Survey of railway maintenance activities from a planning perspective and literature review concerning the use of mathematical algorithms for solving such planning and scheduling problems

Tomas Lidén
October 13, 2014

Contents

1 Introduction 2
1.1 Scope and limitations ................................................. 3
1.2 Organization of the report ............................................ 3
1.3 Terminology and abbreviations ...................................... 4

2 Background 4

3 Railway infrastructure maintenance 6
3.1 Methodology .......................................................... 6
3.2 Maintenance activities ............................................... 7
3.2.1 Categorization and overview ..................................... 7
3.2.2 Contractual forms .................................................. 10
3.3 Possessions ............................................................. 11
3.4 Planning process ....................................................... 12
3.4.1 Planner roles ......................................................... 17

4 A catalogue of planning and scheduling problems 17
4.1 Strategic problems (dimensioning) ................................. 18
4.2 Tactical problems (timetabling/scheduling) ...................... 20
4.2.1 Possession scheduling ............................................ 20
4.2.2 Other tactical problems .......................................... 22
4.3 Operational problems (implementation) .......................... 23
4.4 Summary .............................................................. 25
1 Introduction

This report is written as part of the research project “Efficient planning of railway infrastructure maintenance”, funded by the Swedish Transport Administration (Trafikverket) and conducted within the national research program “Capacity in the Railway Traffic System”\textsuperscript{1}. The aim is to:

- Describe and categorize the maintenance oriented activities that take place on railway infrastructure from a planning and scheduling perspective

- Make a catalogue of interesting planning and scheduling problems as identified during our survey

\textsuperscript{1}www.kajt.org
- Describe and categorize previous work, in the research literature, concerning the use of mathematical algorithms for solving such planning and scheduling problems.

The overall goal of the research project is to develop models and solution methods that can be used to plan and schedule maintenance activities coordinated with the train traffic in such a way that substantial benefits regarding capacity, safety and efficiency is achieved.

1.1 Scope and limitations

In this work we focus on the coordination of train traffic and maintenance activities on a common infrastructure. Hence, we’re less concerned with how to perform the actual maintenance work itself. This is of course an equally important field of study, which includes efficient project planning, lean work organization approaches, automation of manual work, spare part planning, modular components etc. Also we do not intend to study degradation models or how to find accurate estimates to the future maintenance needs. Instead we will regard the maintenance requirements as given (perhaps with some uncertainty) and concern ourselves with the question of how to schedule them in an efficient way together with the wanted train traffic services. Depending on when (in the planning process) the scheduling takes place, the level of flexibility regarding the maintenance and the operative tasks will vary. In the most general case, both types of tasks are planned in the same process step, while in other cases it is a question of fitting one type of tasks into the available slots with minimal adjustment of the already scheduled other ones or such that as much room as possible is left for future needs - be it maintenance or train traffic. Thus one of the important prerequisites is what level of flexibility there is when scheduling the different types of activities.

Our basic assumption is that both operative and maintenance tasks shall be scheduled together in a globally efficient way. Although the purpose of the infrastructure is to enable train traffic, the train operation may not always have precedence over maintenance. Instead they are both needed and dependent on each other - without maintenance the operation will in the end be impossible, without operation the maintenance is of no interest. Thus a healthy balance must exist, which should be reflected in the planning and scheduling processes and tools.

1.2 Organization of the report

In section 2 a general background is given to railway systems and the specific properties that should be considered. Also some terminology, used throughout the report, is listed and explained (along with translations to Swedish). In section 3 railway infrastructure maintenance activities and planning is described and categorized while section 4 lists the planning and scheduling problems we have identified during our survey. Finally section 5 gives an overview of the
applicable research literature and section 6 summarizes our findings and outline
the continued research work.

1.3 Terminology and abbreviations

In this document we will follow the planning terminology as suggested in Rail-
NetEurope (2013a) and RailNetEurope (2013b). Thus the infrastructure owner
is called “infrastructure manager” (IM) and the train operators are called “rail-
way undertakings” (RU). The time reserved for securely accessing parts of the
infrastructure for maintenance actions are called (work) “possessions”.

Further, we will use the British term “turnout” (instead of “switch” or “point”) to
denote the complete system of components that makes it possible for a train
to change tracks. The actual moving parts of a turnout consists of one or more
point machines (or motors) and switch rails. Since maintenance and repair often
concerns these moving parts we will sometimes use the term “switch repair” or
“maintenance of switches”.

The national Swedish infrastructure manager has the name “Swedish Trans-
port Authority” (in English) or “Trafikverket” (in Swedish) which we abbreviate
as TrV.

2 Background

A distinctive property of railway systems is that most activities are exclusive.
Usually you are unable to perform maintenance on the components and sub-
systems (track, power distribution, interlocking etc) currently involved in the
train operation and vice versa. If there are redundancies (e.g. parallel tracks) it
might be possible to perform concurrent operation and maintenance on neigh-
boring parts of the network, usually with some sort of restrictions (e.g. speed
reductions, safety distances). This means that network services can still be
offered during maintenance although the service level might be degraded (e.g.
longer travel times, other routings). Rules and regulations (international, na-
tional and company wide) will set the limitations for how this can be done.
Some countries might allow maintenance work to be carried out on a parallel
track if the train traffic obeys a certain speed limit (e.g. Sweden), while others
might not allow any adjacent train traffic at all (e.g. Holland). Some railway
systems (e.g. certain subway or tram services) might be able to close down op-
eration completely during a couple of (night) hours while others must operate
more or less continuously.

Railway systems have some further complicating properties affecting both
operation and maintenance, that are worth summarizing on this introductory
level:

Interdependency between infrastructure and trains. All rail guided
transportation have a tight coupling between the fixed rail and the moving
wheels, especially when having high weights and/or speed, metal-metal contact
and stiff axles (which practically all rail transportation has). The requirements and tolerances for the track are demanding, both regarding load-bearing (including suspension distribution from rail via sleepers and ballast to the substructure), leveling (lateral and transversal), gauge and displacement. Furthermore the rail surface quality have a crucial importance. All these properties affect riding comfort and degradation speed, both for the trains and the track. Equally important are the requirements and tolerances on the rolling stock (trains, locomotives, wagons and motor units). Flat wheels, slippage, locked brakes or bad roller bearings can cause extensive damages on the track (and thus indirectly affecting other trains). For electrified railways there are equally high demands on the power distribution, both electrically (substation capability, motoring, electric braking, disturbances etc) and mechanically (catenary wire, pantograph etc). Finally, this tight interdependency between infrastructure and trains exists for the complete infrastructure, all rolling stock and the whole transport chain, which makes railways unique when comparing to other transportation modes (shipping, air lines and road traffic).

**Geographic layout of the network and it’s components.** Equipment and crew must be transported to more or less remote locations to perform maintenance. Some of these transportation activities must be done on the infrastructure itself and will thus consume traffic capacity. Furthermore, the different sub-systems will have different geographic layouts, e.g. signal interlocking will not always match the electrification system. Thus a maintenance activity that requires a section of the electrification or interlocking system to be turned off will affect a larger part of the network than one that only requires a specific track or turnout to be blocked for traffic. The way the different sub-systems are partitioned will therefore greatly influence the level of serviceability and maintainability.

**Safety.** Since trains have very long braking distances and run on common tracks, the safety requirements must be high. Sufficient spacing and speed limitation must be guaranteed, both between trains, through turnouts and when approaching occupied or ending tracks. Similarly, safety is crucial when performing maintenance, both for guaranteeing the integrity of the work force as well as trains that are allowed to pass the work site.

**Organization and deregulation.** Several different functional units are involved in making a rail transportation possible, including legislation, design, construction, planning, procurement, infrastructure and rolling stock maintenance, marketing, selling, operation, service and education. Some enterprises, such as Indian Railways (employing about 1.3 million people), cover almost all of these aspects - usually divided into geographical zones or traffic regions. Such large enterprises will always have organizational difficulties and cooperative problems not only due to size, but also due to conflict of interests, economical incentives etc. Primarily in Europe a far reaching deregulation has been go-
ing on since the 1980s, with the overall purpose of opening up for commercial competition in several of the working fields. We will not discuss the pro's and con's of this development, only noting that new requirements and questions are raised and that different roles and responsibilities need to be spelled out and made clear, which is usually beneficial. For infrastructure maintenance this trend has extended the use of maintenance contractors and hence the demands on contractual forms, public procurement as well as planning has increased. The organizational split between infrastructure manager (IM), railway undertakings (RU) and maintenance contractors is a common theme throughout this document.

3 Railway infrastructure maintenance

In this section we first describe the methodology used for collecting the information. Then we categorize the infrastructure maintenance activities followed by a sub-section describing the possessions, which grant access to the infrastructure in a safe way. Finally we explain all the planning process steps, ranging from the very long term to day-of-operation.

3.1 Methodology

The material in this section is based on a series of unstructured interviews with planners, coordinators, technical experts and managers that are involved in planning and performing infrastructure maintenance on the Swedish national railway system. Most of the reference persons comes from the Swedish Transport Administration but several contractors have also been interviewed. The meetings (real life or over telephone) did not have a fixed questionnaire but focused on the following topics: How are the tasks planned and performed, what are the preconditions and effects of the task, what type of equipment and crew is involved, how long possessions are needed/wanted/gotten, what kind of coordinations are done, what are the costs, how large volumes of work is conducted, seasonal variations, suggestions for improvements etc. Each interview has been documented with written notes, which the reference persons has reviewed and corrected.

Rules, regulations, steering documents and guidelines have been collected as well as some statistical material from the IT systems currently in use. Background information and basic facts have been collected from some railway literature (e.g. Bårström & Granbom (2012)) and internet web sites. A good overview of how the railway maintenance is organized and performed in Sweden can also be found in Stenström (2012).

The results are biased towards the Swedish situation and a national railway system, but we try to present the results as generally as possible.
3.2 Maintenance activities

In this sub-section we describe and categorize all major maintenance activities that affects the train operation, normally by requiring possession time. We also make some notes about the contractual forms that are currently used in Sweden.

3.2.1 Categorization and overview

There are different ways of categorizing maintenance activities, based on

- If they are done before or after a fault has been detected, resulting in the classical split into preventive vs corrective maintenance. We will argue that this distinction can be hard to make in some cases and possibly not the best classification for our needs.

- What they consist of, which results in a more practically oriented categorization into diagnostic and restoring actions (further subdivided into technical systems or competencies required). This is often how the maintenance organizations think about and organize their work, e.g. having different crew groups working with inspections and corrections.

- How they are or can be planned, which results in a categorization suitable for the planning tasks. We will present one such categorization model and map the maintenance tasks into it.

