Sons, Inc.

Alfieri, A., Groot, R., Kroon, L., and Schrijver, A. (2006). Efficient circulation of railway rolling stock. *Transportation Science*, 40(3):378–391.

Artigues, C., Bourreau, E., Jost, V., Kedad-Sidhoum, S., and Ramond, F. (2018). Trains do not vanish: the roadef/euro challenge 2014. *Annals of Operations Research*, 271(2):1091–1105.

Artigues, C., Gendreau, M., Rousseau, L.-M., and Vergnaud, A. (2009). Solving an integrated employee timetabling and job-shop scheduling problem via hybrid branch-and-bound. *Computers & Operations Research*, 36(8):2330 – 2340. Constraint Programming.

Benhizia, F. (2012). *Optimisation du plan de transport par planification intégrée des ressources*. PhD thesis, Saint-Etienne, EMSE.

Blasum, U., Bussieck, M. R., Hochstättler, W., Moll, C., Scheel, H.-H., and Winter, T. (1999). Scheduling trams in the morning. *Mathematical Methods of Operations Research*, 49(1):137–148.

Bodlaender, H. L., Kloks, T., and Kratsch, D. (1995). Treewidth and pathwidth of permutation graphs. *SIAM Journal on Discrete Mathematics*, 8(4):606–616.

Brucker, P. and Schlie, R. (1990). Job-shop scheduling with multi-purpose machines. *Computing*, 45(4):369–375.

Cadarso, L. and Marín, A. (2011). Robust rolling stock in rapid transit networks. *Computers & Operations Research*, 38(8):1131 – 1142.

Cadarso, L. and Marín, Á. (2012). Integration of timetable planning and rolling stock in rapid transit networks. *Annals of Operations Research*, 199(1):113–135.

Caprara, A., Kroon, L., Monaci, M., Peeters, M., and Toth, P. (2007). Passenger railway optimization. *Handbooks in operations research and management science*, 14:129–187.

Cardonha, C. and Borndörfer, R. (2009). A set partitioning approach to shunting. *Electronic Notes in Discrete Mathematics*, 35(Supplement C):359 – 364. LAGOS'09 – V Latin-American Algorithms, Graphs and Optimization Symposium.

Cemal, Ö., Lale, Ö., and Yasemin, Y. (2010). Mathematical models for job-shop scheduling problems with routing and process plan flexibility. *Applied Mathematical Modelling*, 34(6):1539 – 1548.

Cicerone, S., D'Angelo, G., Di Stefano, G., Frigioni, D., Navarra, A., Schachtebeck, M., and Schöbel, A. (2009). Recoverable robustness in shunting and timetabling. In *Robust and online large-scale optimization*, pages 28–60. Springer.

Coffman Jr, E., Garey, M., and Johnson, D. (1996). Approximation algorithms for bin packing: a survey. In *Approximation algorithms for NP-hard problems*, pages 46–93. PWS Publishing Co.

Comité Européen de Normalisation (2002). *Transportation. Logistics and services. Public passenger transport. Service quality definition, targeting and measurement - EN 13816:2002.* Comité Européen de Normalisation.

Cornelsen, S. and Di Stefano, G. (2007). Track assignment. *Journal of Discrete Algorithms*, 5(2):250 – 261. 2004 Symposium on String Processing and Information Retrieval.

Dahlhaus, E., Horak, P., Miller, M., and Ryan, J. F. (2000). The train marshalling problem. *Discrete Applied Mathematics*, 103(1):41 – 54.

Dauzère-Pérès, S. and Paulli, J. (1997). An integrated approach for modeling and solving the general multiprocessor job-shop scheduling problem using tabu search. *Annals of Operations Research*, 70:281–306.

Dauzère-Pérès, S., Roux, W., and Lasserre, J. B. (1998). Multi-resource shop scheduling with resource flexibility. *European Journal of Operational Research*, 107(2):289–305.

Demange, M., Di Stefano, G., and Leroy-Beaulieu, B. (2012). On the online track assignment problem. *Discrete Applied Mathematics*, 160(7):1072 – 1093.

Di Stefano, G. e. and Koči, M. L. (2004). A graph theoretical approach to the shunting problem. *Electronic Notes in Theoretical Computer Science*, 92:16–33.

Eggermont, C., Hurkens, C. A., Modelski, M., and Woeginger, G. J. (2009). The hardness of train rearrangements. *Operations Research Letters*, 37(2):80 – 82.

Ehrgott, M. and Gandibleux, X. (2000). A survey and annotated bibliography of multiobjective combinatorial optimization. *OR-Spektrum*, 22(4):425–460.

Fioole, P.-J., Kroon, L., Maróti, G., and Schrijver, A. (2006). A rolling stock circulation model for combining and splitting of passenger trains. *European Journal of Operational Research*, 174(2):1281–1297.

