

Swedish Institute of Computer Science

**Opportunities and challenges
with
new railway planning approach
in Sweden**

SICS Technical Report T2012:11
ISSN 1100-3154

November 2012

Authors:

Malin Forsgren[†], Martin Aronsson[†], Sara Gestrelus[†], Hans Dahlberg[‡]

[†]Swedish Institute of Computer Science (SICS)

[‡]Trafikverket

Opportunities and challenges with new railway planning approach in Sweden

Malin Forsgren[†], Martin Aronsson[†], Sara Gestrelus[†], Hans Dahlberg[‡]

[†]Swedish Institute of Computer Science (SICS), {malin,martin,sarag}@sics.se

[‡]Trafikverket, hans.dahlberg@trafikverket.se

Abstract

Long lead times in railway planning can give rise to a significant discrepancy between the original plan and the traffic eventually operated, resulting in inefficient utilization of capacity. Research shows that the railway sector in Sweden would benefit from a different planning approach in which capacity consuming decisions are pushed forward in time whenever possible. This approach is currently being implemented at Trafikverket, the Swedish Transport Administration. With it follows a number of mathematical opportunities and challenges, some of which will be presented in this paper.

Keywords

railway timetabling, planning, modelling, lean production

1 Disclaimer

This paper summarizes the ideas presented at a talk at the EURO 2012 Conference in Vilnius, Lithuania, July 8-11. The paper was not formally submitted, as no complete proceedings was produced for the conference. Thus, it has not been reviewed by a scientific committee. For related material that has been reviewed, we refer to previously published papers [1, 2, 3, 4].

2 Introduction

A new planning approach called Incremental Allocation (Swedish: *Successiv tilldelning*) is currently being implemented at Trafikverket, the railway infrastructure manager in Sweden [5]. To the typical end customer, i.e. a passenger or goods transportation buyer, the change will mean better quality of services in terms of better punctuality and/or more trains. For Trafikverket and the railway sector, the change means a new way of looking at the timetable, its creation, and its use.

The research behind the new approach was initiated and carried out by SICS, Swedish Institute of Computer Science. Despite the ideas coming from a computer science research institute, the concept of Incremental Allocation does not rely on the implementation of optimizing software or decision support tools, and there are substantial advantages to be gained already by adopting Incremental Allocation using manual methods. To reach the full potential of the new approach, however, optimizing tools will become necessary.

3 Current planning practice

Trafikverket, the Swedish Transport Administration, owns the infrastructure and is responsible for constructing a timetable for the whole country. The railway sector in Sweden is completely deregulated, with approximately 25 railway undertakings (RU:s) competing for capacity (ten of which are freight operators).

The basis for the railway traffic is the yearly train plan, in Sweden referred to as TXX, where XX denotes the year in question. E.g., in July 2012, T12 was in effect. TXX takes effect in December year XX-1 and is valid until December year XX.

The RU:s can request capacity at any time, but in order to compete for attractive train paths, they need to take part in the regular train plan process where next year's train plan is negotiated and established. To do so, they need to submit their applications by mid-April in year XX-1, allowing Trafikverket some months of intensive planning and negotiations with the RU:s before the train plan process is formally finished and the new train plan is published in mid-September.

Once the train plan is officially published, existing train paths may be canceled and new train paths added in the so called ad hoc process that follows, but new train paths are not allowed to affect the already established train paths in any way. In other words, the ad hoc process does not allow an existing train path to be modified to give room to a new train path. In fact, during the ad hoc process, a train path must not be changed for any reason, although cancellations are of course supported.

In practice, train paths are modified all the time during real-time operations, as the dispatchers naturally do not follow the yearly plan slavishly. The fact that the process does not support updating the plans causes some trains to be far from ideally planned, forcing the dispatchers to improvise more than would otherwise be the case.