Preventive vs corrective The European standard EN 13306 for maintenance terminology use the terms preventive and corrective maintenance, for work taking place before and after a fault has been detected. Preventive maintenance is further divided into condition-based and predetermined maintenance, where the former uses measurements and inspections to determine when actions are needed and the latter uses fixed maintenance intervals/schedules. In addition to these categories TRV sometimes use the term operational maintenance\(^2\) for activities that handle normal operational conditions such as snow removal, slippery rail etc (these activities may also be classified as corrective maintenance). Thus we have the following categories:

- Preventive maintenance (before a fault has been detected)
  - Condition-based maintenance, e.g. measurements and inspection, grinding, tamping etc
  - Predetermined maintenance, e.g. exchange of components (light bulbs, batteries, signaling relays etc) on specified intervals (usually specified by the manufacturer/supplier)

- Corrective maintenance (after a fault has been detected), e.g. fixing short circuits, repairing broken fasteners, welding, work after accidents etc

\(^2\)“Drift” in Swedish
• Operational maintenance, e.g. snow removal, handling slippery rail etc

Above we listed tamping and grinding as condition-based, preventive maintenance tasks. This is normally true, since they mostly are done well before any immediate action is needed. But sometimes the deterioration is faster than anticipated and corrective tamping or grinding is needed. Also, it should be noted that condition-based maintenance usually employ several intervention levels, stating the time frames for restoring actions. At the most serious level, immediate actions is needed and operative restrictions may be imposed (lowered speed or train weights). Such immediate actions are considered corrective since the system is faulty (does not operate properly), although technically the components are not yet broken. These example show that it’s not always clear that one type of activity belongs to one of the above categories. In fact, this categorization is mostly used for contractual and budgetary/follow-up reasons.

**Diagnostic vs restoring actions** The maintenance organizations tend to use a more practically oriented classification, where activities are grouped into

- Diagnostic actions, which consists mainly of inspections and periodic measurements,
- Restoring actions, which consists of all repairs, exchanges, remedies etc.

Bundling all restoring actions together is however not very descriptive, since it will include everything from large track renewal projects to small repairs of insulation joints. Hence a further subdivision is necessary.

The frequency of the diagnostic actions as well as the predetermined maintenance and some restoring actions (e.g. grinding) is determined by a set of factors, the most important being the amount of train traffic (volume, weight and speed). Other factors can be surrounding environment, geotechnical standard, age etc. In Sweden the high speed lines are safety inspected 6 times per year while low speed lines with low train weights may be inspected once a year or even less frequently. It should also be noted that inspection might be conducted simultaneously with preventive or prescribed maintenance, a practice that is adopted for catenary wire maintenance.

Inspections and measurements may result in remarks, which calls for some restoring action. The remark will have a time frame for when the restoring action should be performed. At TrV these time frame codes are immediate/acute (A), week (V), month (M) and before next inspection (B). Note that the time frame for the last code depends on the inspection frequency for the affected track section. Also the time frames are somewhat flexible - for example “week” is normally considered to mean “within two weeks from inspection”, “month” to mean “within one-three months” etc. Some of our reference persons consider these relative time frames a bad practice and would rather see that the inspection remark sets an absolute time limit for when the remark should be handled.

Note that the inspection remark does not specify the restoring action to be taken, just that a threshold limit has been reached and that some action
is needed in order to restore the infrastructure to within the prescribed limits. If there are alternate actions, it is up to the maintenance contractor to decide which action that is appropriate. Sometimes there are more than one option, for example spot tamping with a small vehicle (cheap) or tamping a longer section with a dynamic stabilizing train (expensive). Whichever is chosen will be influenced by the contractual form and sometimes negotiations are taking place between the contractor and the IM before the final action is selected.

**Capacity usage and planning horizon** From a planning perspective we may also categorize the actions according to how much infrastructure capacity they consume and how long in advance they are planned. Ideally the highly disruptive actions (requiring very long and exclusive access to the track) should be planned long in advance while actions requiring less possession time can be planned in later stages. In table 1 we list the different activities according to the needed amount of possession time (per work shift) and how long in advance the planning can be done.

Problematic cases are those that require long possession time but have a short planning horizon:

- Catenary wire replacements are highly disruptive and can usually be planned well in advance. But the degradation can in some cases be very quick, especially if the pantographs of the trains are worn, which may lead to urgent need for replacement or repair.

- Tamping can normally be planned well in advance but in cases where preventive maintenance have been neglected tamping might be needed.

<table>
<thead>
<tr>
<th>Possession time</th>
<th>Activity</th>
<th>Planning horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 8h</td>
<td>Catenary wire replacement</td>
<td>2-3 years / urgent</td>
</tr>
<tr>
<td></td>
<td>Track / turnout replacement</td>
<td>2-3 years</td>
</tr>
<tr>
<td>4-8 h</td>
<td>Tamping of tracks</td>
<td>1-2 years / 1 month</td>
</tr>
<tr>
<td></td>
<td>Grinding</td>
<td>1-2 years</td>
</tr>
<tr>
<td></td>
<td>Switch replacement</td>
<td>1-2 years</td>
</tr>
<tr>
<td></td>
<td>Catenary inspection &amp; maintenance</td>
<td>2-3 years</td>
</tr>
<tr>
<td>1-4 h</td>
<td>Tamping of turnouts</td>
<td>1-2 years / 1 month</td>
</tr>
<tr>
<td></td>
<td>Ultra-sonic testing</td>
<td>1-2 years</td>
</tr>
<tr>
<td></td>
<td>Fasteners, joints, rail repair ..</td>
<td>1-2 months</td>
</tr>
<tr>
<td>As train slots</td>
<td>Periodic measurement</td>
<td>1 year</td>
</tr>
<tr>
<td></td>
<td>Fast grinding</td>
<td>1 year</td>
</tr>
<tr>
<td>0-1 h</td>
<td>Inspection</td>
<td>0-2 months</td>
</tr>
<tr>
<td></td>
<td>Signal repair, vegetation etc</td>
<td>0-2 month</td>
</tr>
<tr>
<td></td>
<td>Slippery rail, snow removal</td>
<td>1 year / 0-1 week</td>
</tr>
<tr>
<td>1h - x days</td>
<td>Accidents, urgent repair</td>
<td>none</td>
</tr>
</tbody>
</table>

Table 1: Possession time and planning horizon
with short notice.

- Rail repair[^3] are often due to problems with cracks found by the ultra-sonic testing. In severe cases the possession time might be 4-6 hours.

- Accidents and urgent repair may result in possessions ranging from hours up to several days.

Most maintenance activities are planned and scheduled in the capacity planning processes, such that they are included in the (daily) operational timetable which is handed over to the traffic control centers. Some activities - indicated with a planning horizon of zero in table 1 - may however be carried out directly in the operative phase. Two types of activities are handled in this way:

- Accidents and urgent repair, which are triggered by external events and are managed by the traffic control center.

- Small and quick maintenance tasks, which are triggered and managed by the maintenance contractors. These tasks can either be secured by 1) a possession, given directly by the traffic dispatcher, or 2) a manual train warning procedure handled by the work force itself[^4]. In either case the tasks must be of such nature that it can be ended/cleared with a very short notice, typically within a couple of minutes (when secured with a possession) or less than a minute (when using a manual train warning procedure). The latter handling is hazardous, especially for sections with high-speed traffic, and the use of it is discouraged.

### 3.2.2 Contractual forms

The Swedish railway maintenance market has undergone a quick deregulation since 2001 (Trafikverket, 2012). Today all railway maintenance work on the national infrastructure is performed in contracts signed after open competitive tendering. Three types of contracts are used:

- Re-investment projects, where parts of the infrastructure is re-established to its intended standard. Typical examples can be the replacement of catenary wires, track sections and turnouts. A contract is signed for each project.

- National maintenance contracts. These contracts are used for activities that are performed with a limited but expensive and/or highly specialized set of equipment and crew that operate over the whole infrastructure network. Typical examples are the periodic track geometry measurements, ultra-sonic testing and grinding. The contract lengths are 3-5 years with a prolongation of 2 or 2+2 years. One contract is set up for each distinctive type of activity.

[^3]: Usually includes cutting up the rail, removing a 1-10 meter stretch, inserting a new one, weld it to the existing track and align the track

[^4]: Preferably, the traffic control center is informed about the work that is taking place, but this is not always the case.
• Regional maintenance contracts. These contracts are used for all remaining activities including inspection, predetermined activities, corrective maintenance etc. The contract length is currently 5 years with a prolongation of 2 or 1+1 years. The Swedish infrastructure network is divided into 34 such contracts.

TRV reevaluate and revise the contractual forms continuously, but the tendency is to move more responsibility to the contractors. This requires revising the specifications from a detailed component/technical level to a system/functional level.

3.3 Possessions

All activities that require secure access to the railway infrastructure must obtain a (work) possession\(^5\). A possession shall guarantee that no trains will run on the designated area - usually coinciding with a signaling stretch such as a line blocking, one or more tracks on a station or a complete station between the entry signals. In addition the possession may impose restrictions on neighboring tracks, usually such that passing trains can only run past the work site with a reduced speed. Depending on the capabilities of the train control system and how temporary speed restrictions are communicated such speed limits may have to be imposed for a longer time period (e.g. one or more days) than the actual possession lasts. In addition, there might be speed restrictions after a work has been carried out until a “burn-in/settlement” period has elapsed. Note also that all temporary speed restrictions affect the train operation and should be communicated/negotiated with the train operators if the timetable has to be altered\(^6\).

Each possession is given a unique work id, in the same way as train paths (slots) receive unique train numbers. Just as with train numbers the possession can span several days. Following is a fictitious example:

156630, from AH-22 to OF-11, M-F 06:30 - 07:55, M-F 10:10 - 11:25

where the id is 156630, the possession area is between signals AH-22 and OF-11 and the work is carried out on Monday to Friday between 06:30 and 07:55 as well as between 10:10 and 11:25, giving a total reserved time of 5 * (1:25 + 1:15) = 13h20m divided into 10 schedule parts - allowing for some trains to pass between 08:00 and 10:00.

The possession area may be loosely given in the early planning stages. As an example, a budgeted track renewal project may have a possession between stations A and D for 8 hours every night for two weeks, but the exact work location for each night can not be detailed until the project has been established and planned - some months before the actual work. In fact, this is a challenge for the IM - to dimension appropriate possessions for the early capacity planning

\(^5\)See (RailNetEurope, 2013a) for the definition of “possession”

\(^6\)If the maintenance work is known long in advance, the timetable may include runtime margins that account for the planned speed restrictions.
applications, which will constrain the contractors that plan and perform the actual work. Effectively the “possession design” will impact the project cost for investment and re-investment projects. Furthermore the exact effect on the train traffic can be hard to envisage for such loosely specified possessions (also when the future timetable is unknown). Simulation is a technique that can be used to quantify the capacity restrictions and find suitable runtime margins and slacks that should be added for train slots affected by the possession.

Some possessions are due to civil engineering work performed in the vicinity of the railway, for example on bridges, tunnels or buildings such that train traffic is not possible. These cases offer an excellent possibility to perform railway maintenance in parallel with the surrounding engineering work.

As noted previously, the sectioning of the signaling system and the power distribution network will determine how much of the infrastructure that will be affected by a possession. If the power must be turned off, no electrified traffic will be possible within the same power section. On traffic lines, this is of little importance since the power sections usually follow the signaling. But on stations and marshaling yards several parallel tracks will belong to the same power section. In addition, all tracks that become dead-ended due to the possession will have very little use for the train traffic (other than close to the start and end time of the possession) - also giving an opportunity for coordinated maintenance work.