Freling, R., Lentink, R. M., Kroon, L. G., and Huisman, D. (2005). Shunting of passenger train units in a railway station. *Transportation Science*, 39(2):261–272.

Gallo, G. and Miele, F. D. (2001). Dispatching buses in parking depots. *Transportation Science*, 35(3):322–330.

Gatto, M., Maue, J., Mihalák, M., and Widmayer, P. (2009). *Shunting for Dummies: An Introductory Algorithmic Survey*, pages 310–337. Springer Berlin Heidelberg, Berlin, Heidelberg.

Gilg, B., Klug, T., Martienssen, R., Paat, J., Schlechte, T., Schulz, C., Seymen, S., and Tesch, A. (2018). Conflict-free railway track assignment at depots. *Journal of rail transport planning & management*, 8(1):16–28.

Guyon, O., Lemaire, P., Pinson, E., and Rivreau, D. (2014). Solving an integrated job-shop problem with human resource constraints. *Annals of Operations Research*, 213(1):147–171.

Haahr, J. and Lusby, R. M. (2017). Integrating rolling stock scheduling with train unit shunting. *European Journal of Operational Research*, 259(2):452–468.

Haahr, J. T. and Bull, S. H. (2015). Exact methods for solving the train departure matching problem. Technical report, DTU Management Engineering.

Haahr, J. T., Lusby, R. M., and Wagenaar, J. C. (2017). Optimization methods for the train unit shunting problem. *European Journal of Operational Research*, 262(3):981–995.

Haijema, R., Duin, C., and Van Dijk, N. (2006). Train shunting: A practical heuristic inspired by dynamic programming. *Planning in Intelligent Systems: Aspects, Motivations, and Methods*, pages 437–475.

Hamdouni, M., Desaulniers, G., Marcotte, O., Soumis, F., and Van Putten, M. (2006). Dispatching buses in a depot using block patterns. *Transportation Science*, 40(3):364–377.

Hamdouni, M., Desaulniers, G., and Soumis, F. (2007). Parking buses in a depot using block patterns: A benders decomposition approach for minimizing type mismatches. *Computers & Operations Research*, 34(11):3362 – 3379.

Huisman, D., Kroon, L. G., Lentink, R. M., and Vromans, M. J. (2005). Operations research in passenger railway transportation. *Statistica Neerlandica*, 59(4):467–497.

Jacob, R., Márton, P., Maue, J., and Nunkesser, M. (2011). Multistage methods for freight train classification. *Networks*, 57(1):87–105.

Jacobsen, P. M. and Pisinger, D. (2011). Train shunting at a workshop area. *Flexible services and manufacturing journal*, 23(2):156–180.

Kamenga, F. (2016). Gestion de l'occupation de voie des gares en temps réel. Master's thesis, Université Paris Dauphine – Mines ParisTech.

Kamenga, F., Pellegrini, P., Rodriguez, J., Merabet, B., and Houzel, B. (2019a). Integrating train unit shunting problem, train maintenance and routing. In *12th Word Conference on Railway Research - WCRR 2019, Tokyo, Japan.*

Kamenga, F., Pellegrini, P., Rodriguez, J., Merabet, B., and Houzel, B. (2019b). Train unit shunting : Integrating rolling stock maintenance and capacity management in passenger railway stations. In *8th International Conference on Railway Operations Modelling and Analysis - RailNorrköping 2019, Norrköping, Sweden*, pages 721–740.

Kamenga, F., Pellegrini, P., Rodriguez, J., Merabet, B., and Houzel, B. (2019c). Train unit shunting : Integrating rolling stock maintenance and capacity management in passenger railway stations. In *20th Congress of the French Operations Research & Decision Support Society - ROADEF 2019, Le Havre, France.*

Kamenga, F., Pellegrini, P., Rodriguez, J., Merabet, B., and Houzel, B. (2020). Sequential approaches for solving shunting problems at passenger railway stations. In *21st Congress of the French Operations Research & Decision Support Society - ROADEF 2020, Montpellier, France.*

Kroon, L., Huisman, D., Abbink, E., Fioole, P.-J., Fischetti, M., Maróti, G., Schrijver, A., Steenbeek, A., and Ybema, R. (2009). The new dutch timetable: The or revolution. *Interfaces*, 39(1):6–17.

Kroon, L. G., Lentink, R. M., and Schrijver, A. (2008). Shunting of passenger train units: an integrated approach. *Transportation Science*, 42(4):436–449.

Lentink, R. (2006). *Algorithmic decision support for shunt planning.* PhD thesis, Erasmus Universiteit Rotterdam.

Lentink, R. M., Fioole, P.-J., Kroon, L. G., and Van't Woudt, C. (2006). Applying operations research techniques to planning of train shunting. *Planning in Intelligent Systems: Aspects, Motivations, and Methods*, pages 415–436.

Li, H., Jin, M., He, S., Ye, Z., and Song, J. (2017). Optimal track utilization in electric multiple unit maintenance depots. *Computers & Industrial Engineering*, 108(Supplement C):81 – 87.