3.1 Long lead times

As can be seen in Figure 1, the process involves fairly long lead times even without counting the compulsory preparatory planning that takes place both at Trafikverket and at the RU:s. Trafikverket have to e.g. prepare and publish the Network Statement, and the RU:s obviously need to plan what train paths to apply for in April. Notably, the lead times between planning and actual operations exceed 18 months for train services that are planned for the end of a yearly train plan.

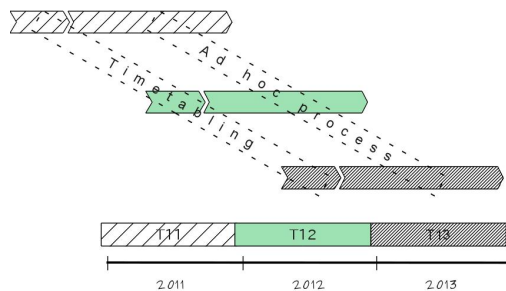


Figure 1: Illustrating the placement in time of the timetabling and ad hoc processes at Trafikverket relative the associated yearly train plan.

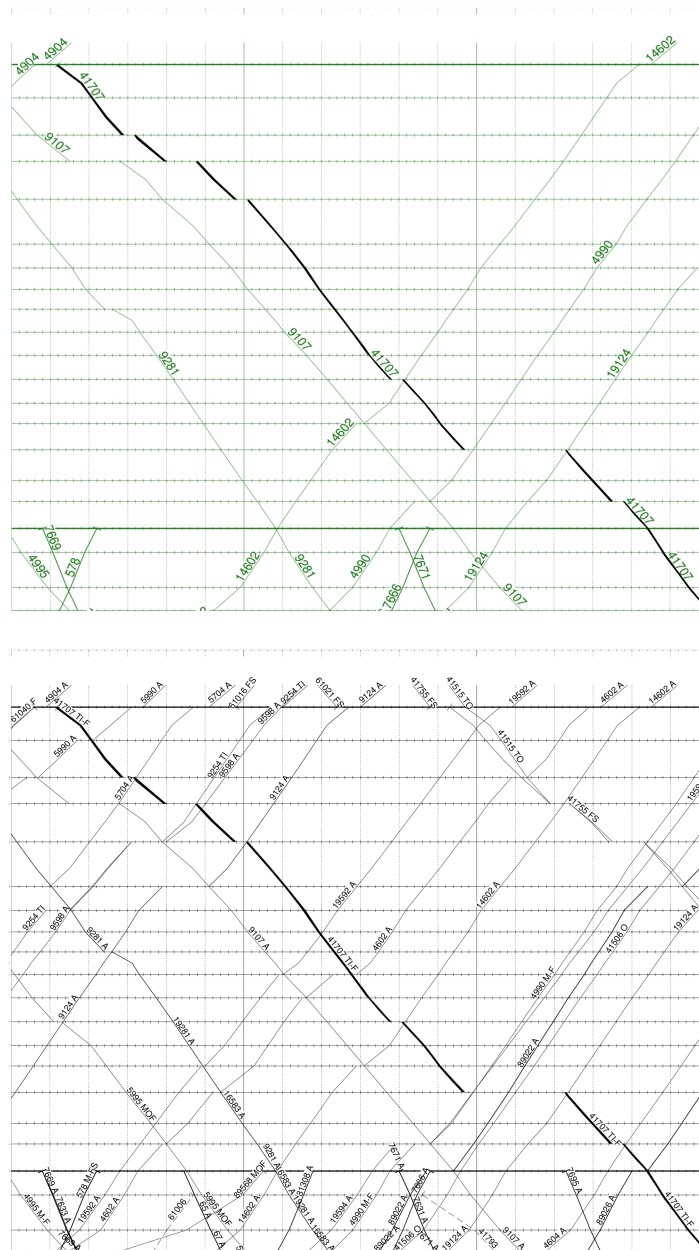


Figure 2: As a consequence of today's principle that train paths should be equal every day the train in question runs, freight train 41707 (highlighted) has a daily plan (top picture) that makes sense from the point of view of the yearly plan (bottom picture), but will clearly not be dispatched that way on this particular day, when most train meetings in the yearly plan will not take place.