Hence the possession should not only describe the work area needed (tracks, signals etc) but the whole affected traffic area. Ideally the maintenance contractor should only need to consider themselves with what components/objects they need access to (and the time needed), while the planning system would keep track of all surrounding objects that should be included in the possession as well as the restrictions imposed on the adjacent traffic tracks\(^7\).

### 3.4 Planning process

Here we describe the complete planning process for obtaining possessions and train operation slots as it works in Sweden today. The work process follows EU guidelines and will probably be quite similar throughout the European countries up till the publication of the yearly timetable. The subsequent steps might differ more between different countries.

The process can be divided in the following steps:

1. Freight corridor planning, where so called prearranged paths (PaP’s) for the international freight trains are established and coordinated with the major possessions (large infrastructure maintenance activities).

2. Preparation and publishing of the network statement, which shall contain all major possessions that the train operating companies should adhere

\(^7\)This functionality, which requires a correct, accurate and date-handled infrastructure register, is lacking in the current planning tools used in Sweden. Instead the possession areas are given as free text in the possession booking record and must be manually verified by capacity and maintenance work planners.
3. Yearly timetable planning, where the regular timetable for all train paths are planned together with the major work possessions.

4. Timetable revision planning, where all dated timetable adjustments are made and final coordination of train paths and possessions should be done.

5. Planning of minor possessions, where plannable work which do not require any train path adjustments are scheduled.

6. Operational planning and control, where the traffic control center will make operative adjustments, authorize unplanned possessions and control all activities (train runs and work) on the railway infrastructure.

Steps 1-5 make up the capacity planning process, while step 6 is the operational phase. In steps 1-4 timetable adjustments and conflicts between different requests are handled, while in step 5 only requests for “spare” capacity should be handled. The handover between step 5 and 6 happens one day before the operational day at TrV.

We will now spell these six steps out in greater detail, describing the process as it is intended to work according to guidelines and regulations, while noting any known deviations from the target process. We make a distinction between:

- Major possessions, which will (or is likely to) be in conflict with one or more train paths and hence requires coordination (handled in steps 1-4)
- Minor possessions, which do not affect the published train paths (handled in steps 5-6)

Whether a possession is major or minor depends on several factors, such as the possession area and its duration, the time-of-day, the train traffic patterns and whether a published timetable exists or not. A very short possession can be considered major as soon as there is a conflict with a (wanted or scheduled) train path. Conversely a possession of several hours could be considered minor if no train paths will run on that part of the infrastructure the same day. The IM will use rule-of-thumb or a specified criteria for which possession to consider in which planning step. Day-time possessions are usually more severe than on the night or over the weekend, while the work cost follows the opposite pattern.

**Freight corridor planning** This process should follow the guidelines given in RailNetEurope (2013b). A number of rail freight corridors (RFC) exist in Europe and to secure a stable rail freight service across them, the traffic is run on so called pre-arranged paths (PaP). These PaP belong to the RFC and have priority over the regular (national) timetable. They are planned roughly one year ahead of the yearly timetable planning and the RFCs should organize two coordination meetings per year: in November and in May, where all concerned IM’s ad RU’s shall participate. Since the appropriate major possessions should be known to the RU’s when making their PaP requests, the specified deadlines are as follows:
In December, T-24m\(^8\), the major possessions should be initially given.

In July, T-17m, the initial PaP applications should be given as well as any further information regarding the possessions, after which the IM’s shall construct the PaP’s (starting in August, T-16m).

In December, T-12m, the final coordination of the possessions and the PaP’s shall be done.

In January, T-11m, the PaP’s should be published (no later than 3 months before the final date for requesting capacity in the yearly timetable planning).

This means that the PaP’s are published one month after the network statement. After publication the PaP’s and the possessions may be updated (after coordination between IM’s and RU’s) in March, T-9m (one month prior to the final date for train path requests), and in August, T-4m (one month prior to the final allocation of train paths).

The criteria\(^9\) for which possessions to include in the RFC coordination are that they will either cause significant changes in the timetable such as rerouting of trains or delays of more than 60 minutes, is a permanent closure of more than 7 days in a row or is partitioned in temporal closures or operational restrictions for more than 30 days in a row. Bigger possessions should be coordinated early while lesser ones may be handled in the later stages of the process.

**Preparation of the network statement** The network statement sets the prerequisites for a timetable period and thus lays the ground for the yearly timetable planning process. It shall be published in December, T-12m, and shall contain all major possessions that the RU’s are expected to adhere to in their planning and train path requests. The deadlines are as follows:

- In August, T-16m, all major possessions should be given.
- In October, T-14m, coordination meetings between the IM and the RU’s are taking place regarding the major possessions.
- In December, T-12m, the network statement is published.

There are general criteria for which major possessions that should be included in the network statement\(^{10}\), but a final decision is taken for each individual case. Between 20-60 major possessions are published by TiV in the network statement.

\(^8\)We use the notation T-24m to denote 24 months before the timetable T starts.

\(^9\)These criteria may differ between the different RFC’s.

\(^{10}\)E.g. permanent closure for one or more days, temporary closures of 6-8 (night) hours for more than a week.
Yearly timetable planning  In the yearly timetable planning, all regular train paths and major possessions should be entered and coordinated. However, the regular timetable is created for one or more typical weeks and for TrV it is currently not feasible to perform a complete coordination and traffic adjustment for all known possessions - that would require a dated planning for more or less the complete year.

For handling of the possessions, the deadlines are as follows:

- In February, T-10m, all the wanted (major) possessions initiated by TrV should be given. For contractors the final deadline is April, T-8m.
- During April to August, T-8m – T-4m, coordination meetings take place between TrV, RU's and contractors regarding the upcoming timetable period T
- In June, T-6m, a proposal for the major possession plan (together with the preliminary regular timetable) is published
- In September, T-3m, the major possession plan (together with the final regular timetable) is published

The criteria for which possessions to request in the yearly timetable process is not firmly stated, but durations of 4-6 hours or longer is a common rule-of-thumb. The amount of possession requests in the yearly timetable process at TrV is between 1000-2000.

Timetable revision planning  In this step the complete dated planning is performed, which includes making adjustments to train paths, handling all conflicts between different requests as well as adding and removing train paths and possessions. Special weekend or holiday traffic is handled as well as changes from RU's and contractors. This is the final step where major possessions should be handled.

TrV currently perform four revisions which divide the timetable into roughly the following periods:

- R1: December to March, which is settled about 8-10 weeks before (R1-10w – R1-8w), i.e. in October
- R2: April to June, settled about R2-10w, i.e. in the end of January
- R3: July to September, settled about R3-10w, i.e. in the end of April
- R4: October to December, settled about R4-15w, i.e. in the middle of June (before vacations start)

The ambition is that all possessions that affect the train traffic shall be coordinated and settled with the affected RU’s and contractors no later than 12-14 weeks before the actual operating day (or week) - a time limit that is not always kept. The limit is set so as to give reasonable time for the RU’s to plan their fleet circulations, crew schedules as well as handling changes in booking systems and towards customers.
Planning of minor possessions  This process is continuous and handles a rolling 8 week period, with weekly increments/handover. The deadlines concerning a specific operating week W, are as follows:

- W-2w is the latest time for application of a minor possession. Apart from the possession application, all work plans and safety documents must be finalized at this time. The minor possession handler shall then verify that there are no train conflicts, that all paper work is in order etc.

- W-1w the possessions should be approved and the plan locked, normally on Thursday for the upcoming week starting on Monday.

The compliance to these deadlines varies. For traffic intense areas (around the major cities) the time limits are well respected while it is less so at other areas. It might even occur that possessions are registered and approved one day before the actual work day.

The typical duration for minor possessions are between 10 minutes up to 1-4 hours. There is currently no statistics regarding the amount of minor possessions, but the total amount of possession requests after timetable publication (including both the major ones handled in the revision planning and the minor ones) is currently around 16 - 17 thousand.

Operational planning and control  One day before the operational day (D-1d) all necessary documents are generated (daily train graph and possession descriptions, including contacts, safety informations etc) and responsibility is handed over to the traffic control centers.

During the operative day (D), unplanned possessions are authorized using a manual procedure (called direct planning), where the dispatcher documents the possession by completing a form, including work id, contact information, description etc. This is done on paper except at one control center where it’s included in an electronic train graph system.

Whenever accidents or situations that require urgent repair happen, the following procedure is used:

1. The dispatcher makes an error report and alerts the operating technician who in turn will contact the appropriate maintenance contractor.

2. The maintenance crew shall immediately (within a contractually agreed time limit) move to the problematic location, make an inspection and return with a proposed repair plan (including time estimates).

3. Meanwhile the dispatcher handles the acute situation and prepares reduction plans.

4. Once the maintenance contractor has given a proposed repair plan the traffic control center can make a complete action plan for the train traffic, passengers and repair work. It can be a delicate balancing act to handle all requests from customers, RU’s and maintenance contractors in a reasonably efficient way.
5. Once the repair work is finished a recovery plan is needed in order to reestablish a more normal traffic situation.

Currently there are no tools that assist the dispatchers in this cumbersome task. However, at one traffic control center, pre-made reduction plans are used. This can be of great help, since the decisions will be quicker, more detailed and easier to communicate (when a suitable plan exists). Furthermore the reduction plans can be better prepared, since they are done under calm conditions and by involvement of several stake holders.

One of our reference persons estimate that each traffic control center, on average, handle 5-10 directly planned possessions per day. Furthermore, all planned possessions will not be utilized in reality, partly due to some over-planning but also due to last minute changes, rescheduling of maintenance crew etc. We have not yet been able to quantify how large this amount can be.

### 3.4.1 Planner roles

The described planning can be organized in several ways. Apart from the traffic control center, TrV have three defined roles: major possession planner, revision planner and minor possession handler. In some cases one person takes on two of these roles (usually the first two).

The major possession planner work with planning steps 1 to 3, while the revision planner handles step 4 and the minor possession handler takes care of step 5.

The first two roles work with major possessions, have close collaboration with traffic planners and handles negotiations with the RU’s. The minor possession handlers are more geared towards the contractors and operative personnel and focus lies on securing that all practical and administrative work regarding issues like safety documents, work planning etc are in place. Previously these tasks were organized together with the traffic control centers, but have now moved to the capacity planning units.

### 4 A catalogue of planning and scheduling problems

Here we list classes of planning and scheduling problems that have been identified during our survey. We try to cover most of the planning steps and resource aspects but make no claim of having a fully complete listing of all types of maintenance oriented planning problems. Some of these problems have been pointed out by our reference contacts as important or demanding to handle. Others have been identified by us as potentially important, due to their seemingly large impact on the overall efficiency of the plan or the process (directly or indirectly). Several of these problems constitute interesting opportunities for future research and/or the use of decision support and planning tools. In some cases we present different approaches to achieving the same overall goal.
We divide the problems into strategic, tactical and operational problems. In the strategic class we put problems concerning dimensioning, localization and organization, which usually concern time horizons of one to several years. Tactical problems include scheduling, timetabling and construction of plans covering a medium long time horizon (weeks to year), often handling resources as categorized, anonymous objects. In the operational class we list problems concerning implementation and effectuation, covering short time horizons (hours to month), usually handling the real individual resources, as well as real-time control. In several cases the described problem does not clearly belong to one or the other of these categories and might, for example, be handled both in the long-term and mid-term planning. In some cases, the long planning process times forces a problem that is tactical or operational in its nature to be handled in the long-term and mid-term steps. Since this can be seen as a deficiency in the planning process, we let the nature of the problem have precedence over the currently used timeline.