Liebchen, C., Lübbecke, M., Möhring, R., and Stiller, S. (2009). The concept of recoverable robustness, linear programming recovery, and railway applications. In *Robust and online large-scale optimization*, pages 1–27. Springer.

Liebchen, C. and Möhring, R. H. (2007). The modeling power of the periodic event scheduling problem: railway timetables—and beyond. In *Algorithmic methods for railway optimization*, pages 3–40. Springer.

Lübbecke, M. E. and Zimmermann, U. T. (2005). Shunting minimal rail car allocation. *Computational Optimization and applications*, 31(3):295–308.

Martello, S. and Toth, P. (1990). Lower bounds and reduction procedures for the bin packing problem. *Discrete applied mathematics*, 28(1):59–70.

Martin, A., Antoine, M., Chalmeau, O., Couly, F., Gourbal, M., Lai, A., Le Potier, A., Périgord, A., and Vialette, T. (2019). *Le marché français du transport ferroviaire de voyageurs: 2018 volume 1*. Autorité de régulation des transports.

Mastrolilli, M. and Gambardella, L. M. (2000). Effective neighbourhood functions for the flexible job shop problem. *Journal of Scheduling*, 3(1):3–20.

Pellegrini, P., Marlière, G., Pesenti, R., and Rodriguez, J. (2015). RECIFE-MILP: An effective milp-based heuristic for the real-time railway traffic management problem. *IEEE Transactions on Intelligent Transportation Systems*, 16(5):2609–2619.

Pellegrini, P., Pesenti, R., and Rodriguez, J. (2019). Efficient train re-routing and rescheduling: Valid inequalities and reformulation of recife-milp. *Transportation Research Part B: Methodological*, 120:33 – 48.

Qi, X., Yue, Y., Han, J., Zhou, L., and Rakha, H. A. (2017). Integrated modelling and optimization of train scheduling and shunting at complex railway passenger stations. In *Transportation Research Board 96th Annual Meeting*.

Ramond, F. (2008). *Planification optimisée des opérations dans les établissements de maintenance du matériel roulant de la SNCF*. PhD thesis, Saint-Etienne, EMSE.

Ramond, F. and Nicolas, M. (2014). Trains don't vanish! roadef euro 2014 challenge problem description. In *EURO XXVI Conference*.

Rétiveau, R. (1987). *La signalisation ferroviaire*. Presse de l'école nationale des Ponts et Chaussées.

Riezebos, J. and Van Wezel, W. (2009). k-shortest routing of trains on shunting yards. *OR spectrum*, 31(4):745.

Sama, M., Pellegrini, P., D'Ariano, A., Rodriguez, J., and Pacciarelli, D. (2016). Ant colony optimization for the real-time train routing selection problem. *Transportation Research Part B: Methodological*, 85:89–108.

Schöbel, A. (2017). An eigenmodel for iterative line planning, timetabling and vehicle scheduling in public transportation. *Transportation Research Part C: Emerging Technologies*, 74:348–365.

SNCF Réseau (2018). *Network statement of the national rail network - 2020 timetable*. SNCF Réseau.

Thorlacius, P., Larsen, J., and Laumanns, M. (2015). An integrated rolling stock planning model for the copenhagen suburban passenger railway. *Journal of Rail Transport Planning & Management*, 5(4):240 – 262.

Tomii, N. and Zhou, L. J. (2000). Depot shunting scheduling using combined genetic algorithm and pert. *Computers in Railways*, 7:437–446.

UIC (2013). *Capacity*. International Union of Railways.

Van Den Broek, J. and Kroon, L. (2007). A capacity test for shunting movements. In *Algorithmic Methods for Railway Optimization*, pages 108–125. Springer.

Van den Broek, R. W. (2016). Train shunting and service scheduling: an integrated local search approach. Master's thesis, Universiteit Utrecht.

Van Wezel, W. and Riezebos, J. (2010). Case study: Advanced decision support for train shunting scheduling. In *Behavioral Operations in Planning and Scheduling*, pages 413–430. Springer.

Winter, T. (1999). *Online and Real-Time Dispatching Problems*. PhD thesis, Technischen Universiät Braunschweig.

Winter, T. and Zimmermann, U. T. (2000). Real-time dispatch of trams in storage yards. *Annals of Operations Research*, 96(1):287–315.

Zhang, G., Gao, L., and Shi, Y. (2011). An effective genetic algorithm for the flexible job-shop scheduling problem. *Expert Systems with Applications*, 38(4):3563–3573.

Zwaneveld, P. J., Kroon, L. G., Romeijn, H. E., Salomon, M., Dauzere-Peres, S., Van Hoesel, S. P., and Ambergen, H. W. (1996). Routing trains through railway stations: Model formulation and algorithms. *Transportation science*, 30(3):181–194.