4 Incremental Allocation – a new planning approach

Railway traffic in Sweden can in general terms be described as mixed and non-periodic, and a large part of the railway infrastructure consists of single track lines. Given these properties, and considering the drawbacks of the current planning practice as described in Section 3, we, researchers at SICS, saw great potential in allowing the plan to be updated continuously. Also, we found that additional freedom to tailor the plan to better suit each individual day can be obtained by making the yearly train plan slightly less rigid in a way that preserves all the important properties of the train services.

4.1 Basic principles

The new approach, called Incremental Allocation, is based on the idea that existing train paths may be modified as long as this does not affect arrivals and departures that are actually important. Which arrivals and departures that are considered important is decided mainly by whether they are of commercial nature or not, but any selection is of course possible. The significant principles for Incremental Allocation is that

- Trafikverket and the RU:s define, in contracts, which arrivals and departures – and other types of event times of relations – that should be respected at all times, *and*
- The selection of event times and relations as specified above should ideally be the smallest possible selection that is still commercially acceptable to the RU

This leaves settling the details of the rest of the events in the train plan to the production planning at Trafikverket, which may involve slightly different solutions for different days of the week, or potentially unique production plans for every day of the year.

The second point above simply means that Trafikverket should be careful when agreeing to put exact departure and arrival times in contracts and not give promises that are not motivated commercially or deemed essential for the internal planning processes of the RU.

It is assumed to be in the interest of both Trafikverket and the RU:s to strive to keep the number of events and relations that are specified in the contracts to a minimum, as the freedom gained by not locking events in time long in advance can be used for optimizing the whole system. Still, the RU:s need sufficiently specific information, adequately long in advance to be able to publish their timetables and to make plans for their crews and vehicles.

Note that Incremental Allocation is a very simple, yet very powerful, concept. It is easily combined with other principles, e.g. Taktfahrplan (clock-face scheduling), as it does not specify what properties the timetable can or cannot have, but only requires that the contracts should leave as much freedom as possible to the production planning phase.

For a detailed account of how Incremental Allocation can be implemented in Sweden, we refer the reader to our report “The Road to Incremental Allocation and Incremental Planning – Content and Potential” [6].

4.2 Main benefits

Separating the important (commercial) properties of the traffic from the technical details about how to live up to the promises specified in the contracts enables all involved actors to focus on the right thing. Instead of agreeing on a detailed plan, and how to produce what

is in the plan, the parties agree on *what* should be produced, and with *what* precision, after which (advisably) detailed production planning takes place to determine *how*.

Agreeing on event times for only a subset of arrivals and departures instead of promising to carry out the entire train plan in the way it was published means that train paths may be modified even after the train plan has been published. This allows for more flexibility during the ad hoc process when it comes to reusing capacity freed up when trains are canceled, e.g. to add more trains, and it allows for optimization of the daily plan in a whole new way compared with the current process. The latter will remedy situations similar to the one displayed in Figure 2 that are common in today's process despite that fact that, as a rule of thumb, train meetings that take place five times or less in a year are usually not regulated.

Also, the fact that a train no longer needs the exact same timetable every day it runs opens up the possibility to free up infrastructure capacity already in the long-term planning process (see Section 5 for a concrete example). The extra capacity can be used for making the plan more robust and more resilient to knock-on delays, or for adding more train paths to the timetable.

5 On the origins of slack

In this section, we show how slack is injected into the train plan as a consequence of the current planning process, potentially even without the planner being aware of it.

We will use a very simple topology with four stations (A, B, C and D) connected by a single line track, and three trains, Train 1, 2 and 3, all displayed in distance-time graphs with time on the x-axis and distance on the y-axis. Stations B and C only serve as meeting locations, and any commercial stops will take place at A or D.