Budai-Balke (2009) suggest another classification scheme, dividing the maintenance planning process into the following decision steps: 1) Budget determination; 2) Long-term quality prediction; 3) Project identification and definition (diagnosis); 4) Project prioritization and selection; 5) Possession allocation and timetabling of track possession; 6) Project combination; 7) Short term maintenance and project scheduling; and 8) Work evaluation and feedback loop. These steps are identified from the practices used in the Netherlands and United Kingdom. Roughly the steps 1-4 are strategic problems, steps 5-7 (partly) tactical and parts of 7 operational. At the end of this section we will map our problem catalogue and the different classifications together in a common table (see 4.4).

4.1 Strategic problems (dimensioning)

Maintenance dimensioning The classic maintenance planning problem is to decide what to do when, e.g. when should tracks be exchanged, grinded, tamped etc. This is a question of dimensioning the maintenance volumes and allocate them over the infrastructure network while considering restrictions regarding economy and resources as well as the train traffic volumes. The overall system capacity and service level should ideally be considered as well as forecasts regarding degradation.

At TrV a national multiyear business plan is made for the replacements of tracks, turnouts, catenary wires and bridges. On the regional level, maintenance engineers monitor the infrastructure status and devise plans for the regular maintenance work (e.g. grinding, tamping etc). The level of coordination between different technical systems (track, signaling, power distribution) varies significantly over the country. Furthermore the appraisal of benefits (such as network capacity, riding comfort, reduction in future repair costs and delays) is not well developed. Instead criticality and cost will be the primary evaluation criteria, resulting in a strategy to do the most urgent things that fit into the budget.

A few examples of network-wide dimensioning concerning maintenance and
traffic have been found (see 5.1.1). Plenty of research work (see 5.1.2 and 5.1.3) has been done regarding life-cycle-cost studies (including both maintenance and traffic), maintenance interval studies (considering degradation due to traffic) on typical line stretches or for specific maintenance activities (tamping, grinding) and renewal planning/scheduling.

**Contract design** The infrastructure manager will split the maintenance into several different contracts, some being national, other regional, some concerning a specific type of maintenance and others bundling smaller tasks together. How these contracts are designed will greatly impact cost, quality and efficiency over long time horizons. The factors that can be varied are

- Scope e.g. geographical split, contract lengths and maintenance content
- Form e.g. detailed activity specifications, task-oriented volumes or functional requirements, tendering process
- Terms e.g. economy, fixed/variable pricing, incentives, indicators and statistics

Several qualitative and comparative studies have been performed on one or more of these factors but so far we have not seen any work where they are treated as a quantitative design problem. This might be a very interesting research field, given the great impact the contract construction has. In our interviews the questions concerning contract design, especially scope and incentives, also have been brought up several times, which further indicate the importance of this field.

**Maintenance resource dimensioning and localization** A large number of maintenance resources (equipment, crew and material) are needed. In order to guarantee adequate service levels (e.g. response times, mean-time-to-repair etc) these must be dimensioned and localized in an efficient way. To a large extent this is the concern of each maintenance contractor, but sometimes pooling may be used (for expensive equipment) or central warehousing prescribed by the infrastructure manager. In these cases the planning of these resources becomes a common concern. When several contractors collaborate regarding equipment or perform work sharing, interesting questions arise regarding how to split the benefits of doing so.

We have found no references focusing specifically on these issues for railway maintenance, but in other domains such as airlines, logistics, network and rescue planning etc a lot of references can be found. In Peng et al. (2013) however, a simplified version of an operational planning tool for inspection teams, is used for yearly dimensioning purposes.
Network and planning robustness 11 As with all planning, scheduling, real-time operation and maintenance, the issue of uncertainty is crucial. Due to the long planning process times this is even more so for railway infrastructure maintenance. A prognosis of the possession time must be made long before the detailed work planning has started, the actual infrastructure status may change substantially before the actual work is done, the resource situation and/or budget restrictions may change, new funding might emerge, etc. How to cope with these uncertainties is a challenging problem, so as to achieve reasonable robustness and flexibility both regarding the infrastructure network as well as the planning process.

Here we see several interesting research opportunities, concerning for example sensitivity, probability and effects of changes, levels and consequences of margins, planning details, capacity usage, cancellation of planned work, late introduction of new work etc. Similar robustness issues has received a lot of attention regarding railway traffic planning in recent years but much less work has been done regarding infrastructure maintenance.

In our literature study we have found no references having robustness as the main theme and only a few ones that consider the subject partly (see table 2).

4.2 Tactical problems (timetabling/scheduling)

In this section we dedicate a specific sub-section to possession scheduling questions, since there are several ways of addressing this issue and it is the focal point for coordination with the train traffic. Afterwards we will discuss some other tactical planning problems.

4.2.1 Possession scheduling

Major possession scheduling As described in 3.3 and 3.4 the scheduling of major possessions, coordinated with the train traffic, is perhaps the key planning problem regarding railway infrastructure maintenance since it 1) is conducted all the way from freight corridor to timetable revision planning, 2) has a fundamental impact on the traffic capacity and 3) frames the work planning and cost conditions.

The current practice in Sweden requires the possession time to be very detailed (on the minute) but the area can be somewhat loosely given. Ideally it should be sufficient to specify a rough start time (date, day/evening/night, window etc) and the min/max duration, perhaps given as alternate durations depending on how many slots the possession is split into. Such loose possession requirements would call for more advanced scheduling models and tools. TrV is running a project that aims at introducing such such loose requirements for the train slots (see Forsgren et al. (2012)) and studies have been done on how to address maintenance possessions as well (see Forsgren et al. (2013)).

11 It can be argued that robustness is not a problem area of its own, but should rather be treated within each specific planning problem. Due to its importance we do however want to highlight it specifically and suggest that it be treated as a strategic question.
Several research references (see 5.1.4) consider different aspects of this problem, but there is still plenty of interesting questions left, such as what level of flexibility to allow for the trains and the possessions, how to consider maintenance resources, how to achieve robust plans etc.

**Regular possession pattern construction** Instead of acquiring a minor possession for each and every regular maintenance activity it can be beneficial to preschedule patterns of well-sized possessions that give access to every part of the infrastructure with certain intervals, say once every four to six weeks. Thus several activities will be coordinated to these slots, the capacity planning less burdened and the traffic situation more predictable. The construction of these possession patterns (which can almost be regarded as a strategic problem) is however not trivial. The maintenance volumes must be known and several factors need to be considered, like possession sectioning and timing, work hours, coordination with traffic and other (major) possessions, capacity restrictions past the work place etc.

This model has been used in the Netherlands (see van Zante-de Fokkert et al. (2007)). In Sweden, a similar approach called “channel templates” is recommended on double track lines where maintenance work should be done on the up/down track on odd/even weeks (so as to minimize the need for trains to cross over from one track to the other). Currently a more structured approach called “service windows” is under consideration, where the idea is to construct these possession patterns before the network statement is published and have them given as preconditions for the maintenance contracts.

**Possession coordination** There are plenty of possibilities for scheduling possessions efficiently, for example possession sharing by two or more work tasks, synchronizing possessions to tracks unusable for traffic or “in the shadow” of other possessions, performing small tasks close to or together with larger ones, create possession patterns that fit well with the traffic patterns etc. Some of these ideas have been studied in the research literature (e.g. in Budai-Balke (2009)).

**Maintenance work coordination** A similar approach is to coordinate the work content so as to minimize the possession time and cost (see e.g. Jenema (2011) and Peng & Ouyang (2014)). Examples are so called opportunistic maintenance, project combination etc.

An obstacle for both possession and work coordination appears when several (possibly competing) contractors and operators are involved. Each and every organization will want to run their own business as efficiently as possible, but usually regard cooperations as time consuming, cumbersome and costly. Hence there are interesting questions regarding how processes, incentives and cost sharing can be set up in a deregulated market so as to encourage globally efficient solutions, rather than suboptimal ones.
Timetable compression Instead of scheduling possessions early, the timetable construction could be done so as to guarantee a certain traffic free time - either as requirements/constraints or as a separate goal to maximize. These train free slots would then be available for ad hoc maintenance work later in the planning process. This approach would treat some maintenance aspects as a subproblem/addition to the normal train timetable construction.

4.2.2 Other tactical problems

Maintenance vehicle and team routing Several maintenance activities require specialized teams with large, expensive equipment and machinery - often constructed as multi-vehicle trains that operate directly on the track, ballast, catenary wire etc. Examples are machines for track and catenary measurement, ultrasonic inspection, tamping, cleaning, grinding, track and sleeper replacement etc. It is essential that these machines and teams are used efficiently in order to justify their high cost. Furthermore they will consume track capacity when parked on stations and side tracks (e.g. when waiting for their work possession slots). Interrupted work slots may also require longer setup times for repositioning the machinery to previous work location.

The maintenance vehicle and team routing problem should consider the work tasks to be performed, transportation from and to depots, interruptions, other train traffic, crew requirements, machine service needs and coordination with other machines (either performing the same or other type of tasks). A special type of coordination is the one between measurement/inspection cars and repairing machines, e.g. how close in time should a measurement train run past a track that has been tamped. Grinding will also become better if preceded by tamping.

Several references treat this problem from a North American perspective (where one company operate and maintain the whole railway system), while European research usually focus on a specific type of machine or maintenance region - see 5.1.5 and 5.1.6.

Maintenance rescheduling As pointed out previously, uncertainty is always prevalent. In cases when budget changes drastically or urgent repair is needed, a rescheduling is needed. Questions regarding which jobs to change or cancel, how to reschedule resources, the best way to get back to the master plan will then need an answer - often within a short response time.

The amount of changes - both regarding maintenance and train traffic - is affected by the construction of the original plan, which can be seen as a robustness issue. The rescheduling problem might appear both as a tactical and an operational problem.

We have found no references regarding maintenance rescheduling.

Scenario planning Another option for handling uncertainty is to prepare recovery or reduction scenarios. TrV today prepare reduction plans for handling harsh weather conditions (snow, storms etc) in which train cancellations and
traffic reductions are given. Sometimes snow and ice removal tours are also prescheduled.

Recovery scenarios could also be prepared for important and crucial accident or repair situations, e.g. a derailing or catenary wire breakage on specific line stretches. Instead of solving these situations in real-time (stressful and with scarce information), better recovery plans can be prepared beforehand, considering factors like fair handling of operators, contractors and customers, knock-on effects, vehicle circulations, emergency handling, work access and passing traffic.

Another type of scenario planning concerns preparations for conceivable maintenance work generated by measurement and inspection cars and activities. It can, for example, be wise to prepare for a certain amount of follow-up repair work after ultra sonic testing has been performed. These follow-up plans might involve necessary equipment, crew and possession times.

We have found no references regarding scenario planning.

4.3 Operational problems (implementation)

Maintenance project planning When a specific maintenance project shall be performed there are several detailed plans to coordinate regarding equipment, crew and material. If (minor) possessions have not yet been reserved, the planner needs to find the best (or least bad) time slot, including transportation, get necessary permits, document the operational restrictions around the work area etc. Crew schedules shall be done according to applicable work regulations and necessary equipment and material shall be ordered and prepared.

Of specific interest here is the coordination with the train traffic and other work activities - not necessarily performed by the same organization. Although most of the planning is taking place at each contractor, there is need for cooperation with other organizations and the infrastructure owner - who will need good support in their capacity planning tools.

The only reference we’ve found regarding these issues is the EU-funded project AutoMain (see AUTOMAIN (2013)).

Work timing and resource scheduling Safety and maintenance inspections shall repeatedly visit all parts of the infrastructure - with varying frequencies. To some extent these visits should be coordinated with other maintenance activities (e.g. reasonably close in time from planned repair or restoring actions). Efficient inspection work tours should be constructed that will cover the prescribed requirements.

Inspections and diagnostic measurements generate a large number of small repair and corrective maintenance actions, each having varying time frames and restoration time. The contractors are then faced with the problem of efficiently selecting the exact timing of each action, bundling the tasks together in work packages, assigning them to repair crews and finding suitable possessions. If well managed, all known inspection remarks are considered (not just the urgent ones), crew are matched better to the tasks (considering capabilities, travel etc) which should result in a smaller backlog and more efficient work
force utilization. If the workload is overwhelming, resources are scarce or a larger maintenance/renewal is planned, some repair actions might be deferred (or even cancelled) - given that the risk of failure is kept within acceptable limits. These problems are very similar to the previously described “maintenance work coordination” and “maintenance vehicle and team routing” problems since they concern the coordination of tasks, routing of resources etc, but the planning horizon is short-term and level of detail is more fine-grained (regarding resources, time and tasks). These problems mostly concern the maintenance contractors.

Another type of timing decision can be seen for recurring volume components (like sleepers and fasteners). Here it might be acceptable to have a certain number of failed units per rail length (if spaced sufficiently), which gives room for selecting a limited number of units to repair per intervention occasion and deferring the rest to a later upgrade action.

Some research have been found regarding these issues (see 5.1.7).

**Track usage planning** As part of the timetabling process a detailed track usage planning must be made, especially at large stations and railway yards. Of particular interest for the contractors is how the trains, locomotives and wagons are parked (when not performing train operation), since maintenance sometimes will be needed on these parking tracks (or nearby tracks, turnouts, signaling or power systems). If the tracks are not cleared, the work will be delayed when the rolling stock has to be moved. Technical problems might also occur for locomotives having their pantograph and electrical system active, when the power system is turned off during maintenance and switched on after completion.

Thus, the track usage planning must be revised (performed by the infrastructure manager) so as to allow for the maintenance work to take place (done by the contractor). Snow removal on large railway yards is a specific action where this replanning is of crucial importance, especially if adjacent tracks are needed for snow collecting vehicles.

**Real-time maintenance/operational control** The traffic control center handles all train traffic and all work possessions. Additionally, they will manage error reports, accidents and urgent incidents, deciding when to call in maintenance contractors and how to handle all operative situations. Unfortunately there are currently few planning and decision support tools to assist the dispatchers in this tremendous task.

Quite a bit of research has been conducted regarding the handling of the train traffic, passengers and goods. But the only finished work\(^{12}\) about operational planning of maintenance possessions and train traffic found so far is Albrecht (2009); Albrecht et al. (2013).

---

\(^{12}\)A PhD project at the Technical University of Denmark, together with Bane Danmark was started 2013 with the intent of studying this problem domain.
4.4 Summary

In table 2 we summarize our catalogue of planning and scheduling problems and categorize them according to class, decision step and planning step. In the last column we list cross references to individual papers or the different lines of research (with number of references) containing optimization methods for these problem classes, further described in our literature review (section 5).

From this we can identify several interesting research opportunities were no or very little work has been identified, such as network design (concerning maintenance issues), contract design, resource dimensioning and localization, robustness, rescheduling, scenario planning, project planning, track usage and real-time operational control. Even for an area like possession scheduling were several references have been found, we see plenty of room for more contributions.

5 Literature review concerning the use of optimization models for railway maintenance planning and scheduling

In the following we will summarize some 60+ research references regarding the above subject. Duplicate or very similar publications has been skipped as well as papers with vague or little research contribution (according to our understanding). We will group them into lines of research, roughly listed in the strategic, tactical and operational order. Within each problem type, we treat the references in their historical order. We do not include any general research work such as pure scheduling, routing, maintenance, reliability, life-cycle costing etc. Instead we limit ourself to railway oriented research about optimization models for maintenance planning and scheduling.

For each line of research, we summarize the references in a table, listing the characteristics along with a mapping to the categorizations given in the previous section (4). When describing the different models we will use the notation \(<\text{period length}>/<\text{time step}>\) for describing the length of the schedule and the time resolution. As an example, 1m/1d shall be interpreted as a schedule horizon of one month divided into 1 day time slots.

Finally, we will present a histogram, showing the how the number of publications have evolved over time which indicate the (well-deserved) growing interest in this vast, challenging and interesting field.

A good introduction to practical maintenance planning issues, in a North American setting, is given in Aspebakken et al. (1991). They describe a business process reengineering project conducted at Burlington Northern, the planning tools introduced and the aspects considered both in tactical (2-3 months prior to day of operation) and operational planning (48 h prior). Planning and conflict resolution is performed manually but supported by a rule based graphical software tool.

Another good overview paper is Ferreira (1997) which covers most planning
<table>
<thead>
<tr>
<th>Problem type</th>
<th>Class</th>
<th>Decision step</th>
<th>Planning step</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintenance dimensioning</td>
<td>Strategic</td>
<td>D1, D2</td>
<td>–</td>
<td>ND: 2 (5.1.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF: 14 (5.1.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RS: 5 (5.1.3)</td>
</tr>
<tr>
<td>Contract design</td>
<td>Strategic</td>
<td>D3, D4</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Resource dimensioning &amp; localization</td>
<td>Strategic/Tactical</td>
<td>D1, D2</td>
<td>–</td>
<td>Peng et al. (2013)</td>
</tr>
<tr>
<td>Robustness</td>
<td>Strategic/Tactical</td>
<td>D4-D8</td>
<td>P1-P6</td>
<td>Higgins (1998); Lake &amp; Ferreira (2002); Zhang et al. (2013a)</td>
</tr>
<tr>
<td>Possession scheduling</td>
<td>Tactical/Operational</td>
<td>D5-D7</td>
<td>P1-P5</td>
<td>PS: 15 (5.1.4)</td>
</tr>
<tr>
<td>Vehicle &amp; team routing</td>
<td>Tactical/Operational</td>
<td>D5-D7</td>
<td>P3-P5</td>
<td>DS: 8 (5.1.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VR: 10 (5.1.6)</td>
</tr>
<tr>
<td>Rescheduling</td>
<td>Tactical/Operational</td>
<td>D4-D7</td>
<td>P4-P6</td>
<td>–</td>
</tr>
<tr>
<td>Scenario planning</td>
<td>Tactical/Operational</td>
<td>D7-D8</td>
<td>P4-P6</td>
<td>–</td>
</tr>
<tr>
<td>Project planning</td>
<td>Operational</td>
<td>D7</td>
<td>P5</td>
<td>AUTOMAIN (2013)</td>
</tr>
<tr>
<td>Work timing &amp; resource sched.</td>
<td>Operational</td>
<td>D7</td>
<td>P5</td>
<td>WT: 5 (5.1.7)</td>
</tr>
<tr>
<td>Track usage</td>
<td>Operational</td>
<td>D5, D7</td>
<td>P4-P6</td>
<td>–</td>
</tr>
<tr>
<td>Real-time operational control</td>
<td>Operational</td>
<td>D7, D8</td>
<td>P6</td>
<td>Albrecht et al. (2013)</td>
</tr>
</tbody>
</table>

Column 3 lists the applicable decision steps as suggested by Budai-Balke (2009), where D1=Budget determination, D2=Quality prediction, D3=Project identification & definition, D4=Project prioritization & selection, D5=Possession allocation & timetabling, D6=Project combination, D7=Short term maintenance & project scheduling, and D8=Work evaluation & feedback loop.

Column 4 lists the capacity planning steps (see sub-section 3.4), where P1=Freight corridor planning, P2=Network statement, P3=Yearly timetable, P4=Timetable revision, P5=Minor possessions, and P6=Operational planning and control.

Column 5 lists references, where ND=Network design, SF=Service life & maintenance frequencies, RS=Renewal scheduling & project planning, PS=Possession scheduling, DS=Deterioration-based maintenance scheduling, VR=Maintenance vehicle routing & team scheduling, and WT=Work timing & resource scheduling.
problems and aspects faced by a deregulated and timetable governed railway operation, such as those in Europe and Australia.

Several software tools for asset management and maintenance planning, have been developed for railway infrastructure purposes over the years (e.g. REPMAN, TRACS, MARPAS, ECOTRACK, SOG) but few include any advanced mathematical optimization modules (although often outlined). We do not include any further survey of these planning and decision support systems here, instead referring interested readers to Guler (2013).

5.1 Lines of research

5.1.1 Railway network design for optimal traffic vs maintenance balance

A strategic approach for dimensioning the maintenance volumes is to account for this in the network design decisions. Lai & Barkan (2011) describe a decision support framework focusing on the network capacity planning problem, where maintenance cost is included in the flow cost of running trains. This type of model could be extended such that the maintenance cost is affected by the investment decisions taken - especially if a multi-period modeling is used.

Lai et al. (2013) consider the problem of assigning the appropriate track class to each link in the network. The different track classes will determine the service level (train speed and riding comfort) as well as the maintenance cost. The proposed model minimizes the sum of the maintenance and the transportation cost, but a multi-objective approach could also be adopted. The initial model is a non-linear multi-commodity network flow model, which is linearized into a large MIP model. Lagrangian relaxation is used in order to decompose the problem to a manageable size for large problems.

In table 3 these references are summarized.

A related type of work is network simulation models (see for example Simonson et al. (2000) and Podofillini et al. (2005)) for studying the network capacity (delays etc) considering the state of each track section, degradation, speed regulations due to track status (probabilities for detection, repair etc), maintenance work and train traffic.
5.1.2 Service life and maintenance frequency determination

An infrastructure manager needs to know the service life of different components in order to assess the long-term economy and for planning future renewal projects. In addition, policies for preventive maintenance and renewal intervals are needed. For this purpose, life-cycle cost models are developed which consider a standard unit of track with some given traffic load. Several such studies and models have been developed.

Lamson et al. (1983) study a heavy haul freight line and develop two models for finding optimal intervals for rail maintenance and renewal actions, one for curved track and the other for straight segments. The first model concerns grinding as maintenance action and rail exchange as renewal and is solved with a dynamic programming approach. The other model calculates the life cycle cost for the whole railway life (some 60 years) depending on the renewal frequency (in number of years) while considering maintenance and renewal costs, different interest and taxation rates etc.