For the sake of simplicity, we assume that the running times of Train 1 and Train 3 should be kept as small as possible, and that the departure times given in the applications are the earliest acceptable starting times for the trains.

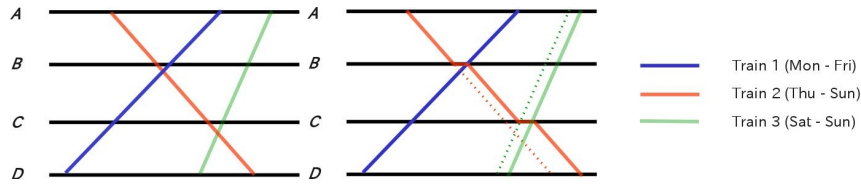


Figure 3: Schematic example of applications for three different trains (to the left), each with a different calendar (legend displayed to the right), and a conflict-free timetable for them (in the middle). The dotted lines indicate the train paths originally applied for, for convenient comparison.

An example of the applications for train paths for the three trains is displayed to the left in Figure 3. The graph to the right in the same figure shows the yearly plan after all conflicts

have been resolved in a straightforward way, taking into account the current principle that the trains should have the same train paths every day they run.

The catch here is that Train 2 will meet precisely one other train every day, never two, since the calendars of Train 1 and Train 3 do not overlap (see the legend in Figure 3). Yet, both meetings have to be regulated in this way, or else Train 2 would need different train paths for different days.

Figure 4 below shows the most efficient solution for only Train 1 and Train 2. In this solution, Train 2 will of course not spend any time at C waiting for Train 3, since Train 3 does not run.

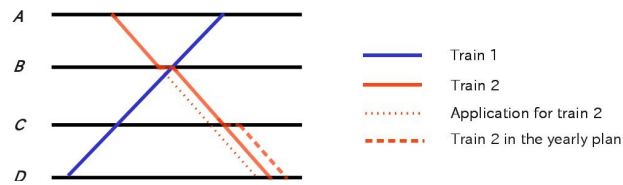


Figure 4: The required capacity for the train meeting that will take place on Thursdays and Fridays is visualized with the respective train paths shown with continuous lines. The original application and the solution for all three trains in the yearly plan are displayed in dotted and dashed lines for comparison.

Similarly, in Figure 5, a more efficient solution for Train 2 and Train 3, compared with the initially suggested yearly plan, is displayed. On Saturdays and Sundays, this timetable could be adopted, giving Train 2 a slightly better running time than in the suggested yearly train plan. Note that the train path for Train 3 in this scenario coincides with the path that the RU applied for.

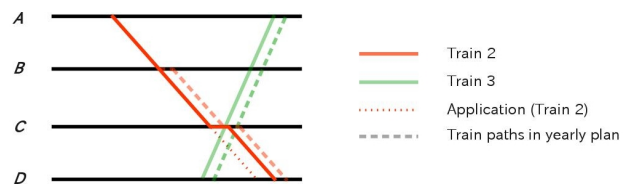


Figure 5: The required capacity for the train meeting that will take place on Saturdays and Sundays is visualized with the respective train paths shown with continuous lines. The original application (where it does not coincide with the suggested train paths) and the solution in the yearly plan are displayed in dotted and dashed lines for comparison.

If we allow Train 2 to be split into two different trains, we can simply replace the initially suggested train path in the yearly plan with the two different solutions for Train 2 and adjust the train path for Train 3 accordingly (see Figure 6).

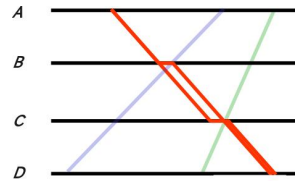


Figure 6: The two solutions for Train 2, merged (both displayed in a darker shade than the shade used for train paths for Train 1 and Train 3).