Meier-Hirmer et al. (2005) describe a method for choosing an optimal combination of inspection interval and intervention level for tamping on a high-speed line. Degradation is modeled with a Gamma distribution and calibration is done with real measured data for the a TGV line.

Zhao et al. (2006a) study ballast tamping and renewals jointly in the same life cycle model to determine the optimum service life and number of preventive maintenance occasions. Three different maintenance policies are considered (fixed intervention level, fixed time interval and variable maintenance occasions).

Podofillini et al. (2006) study ultra-sonic and manual inspection of rail cracks and the consequential maintenance (grinding or repair). A Markov model describes all possible states from initial detection (possibly faulty) to different levels of failure, including the maintenance action decisions. A multi-objective optimization model for minimizing maintenance cost as well as probability of rail breakage is used in order to decide on inspection interval times and waiting times before a repair action is performed. A genetic algorithm is used for producing the Pareto front. The method can be used to select the optimal strategy for obtaining a desired risk level for rail breakage.

The Markov failure model is further studied in Lyngby et al. (2008), first for rail cracks where it is calibrated to a specific line and used for selecting inspection intervals to reach a wanted mean-time-to-failure and cost level. A similar model is used for track geometry inspection to evaluate whether inspection intervals can be increased for a given line.

Antoni & Meier-Hirmer (2008) and Antoni (2009) develop a continuous statistical model that can be used for different types of failing components with replacement maintenance. They apply the model to service life calculations for track, signaling, and contact wire systems.

Meier-Hirmer & Pouligny (2008) study the impact of preventive grinding on the total maintenance cost. A life cycle cost model is used for calculating the optimal preventive grinding cycle and grinding depth for different classes of rail profiles.
Liu et al. (2014) develop a cost model for determining the optimal annual ultra-sonic rail inspection frequency. The cost components (collected from different resources) are: 1) inspection, 2) repair (defects and breakage) including train delays and 3) derailment damage including train delays. The cost components have a straightforward dependency on the inspection frequency, which thus can be found by an arg-min calculation.

Gustavsson et al. (2014) study a maintenance scheduling model with interval cost and its theoretical properties. The model is applied to four different cases, one of which is rail grinding. The model captures the fact that more grinding is needed (and hence costs more) the longer time the track deteriorates. A grinding strategy is constructed that select timing intervals (measured in mega gross tons) and number of grinding passes for track sections with different radii. When the setup cost increases, the model will schedule several curves to be grinded at the same time.

Chen et al. (2013, 2014) study the maintenance of electrical power feeding systems. A multi-objective optimization for balancing cost vs reliability is used, where time between maintenance occasions are decided while restoration levels are given as input.

In some recent work, such as Andrews et al. (2014) and Zhang et al. (2013a) (further described in 5.1.3) the use of Petri-Nets have been proposed instead of Markov models.

In table 4 on the following page these references are summarized.

5.1.3 Renewal scheduling and project planning

This family of problems consider a real infrastructure network with varying track quality, traffic load etc. Thus the maintenance and renewal jobs must be jointly planned over a long time period. Often it is beneficial to combine projects to minimize the track closures and reduce work costs.

Lévi (2001) describe the work carried out at SNCF in order to find an optimal balance between maintenance and renewals for each section of their network and thus obtain a 5 year budget. Statistics from historical data could be used for several components, but for sleepers and rails probability models were developed in order to assess the amount of repairs. A rule-based strategy was implemented and used in a simulation program in order to find good estimates for renewal plans.

Zhao et al. (2009) consider ballast, sleeper and rail renewal on a set of track segments over a multiyear period which is shorter than the service life of the components. The objective is to maximize the benefit of synchronizing these renewals. A genetic algorithm approach is used for solving the problem and demonstrated on a 10y/1y * 30 segment instance.

Andrade & Teixeira (2011) study how to schedule preventive tamping (per section) and track renewals (per link). Train traffic is handled as flow volumes and track degradation will reduce the max speed and give longer travel times, which increases the traffic cost. A bi-objective model is used to minimize both maintenance cost and traffic (delay) cost while constraining the maximum
Table 4: References about service life & frequency determination

<table>
<thead>
<tr>
<th>Reference</th>
<th>Model</th>
<th>Activity</th>
<th>Details</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhao et al. (2006a)</td>
<td>LCC</td>
<td>Tamping &amp; renewal</td>
<td>Exponential deterioration</td>
<td>Strat.</td>
</tr>
<tr>
<td>Liu et al. (2014)</td>
<td>Sum of cost components</td>
<td>Ultra-sonic inspection</td>
<td>Models from other references</td>
<td>Strat.</td>
</tr>
<tr>
<td>Gustavsson et al. (2014)</td>
<td>Scheduling with interval costs</td>
<td>Grinding</td>
<td>Discrete time = traffic load. IP formulation</td>
<td>Strat.</td>
</tr>
</tbody>
</table>

Legends: LCC = Life cycle cost; IP = Integer program; GSQP = Global sequential quadratic programming; GA = genetic algorithm; TS = tabu search.
Table 5: References about renewal scheduling and project planning

<table>
<thead>
<tr>
<th>Reference</th>
<th>Model</th>
<th>Objective</th>
<th>Variables</th>
<th>Techniques</th>
<th>Data</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang et al. (2013a)</td>
<td>Project scheduling</td>
<td>Possession time, over- or underuse, schedule robustness</td>
<td>Start time</td>
<td>Genetic algorithm, Petri-Net</td>
<td>TN</td>
<td>Strat.</td>
</tr>
</tbody>
</table>

Legends: MO = multi-objective; LCC = life cycle cost; RN = real data for a network; TL = theoretical data for a line; TN = theoretical data for a network; RL = real data for a line.

...amount of tamping and renewal in each time slot. Simulated annealing is applied as solution method to produce the Pareto front. Experiments are done for a 30y/3m plan on a theoretical network with 6 nodes and 8 links (2.7 km long).

In the follow-up paper Caetano & Teixeira (2013), three track (work) components are studied, namely ballast (tamping), sleepers (spot repair) and rail (spot repair). Renewals can either be done for each component or by combining the work. The unavailability due to track possession for each type of job (or combination) is minimized as one objective and the other is to minimize the life cycle cost of all components, while constraining the yearly budget as well as yearly track renewal length. The Pareto front is constructed with a genetic algorithm. Experiments are conducted on a real life instance - the 336 km line between Lisbon and Oporto, divided into 21 segments with heterogenous traffic and component ages, constructing a 20y/3m plan. The results can be used for estimating the yearly budget needs as well as suggesting the timing and combination of different renewal projects.

Zhang et al. (2013a) study the planning of renewals and large repair jobs, with a special focus on the uncertainties concerning actual deterioration and exactly when the maintenance works will start (due to limited resources). A genetic algorithm is used for generating different plans while a Petri-Net model is used to model the detailed deterioration and scheduling of activities, giving the actual cost and performance (fitness) of the plan. Experiments are done on a theoretical 3y/1d plan with 1-3 maintenance teams and activities that take 3 months to conduct.

In table 5 these references are summarized.
5.1.4 Possession scheduling

Ruffing (1993) describe a software tool, containing an automatic procedure for finding work possession windows and adjusting train slots in an existing timetable, using a sequential and iterative construction heuristic approach. This paper is one of few that handle operational restrictions (slow passing) for trains that run through or beside maintenance work sites.

Higgins (1998) present the problem of scheduling maintenance jobs and assigning them to work crew on a single track line with a given train traffic timetable. The objective is a weighted sum of expected interference delay (train delays due to late ending job as well as job delays due to late trains) and prioritized finished times (as many trains as possible should run on better track). The resulting non-linear model is solved heuristically with a greedy construction phase followed by tabu search. Experiments are reported for a 4d/1h plan on a 300 km line with 120 trains and 80 jobs. The work is continued in Lake et al. (2000) where the objective is to minimize the work cost, work-splitting is introduced and simulated annealing is used in the improvement phase. A small problem with 20 track segments, 5 jobs and a limited number of trains is scheduled for a 1w/1h period. Further studies are done in Lake et al. (2002) and Lake & Ferreira (2002) where comparisons are made between different metaheuristics (simulated annealing, local search, tabu search) and some other variants of the objective function.

Cheung et al. (1999) present a constraint programming model for constructing weekly track possession schedules, implemented and used in practice at the subway of Hong Kong Mass Transit. A set of maintenance jobs are to be scheduled to (nightly) work shifts when no regular train traffic is taking place. The jobs, prioritized according to work content, are given as requests for a location, resource usage, duration and possible work period. The objective is to assign as many requests as possible, based on priority rules, while respecting resource constraints (tracks, locomotives, drivers and equipment) and several job assignment rules.

van Zante–de Fokkert et al. (2007) describe how the maintenance work planning was reorganized in the Netherlands by dividing the track into work zones (see den Hertog et al. (2005)) and constructing regular work possession patterns to give one 5.5 hour access to the track once every 4th week, where all normal maintenance should be performed. The problem was solved in two steps by 1) grouping adjacent work zones into so called single-track grids (STG), and 2) assigning these to specific nights. The first step was done manually, although two references are given for optimization-based approaches. The second step is solved with a MIP model using a hierarchical objective function that minimizes a) the number of nights with maintenance and b) the sum of the max work load in an attempt to achieve an even resource usage for the contractors.

The PhD thesis Budai-Balke (2009) study the problem of clustering preventive maintenance possession for small routine tasks and larger projects together in order to minimize possession and maintenance cost, without considering the train traffic. An exact MIP formulation takes too long to solve. Hence, some
greedy construction heuristics are developed. Later, the use of genetic algorithms and different hybridization schemes are investigated. Experiments are made for 2y/1w generated instances on a single line. Pouryousef et al. (2010) develop the model further by considering several segments, enforcing strict frequencies on routine tasks, handling traffic restrictions with cost penalties and using realistic input data.

Jenema (2011) develop a similar model for scheduling prescribed maintenance onto available train free slots, considering work zone and catenary system sectioning, maintenance and operative requirements as well as possible coordination with larger work projects. The MIP model is applied to a midsize railway junction, using realistic input data and producing a 1y/1d plan (although only a few train free slots are available per week) with given setups of possession lengths and combination possibilities.

Boland et al. (2013) study the problem of adjusting a maintenance plan for a complete transportation chain, consisting both of a railway and a terminal network. The maintenance tasks, which reduce the link capacities, shall be scheduled so as to maximize the transportation throughput. A secondary objective is to minimize the perturbation of the initially given task times. A time-expanded network model is used where the flow represents train loads and capacity restrictions can be complete (e.g. track closures) or partial (e.g. single track operation or speed restrictions). Experiments are made for two 1y/0.5h plans with roughly 1000 adjustable maintenance tasks that can be moved +/- 7 days in half hour steps. To achieve a tractable problem size, heuristic techniques are applied for: 1) reducing the number of potential start times (by analyzing tasks pair wise) and 2) solving a sequence of smaller sub-problems (by applying a rolling time horizon). In the accompanying paper by Boland et al. (2014) the coordination of maintenance tasks is further studied and several local search heuristics (greedy, randomized and for single or multiple job adjustments) are developed and incorporated into the maximum flow solution process.

The only references about scheduling both trains and track possessions in the same model we have found, are in Ruffing (1993) (see above), Albrecht et al. (2013) (based on the PhD thesis Albrecht (2009)) and Forsgren et al. (2013). In all cases, a small number of maintenance possessions shall be introduced into an existing train timetable, allowing different types of adjustments to the trains. Forsgren et al. (2013) address the tactical timetable revision planning case, handle a network with both single and multitrack lines, allow trains to be rerouted or cancelled, consider different running times depending on train stops and use a clique-based MIP model approach to solve the scheduling problem. Albrecht (2009) address the realtime operational control case for a single track line with long-haul freight traffic, allow train times to be adjusted but not cancelled, study several different mathematical formulations (MIP, set partitioning with branch & price, heuristics) and settle for a probabilistic meta-heuristic called Problem Space Search.

The detailed movement scheduling described in AUTOMAIN (2013) is reported to adjust both maintenance and train runs (see 5.1.6 on page 36), but the model is not made public and hence omitted here.
In table 6 on the next page these references are summarized.

5.1.5 Deterioration-based maintenance scheduling

A nice overview of this field is given in Ferreira & Murray (1997) who also discuss some aspects of deregulation. In this sub-section we will focus on the long-term tactical planning issues. Some references consider deterioration for short-term operational scheduling as well, which will be covered in 5.1.7 on page 39.

Murakami & Turnquist (1985) describe a general multi-period non-linear maintenance resource allocation model. A single objective is used that minimizes the facility deterioration. The model is solved with a generalized reduced gradient method, using recursive formulas for calculating partial derivatives. The model is applied to tamping where a set of machines are scheduled on a railway system with three lines, so as to optimize the track quality index over time. Results are given for 1y/3m and 3y/3m problems.

Miwa (2002) consider the scheduling of one tamping machine and presents a MIP model that handles several resource and business constraints. The planning period is 6m/10d applied to a network consisting of 150 km track divided into 300 segments, covered by 5 depots. A sequence of relaxed problems are solved in order to obtain a close to optimal schedule. The method is further developed in Oyama & Miwa (2006) where an initial grouping model is used for bundling segments into maintenance units followed by the tamping scheduling model, now considering a complete year. The objective is to maximize track quality. Experiments are presented from several real life instances and the method has been used in practical planning in Japan.

In the three papers Vale et al. (2010), (2011) and (2012) the tamping maintenance problem is studied. Interventions must start/end on straight rails, but other resource limitations are not considered. Planning is made for 2-5y/90d since inspection is done every third month. Different objectives are used in the papers, such as minimal tamping (within allowed track quality limits) and minimal total cost (calculated as net present value).

Famurewa (2013) also addresses tamping scheduling. A greedy construction heuristic is used to schedule two types of tamping machines (called preventive and corrective tampers), making as much use as possible of each 4h night shift. Experiments are done for a 130 km stretch, divided into 200 m segments and a simulated 2y/1d schedule. The paper specifically study the consequences of different levels of preventive tamper usage. A similar problem is studied in Quiroga et al. (2011), where a set of yearly tamping campaigns are scheduled so as to achieve the best track quality. This paper also addresses the issue of balancing preventive and corrective tamping. Special focus is given to the prediction model since there is great uncertainty in track geometry degradation development over time. This degradation uncertainty is also studied in Andrade & Teixeira (2012) where a Bayesian approach is used.

In table 7 on page 36 these references are summarized.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Model</th>
<th>Objective</th>
<th>Variables</th>
<th>Techniques</th>
<th>Data</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ruffing (1993)</td>
<td>Timetable adjustment</td>
<td>Min delay</td>
<td>Start times, departures</td>
<td>Sequential constr. heuristic</td>
<td>RN</td>
<td>Tact.</td>
</tr>
<tr>
<td>van Zante–de Fokkert et al. (2007)</td>
<td>Possession pattern ↔ work night assignment</td>
<td>HO: Maint. nights, work load</td>
<td>Assignment</td>
<td>MIP</td>
<td>RN</td>
<td>LT</td>
</tr>
<tr>
<td>Jenema (2011)</td>
<td>Assign prescribed maint. to train free slots</td>
<td>Possession, maint., project and closure cost</td>
<td>Assignment</td>
<td>MIP</td>
<td>RJ</td>
<td>LT</td>
</tr>
<tr>
<td>Boland et al. (2013)</td>
<td>Adjust maintenance schedule</td>
<td>Max traffic flow and min adjustment</td>
<td>Task start times</td>
<td>Network flow w heuristic reductions</td>
<td>RN</td>
<td>1y/0.5h</td>
</tr>
<tr>
<td>Boland et al. (2014)</td>
<td></td>
<td></td>
<td>Network flow w heuristic reductions</td>
<td>RN</td>
<td>1y/1h</td>
<td>LT</td>
</tr>
<tr>
<td>Allbrecht (2009); Albrecht et al. (2013)</td>
<td>Schedule maint. possession &amp; adjust train schedule</td>
<td>Total train and maintenance delay</td>
<td>Departures, start times</td>
<td>MIP, Set partitioning, branch &amp; price, PSS</td>
<td>RL</td>
<td>1d/1m</td>
</tr>
<tr>
<td>Forsgren et al. (2013)</td>
<td></td>
<td>HO: Conflicts, cancellations, delays</td>
<td>Start times, alternate routes, cancellations, departures</td>
<td>MIP w clique based cuts</td>
<td>RN</td>
<td>1d/1m</td>
</tr>
</tbody>
</table>

Legends: WS = weighted sum; HO = hierarchical objective; TS = tabu search; SA = simulated annealing; LS = local search; MIP = mixed integer linear program; GA = genetical algorithm; PSS = problem space search heuristic; RN = real data for a network; RL = real data for a line; RJ = real data for a junction; TL = theoretical data for a line; ST = short term planning; LT = long term planning; MT = medium term (revision) planning
Table 7: References about deterioration-based maintenance scheduling

<table>
<thead>
<tr>
<th>Reference</th>
<th>Model</th>
<th>Objective</th>
<th>Variables</th>
<th>Techniques</th>
<th>Data</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murakami &amp; Turnquist (1985)</td>
<td>Maintenance resource allocation and scheduling</td>
<td>Weighted track quality</td>
<td>Machine assignment and time</td>
<td>Non-linear MIP, reduced gradient method</td>
<td>RN 3y/3m</td>
<td>Tact. LT</td>
</tr>
<tr>
<td>Vale et al. (2010, 2012)</td>
<td>Activity scheduling</td>
<td>Tamping actions</td>
<td>Section ↔ time assignment</td>
<td>Linearized MIP</td>
<td>RL 2y/3m</td>
<td>Tact. LT</td>
</tr>
<tr>
<td>Vale et al. (2011)</td>
<td></td>
<td>Tamping cost</td>
<td></td>
<td></td>
<td>RL 5y/3m</td>
<td>Tact. LT</td>
</tr>
<tr>
<td>Quiroga et al. (2011)</td>
<td>Maintenance scheduling</td>
<td>Track quality</td>
<td>Starting &amp; ending depot / section</td>
<td>Construction heuristic</td>
<td>RL</td>
<td>Tact. LT</td>
</tr>
</tbody>
</table>

Legends: MIP = mixed integer linear program; IP = integer program; RN = real data for a network; RL = real data for a line; LT = long term planning.

5.1.6 Maintenance vehicle routing and team scheduling

Several papers have been written about the maintenance team (or gang) scheduling problem (also called curfew planning since the jobs require several work shifts and will impose track closures and reductions in traffic). This is applicable for railroad operators that own the infrastructure and organize all the major maintenance activities or for contractors with network wide responsibility of one or more maintenance tasks involving several teams/vehicles. A given set of maintenance jobs are to be scheduled on a set of maintenance teams with given capabilities, equipment and home locations. All the suggested models handle traffic considerations by imposing constraints on which jobs that can be done simultaneously. The objective is to get good routings/schedules for the maintenance crews such that the total cost is minimized, either measured as work/travel costs and/or amount of soft constraints broken. A yearly plan is constructed, either divided into weeks or days.

Li et al. (2009) is the first information found on this subject, but unfortunately it’s only a presentation with uncommented slides.


Nemani et al. (2010) study three complete models (time-space network, set partitioning with alternative work duties per project and column generation with team routes as columns) and a decomposition approach based on the set
partitioning model, which gives the best result. The models differ regarding ability to handle distance constraints/cost and project "crashing" (getting shorter durations by assigning two or more teams) together with several intermingled side/business constraints.

Boğ et al. (2011) use a MIP model that is solved iteratively as a series of one-week or k-week problems with or without backtracking. The objective is to reduce constraint violations. The method is suitable for interactive use in a scheduling tool. The approach resembles a feasibility problem which indicates that constraint programming possibly could fit as search method.

Peng et al. (2011) use a time-space network formulation to minimize travel and penalty costs. The solution steps consist of a clustering method, then local search (considering teams pairwise) followed by a path-relinking strategy (swapping jobs between teams) for improving the solution. In Peng & Ouyang (2012) the work is continued by introducing several more side constraints to the model and development of a better solution method. An initial solution is obtained from a relaxed/simplified problem which is improved by performing parallel randomized local search where blocks of work are interchanged between teams. The results greatly improve both regarding solution quality and performance.

Borraz-Sánchez & Klabjan (2012) use a job-time formulation and a two step solution approach. The first step produces an initial schedule with a dynamic programming method followed by an insertion and swap handling for the unassigned jobs. The second step improves the solution by extracting sequences of jobs, recombining into new subsequences and solving an insertion IP model (which allows for shifting the start times of the job sequences). In this way they are able to solve a 365-day problem with 1000 jobs and 10-100 teams within 2,5 hours.

In the AutoMain project (see e.g. AUTOMAIN (2013)) a multilevel, modularized planning system is developed consisting of: a) Job scheduling and combination for a 1-3 year plan, typically concerning track inspection cars, tamping and grinding; b) Path finding for each machine transportation between work sites/depots; and c) Detailed movement scheduling/timetabling with possible adjustments of regular operation trains. The modules are integrated such that a complete planning can be achieved automatically. The two main optimization modules is the job scheduling part (using local search and simulated annealing) and the detailed movement scheduling (using a MIP model, however not yet disclosed).

Peng & Ouyang (2014) describe a model for clustering (small) maintenance jobs into week long projects. This is done as a preprocessing step before team scheduling is performed. The total duration of all projects (measured in number of project weeks) plus penalties for violating soft constraints is minimized. The problem is cast as a vehicle routing problem with several side constraints and solved heuristically with a greedy construction phase followed by repeated local improvements and feasibility adjustment. Instances of over 2000 jobs can be clustered into projects for five types of teams in less than 10 seconds. Further details can be found in the PhD thesis Peng (2011).