However, since the RU applying for the train path had not originally applied for two different train services, it is likely that it is preferable to see the two variants not as two different trains, but as production variants of the same train service, and aim for a common arrival time at D. The latest arrival time of the two variants at D is then chosen, thereby introducing some slack on the track between B and D for Train 2 on Thursdays and Fridays.

Note that size of the slack in the solution with the merged variants would be nowhere near the size of the slack given to Train 2 in the first suggestion for the yearly plan, as there is only a slight difference between the arrival times of the variants.

6 Opportunities and challenges

The greatest opportunity with Incremental Allocation is also its greatest challenge: Making constructive use of the increased flexibility in the system. As indicated in the introduction, Trafikverket will undoubtedly be able to exploit parts of this newly gained flexibility with manual methods. The great challenge lies in exploiting this flexibility to its full potential.

6.1 Short-term planning and real-time operations

Closer in time to the real-time operations, and during the actual dispatching, the advantages of having separated “what?” from “how?” become obvious:

1. The daily plan can be optimized
2. The plan agrees with how the trains will actually be dispatched, meaning that it stands a good chance of steering the traffic towards the optimal plan referred to in (1) above
3. Everyday, minor perturbations are absorbed silently to a higher extent, as only the ones that affect the subset of events that are in the contracts have to be reported

4. In the case of major disturbances, the goals for the traffic are better known and the dispatchers can more easily steer towards fulfilling them

The optimization of the daily plan in (1) should ideally take the whole network into consideration at once. This is not yet computationally tractable due to the size of the problem in combination with the level of detail considered in the optimization model. Compared with the approach at Trafikverket today, where no optimization at all takes place, our current models still offer significant opportunities for the improvement of the daily plan.

If the plan is logical from the point of view of the dispatcher, as (2) implies, there will be less need for the dispatchers to change it, which increases the chances that the plan will be followed.

The third point (3) means that less focus is spent on everyday variations that do not cause delays that matter to the end customer. Such delays should not even be called delays, but possibly instead be considered as deviations from the production plan.

Finally, the fourth point (4) means that the dispatchers, with or without computerized decision support, are more free to handle delays when any deviation from the production plan is acceptable as long as the fulfillment of the contracts is maximized.

How to minimize delays in a railway network is of course subject to lots of ongoing research in the field. Our research focuses on off-line scheduling, not on re-scheduling trains in the event of delays. Recently, however, we have begun to explore the topic of robustness in timetables, which in some respects can be seen as a related area of research [2].

6.2 Long-term planning process

While Incremental Allocation clearly paves way for more opportunities to “polish” the details through optimization of the daily production plans, the yearly plan will continue to provide the basic structure for the traffic. Identifying and dealing with inefficient use of capacity already in the long-term process has the potential of improving this basic structure, resulting in a capacity utilization that is significantly more efficient.

Splitting trains into two or more daily variants to make more efficient use of the infrastructure in the manner described in Section 5 is done today, but to a limited extent. After some informal inquiries at Trafikverket about their gut feeling about how widespread this practice really is, we have identified some reasons why it is not a widely used measure:

1. It is a very time-consuming task to identify relevant opportunities and to perform the actual splitting
2. The RU:s prefer fewer variants to keep track of
3. There are not enough train numbers available to support more variants

The first reason (1), we hope to remedy with algorithms and heuristics, which is one of the mathematical opportunities and challenges that we are currently facing.

The second (2) refers to the fact that in today’s process, splitting trains into variants means that these variations have to be communicated as different trains to the RU. This will no longer be a problem when Incremental Allocation has been implemented, as long as the splitting is done within the bounds of what will eventually be specified in contracts.

Of course, if the train variants of a certain train service imply that it will not be possible for the RU to have the same timetable for the train every day, this will still need to be negotiated with the RU in question. But to a much greater extent than today, Incremental Allocation gives Trafikverket the freedom to take these kind of measures without involving the RU:s, since the production details are solely up to Trafikverket anyway.

The third reason (3) is hopefully a passing problem, as the mainframes are being replaced at Trafikverket. In the current system, safety regulations state that every variant during the course of the year needs a unique train number, and there is a technical restriction on how many train numbers that can be handled by the mainframe system.

7 Discussion

In this paper, we have identified some of the benefits of adopting the principle of Incremental Allocation (Swedish: *Successiv tilldelning*) for railway timetabling at Trafikverket, the infrastructure manager in Sweden. Incremental Allocation will improve both the planning process and the resulting timetables, and no new algorithms or tools need to be developed or implemented for this concept to work. However, in order for Incremental Allocation to reach its full potential, more research and more advanced optimization techniques will be necessary.

One of the major benefits of Incremental Allocation is that it frees up capacity. The main challenge described in this paper is how to use this freed-up capacity in a constructive way. We argue that the first step is to identify it, which is a hard problem in itself, and then an active and wise decision about what to do with it has to be made. We are working with these issues but do not have any solutions ready for how either of these things can be done yet.

A valid objection to splitting trains into variants in a timetable in the way described in Section 5 is that there is a risk that it will remove too much slack from the train plan. For this reason, the decision to introduce variants to optimize the running times of the trains in the yearly plan should be made with great care. However, we must not forget that there is nothing magical about the placement of slack that occurs “naturally” due to today’s planning practice, and thus nothing prevents us from redistributing it (or removing some of it) in any way we see fit.

If the yearly plan in Figure 3 was in use, during real-time operations the dispatchers would most likely not ask Train 2 to stop at C on the two days of the week when the meeting between Train 2 and Train 3 does not take place, but instead let it arrive early at D. Arriving too early is however not a good thing. Train 2 might be dispatched early also from D and onwards, in which case it continues to deviate from the plan, possibly disturbing other trains.

For the main example in this paper, hidden capacity was revealed by identifying that one of the trains involved in train meetings will never meet both of the other trains on the same day. We believe that it is important to detect and stop these “capacity leaks”. Trafikverket would benefit from a tool that could identify this situation for what it is, if only to provide the planner with the information that the “old” way of scheduling the trains will lead to a considerable amount of slack for Train 2. If this slack is unwanted, the planner should be given the option to shorten the running time of Train 2 instead.

References

- [1] Forsgren, M., Aronsson, M., Kreuger, P. & Dahlberg, H., The Maraca: A tool for minimizing resource conflicts in a non-periodic railway timetable. *Proceedings of the 4th International Seminar on Railway Operations Modelling and Analysis*, Rome, Italy, 2011.
- [2] Gestrelus, S., Aronsson, M., Forsgren, M. & Dahlberg, H., On the delivery robustness of train timetables with respect to production replanning possibilities. *Proceedings of the 2nd International Conference on Road and Rail Infrastructure (CETRA)*, Dubrovnik, Croatia, 2012.
- [3] Forsgren, M., Aronsson, M., Kreuger, P. & Franzén, T., A novel approach to railway planning. *Proceedings of the 16th ITS World Congress*, Stockholm, Sweden, 2009.
- [4] Forsgren, M., Aronsson, M., Gestrelus, S. & Dahlberg, H., Using timetabling optimization prototype tools in new ways to support decision making. *13th International Conference on Design and Operation in Railway Engineering (COMPRAIL)*, New Forest, UK, 2012.
- [5] Trafikverket, Aktuella projekt – Successiv Tilldelning. Project web page, 2012. <http://www.trafikverket.se/Foretag/Trafikera-och-transportera/Trafikera-jarnvag/Successiv-tilldelning/>.
- [6] Aronsson, M., Forsgren, M. & Gestrelus, S., The road to incremental allocation and incremental planning – content and potential. Technical report T2012:09, Swedish Institute of Computer Science, 2012.