In table 8 on the following page these references are summarized.
Table 8: References about maintenance vehicle routing and team scheduling

<table>
<thead>
<tr>
<th>Reference</th>
<th>Model</th>
<th>Objective</th>
<th>Variables</th>
<th>Techniques</th>
<th>Data</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gorman &amp; Kanet (2010)</td>
<td>Maintenance vehicle routing and team scheduling</td>
<td>Total team cost</td>
<td>Team and time ↔ task assignment</td>
<td>TSN, CP, GA</td>
<td>TN</td>
<td>Tact. LT</td>
</tr>
<tr>
<td>Nemani et al. (2010)</td>
<td>Work cost + constraint violations</td>
<td>Work cost + constraint violations</td>
<td>TSN, SPP, CG, decomposition heuristic</td>
<td>RN</td>
<td>1y/1w</td>
<td></td>
</tr>
<tr>
<td>Boğ et al. (2011)</td>
<td>Track closures + constraint violations</td>
<td>Travel and work (arc) assignment, constraint violations</td>
<td>MIP, sequence of heuristic reductions</td>
<td>RN</td>
<td>1y/1w</td>
<td></td>
</tr>
<tr>
<td>Peng et al. (2011)</td>
<td>WS: Travel cost + constraint violations</td>
<td>Travel and work (arc) assignment, constraint violations</td>
<td>TSN, iterative heuristic</td>
<td>RN</td>
<td>1y/1w</td>
<td></td>
</tr>
<tr>
<td>Peng &amp; Ouyang (2012)</td>
<td>-</td>
<td>-</td>
<td>TSN, decomposition, LNS</td>
<td>RN</td>
<td>1y/1w</td>
<td></td>
</tr>
<tr>
<td>Borraz-Sánchez &amp; Klajban (2012)</td>
<td>WS: Travel cost + constraint violations</td>
<td>Job-time network, DP + LS, LNS + IP</td>
<td>RN</td>
<td>1y/1d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUTOMAIN (2013)</td>
<td>Possession time, target dates, machine usage, delays</td>
<td>Task time, movement path &amp; time, train adjustments</td>
<td>LS, SA; shortest path; MIP</td>
<td>RN 1-3y/1w-1h</td>
<td>Tact. LT ST</td>
<td></td>
</tr>
<tr>
<td>Peng &amp; Ouyang (2014)</td>
<td>Clustering jobs into projects</td>
<td>Total project duration + constraint violations</td>
<td>Travel and work (arc) assignment, constraint violations</td>
<td>Constr. heuristic + LNS</td>
<td>RN</td>
<td>Tact. LT</td>
</tr>
</tbody>
</table>

Legends: WS = weighted sum; TSN = time-space network; CP = constraint programming; GA = genetic algorithm; SPP = set partitioning problem; CG = column generation; MIP = mixed integer linear program; DP = dynamic programming; LS = local search; LNS = large neighborhood search; IP = integer program; RN = real data for a network; LT = long term planning; ST = short term planning.
5.1.7 Work timing and resource scheduling

Zhao et al. (2006b) study the reliability regarding rail sleeper failures. As long as single sleepers are failing the system can still operate, but when two or three adjacent ones fail, corrective maintenance must be carried out immediately. If inspection show that a set of single sleeper failures exists, there is a decision regarding which should be repaired in order to maintain a wanted reliability of the total system. A k-out-of-n probability model for consecutive sleeper failure is developed and used in an optimization model that can be used after each inspection to determine exactly which sleepers to immediately repair (and which to defer until doing a major upgrade) in order to uphold the wanted reliability. This is an interesting approach to maintaining sufficient quality of service at a minimum corrective maintenance cost. The same type of reasoning should be able to do for fasteners and perhaps other infrastructure components.

Peng et al. (2013) (based on Peng (2011)) describe the problem of constructing work tours for rail inspection teams in order to periodically examine the status of the infrastructure network. This is similar to the periodic vehicle routing problem, but where the periodicity requirements are non-strict (time intervals between visits are allowed to vary with days or even weeks). Several complicating side constraints must also be handled regarding non-simultaneity, time windows and non-preferred assignments. The main application is to weekly construct a Sw/c (continuous time) schedule, where the objective is to minimize the variance from the wanted periodicity and a sum of soft constraint violations. A relaxed version of the model can also be used to solve yearly problems suitable for what-if analysis regarding crew dimensioning, localization etc. The solution algorithm incrementally adds tasks to the schedule by increasing the time horizon and performing a large neighborhood search consisting of different task interchanges.

Zhang et al. (2013b) present a model for constructing a monthly schedule for maintenance teams, based on the track segment status obtained from inspection cars and the subsequent deterioration during the operational period. The objective function include costs for: 1) unsafe operation (segments having unacceptable state), 2) unused component life (due to early maintenance), 3) maintenance cost and 4) travel cost. A genetic algorithm approach is used for finding a heuristic solution, which is demonstrated for tamping (always conducted on nightly 4h shifts) on a small regional network with 10 stations, 8 junctions, 16 sections and 900 segments.

He et al. (2014) consider how to schedule the rectification of track geometry defects, after they have been identified by periodic measurements. Critical defects must be handled directly (within a given due date) but for non-critical defects there is a choice which ones to rectify and which ones to postpone. Based on a statistical prediction model for the degradation and a similar prediction model for the possible derailments due to unresolved track defects, a linear MIP model is developed for scheduling the rectification tasks, with the objective of minimizing either total cost (including rectification as well as derailment costs) or the risk of derailment. Experiments are done for several real data instances of
Table 9: References about work timing and resource scheduling

<table>
<thead>
<tr>
<th>Reference</th>
<th>Model</th>
<th>Objective</th>
<th>Variables</th>
<th>Techniques</th>
<th>Data</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhao et al. (2006b)</td>
<td>k-out-of-n consecutive component repair</td>
<td>Sum of immediate repairs</td>
<td>Unit selection</td>
<td>Non-linear IP, steepest gradient</td>
<td>small TL</td>
<td>Oper.</td>
</tr>
<tr>
<td>Peng et al. (2013)</td>
<td>Scheduling periodic inspections</td>
<td>Periodicity variance + constraint violations</td>
<td>Travel and work (arc) assignment, constraint violations</td>
<td>VRP, increasing horizon + LNS</td>
<td>RN, 8w/c or 1y/c</td>
<td>Oper.</td>
</tr>
<tr>
<td>Zhang et al. (2013b)</td>
<td>Condition-based maintenance scheduling</td>
<td>Unsafe operation, unused comp. life, maint. &amp; travel cost</td>
<td>Team &amp; date ↔ task assignment</td>
<td>GA</td>
<td>RN 1m/1d</td>
<td>Oper.</td>
</tr>
<tr>
<td>He et al. (2014)</td>
<td>Scheduling defect rectification</td>
<td>Total cost or risk of derailment</td>
<td>Task and time selection</td>
<td>MIP</td>
<td>RN 1m/1d</td>
<td>Oper.</td>
</tr>
<tr>
<td>Heinicke et al. (2014)</td>
<td>Routing corrective tamping</td>
<td>Travel cost + service limit penalty</td>
<td>Travel / work assignment</td>
<td>Linearized MIP, relaxation + sub-tour elimination</td>
<td>TN xw/1m</td>
<td>Oper.</td>
</tr>
</tbody>
</table>

Legends: IP = integer program; GA = genetic algorithm; VRP = vehicle routing problem; LNS = large neighborhood search; MIP = mixed integer linear program; TL = theoretical data for a line; RN = real data for a network; TN = theoretical data for a network.

varying size from a major US nation-wide network to produce a 1m/1d schedule.

Heinicke et al. (2014) describe a similar problem of routing one or more machines for corrective tamping of the track, where the locations might have gotten (or will get) a service limitation (in the form of speed reductions) until the geometrical defects have been rectified. This incurs a penalty cost per day until the tamping has been performed. The problem is modeled as a vehicle routing problem with time dependent penalty costs. Several different MIP formulations are studied and compared on a family of theoretical test cases.

In table 9 these references are summarized.

5.2 Summary

The application of optimization methods to planning and scheduling of railway infrastructure maintenance has historically received very little attention. But since 2005, a steady increase of publications can be observed - see table 10. As previously noted, several interesting planning problems have not been studied at all and even where more work has been conducted there seems to be plenty of opportunities for deeper studies, especially regarding the coordination with train traffic. Further work could also be done regarding other solution techniques, hybrid approaches, more resource considerations, stochastic methods, robustness, multi-objective approaches etc.
Table 10: Research publications over time

<table>
<thead>
<tr>
<th></th>
<th>ND - 5.1.1</th>
<th>SF - 5.1.2</th>
<th>RS - 5.1.3</th>
<th>PS - 5.1.4</th>
<th>DS - 5.1.5</th>
<th>VR - 5.1.6</th>
<th>WT - 5.1.7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1990</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>1990 - 1999</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>2000 - 2004</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>2005 - 2009</td>
<td>7</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td>2010 - 2014</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>14</td>
<td>5</td>
<td>15</td>
<td>8</td>
<td>9</td>
<td>5</td>
<td>58</td>
</tr>
</tbody>
</table>

Legends: ND=Network design; SF=Service life & maintenance frequencies; RS=Renewal scheduling & project planning; PS=Possession scheduling; DS=Deterioration-based maintenance scheduling; VR=Maintenance vehicle routing & team scheduling; and WT=Work timing & resource scheduling.

6 Summary and continued work

In this document we have introduced the field of railway infrastructure maintenance and its coordination with train traffic. We have presented

- An overview and classification of the different maintenance activities, the interaction with train traffic and how access is granted to the infrastructure through possessions, the contractual forms and a detailed description of the capacity planning process used in Sweden,

- A catalogue of interesting and crucial planning and scheduling problems concerning this field, and

- An extensive literature review, which has been mapped to the catalogue of problems.

We have identified several areas where no research has been found (see section 4.4) and also indicated some possible future directions for continued work in the already explored fields.

In our research project we will now focus on the area of “service windows”, where we aim at developing new optimization methods for constructing efficient patterns of regular work possessions.

We would also like to suggest some further projects that could be considered as possible continuations within the research program “Capacity in the Railway Traffic System”, such as:

- Collection and statistical analysis of capacity usage for maintenance of the Swedish network, classified into different categories and compared to the volumes of train traffic. In this work, studies should also be done on when and how possessions are applied for, adjusted, finalised and utilised (or possibly inhibited). Data feeding channels from the planning systems developed in other research projects, should be possible to use in this type of work.
• Analysis of how infrastructure maintenance affects and should be treated in the economical assessment of railway and transportation system investments. It is clear that work possessions will consume infrastructure capacity and influence both travel times and delays. But how these effects shall be modelled and measured is not evident.

Railway and maintenance specialists, for example at JVTC in Luleå, could perhaps consider studying:

• Application and calibration of different deterioration models regarding ballast, sleepers, rail and catenary wires, with the purpose of reviewing the current strategies, inspection and maintenance frequencies together with renewal planning. This work should consider the risk levels for failures and undetected deficiencies and be the first step toward a maintenance dimensioning and planning model where service quality, risk and cost efficiency can be accurately controlled. Several of the research references mentioned in this document could be the basis for such a project.

In parallel, Trafikverket might consider (if not already doing so) projects for

• Collection and analysis of inspection remarks and error reports, with the purpose of finding a model for assessing the quality level of the major infrastructure components. Considerations should be made to the traffic load, maintenance level (as found in the above mentioned study), different technical concepts and age. This knowledge is needed in order to confidently plan the future maintenance activities.

Finally, we hope that this survey will inspire others and spur further research and development in this interesting, challenging and much needed field of work.

References


