An evaluation of algorithms and systems for Computer-Aided Train Dispatching
P. Hellström¹, B. Sandblad², I. Frej³, Anders Gideon³
¹Högskolan Dalarna, School of Engineering, S-78188 Borlänge, Sweden.
Email: Peter.Hellstrom@cmd.uu.se
²CMD, Center for Human-Computer Interaction, Uppsala University, Lägerhyddvägen 18, S-752 37 Uppsala, Sweden
Email: Bengt.Sandblad@cmd.uu.se
³Swedish National Rail Adm., Engineering Department.

Abstract

A distinctive feature of the train control process is that it is fairly easy to handle when trains are running according to their timetables and no disturbances occur, but is extremely difficult to handle in an optimal way in case of disruptions. Within this research area, one focus has been on development of models, algorithms and systems, which support optimal computer-aided train dispatching. This paper summarises an evaluation of the properties, in general and with respect to rules and conditions in Sweden, of one of the more important of the existing algorithms. We have also implemented it as a minor prototype system in order to, "off-line", evaluate its possibilities and limitations concerning the dynamic aspects of train traffic control.

1 Introduction

The first system for centralised traffic control (CTC) was installed in 1927 on the New York Central Railroad. This system permitted a single
dispatcher to control the operation of trains in a territory without the use of train orders, deciding the location of every train meeting. Seventy years later the train dispatcher still is the most important factor in the process of train traffic control, but now the main question being how to best utilise the potential of more recent information and automation technologies, i.e., how to change supervisory control of complex systems towards a new quality of automation and co-operative man-machine decision making, see also Andersson[4] and Sandblad[5].

A dispatchers chances to, in an active way, plan and control the train traffic are often crucial for how different disturbances effect train delays. Within this research area, one focus has been on development of algorithms and systems, which support optimal computer-aided train dispatching (CATD), see Jovanovic[1] and Peterson[2]. The objective of the algorithms are normally to minimise total weighted train delay. Another focus, less covered in literature, has been on using systems, based on so called priority-based train dispatching. This is the case especially in commercially available real-time systems.

Despite the fact that there is a substantial literature on the problems of optimal timetabling and control, there are many issues concerning different dynamic aspects seldom dealt with in literature. Therefore we have implemented one of the more promising models and tested it with data from a single track main line in the middle of Sweden. This is done in order to evaluate its possibilities and limitations in a dynamic environment. This paper gives a description of the model chosen including an analysis of its pros and cons in general, and an example of how to adapt the model to the dynamic behaviour of real-time train dispatching.

2 The real-time CATD-system

A typical CATD-system has the following features:
- It serves a line between two major stations or terminals. There are normally several meetpoints on that line.
- Every train has a schedule for that line. That schedule represents the initial optimum.
- It’s continuously served with new train positions from the signal systems.
The dispatcher can, if there has been a disturbance, ask the system to calculate and present him a new optimal schedule. The dispatcher makes the final decision whether to use it or not.

A CATD-system that in real-time calculates a new optimal schedule has to be based on relatively advanced optimisation algorithms. In fact, these algorithms are one of the stumbling blocks when designing and constructing such a system. Research is still on finding improved heuristic solutions. The main published work within this area is done by Jovanovic[1].

In a previous paper Hellström[3], we have presented a survey, based on a set of interviews of train dispatchers and other experts on train traffic control. We there showed that there is a great potential for improved decision making in the process of traffic control. Improvements that would result in a better on-time performance. The survey also indicated that the most promising area for an introduction of a decision support system is the single track area. The main reasons are that a dispatcher there has more time to consider which decision to make, and also that the on-time performance of the traffic is very sensitive to the choice of decision. We also found that the dispatchers are not able to fully consider the effects of all possible alternative actions in case of disturbances, and that they have problems predicting the consequences of their various decisions.

The objective with this work is to test one promising model and find out if it can be adapted to handle some of the most frequently occurring traffic disturbances in the daily work. Jovanovic[1] model was chosen because it was the most exhaustive and versatile found in literature at the time the decision was made.

Finally there are also some other fields of application of these kinds of models worth mentioning. The most natural use of a model of this kind is of course on a tactical level as an aid in the timetable construction process, as shown by Jovanovic[1], but it could also be a valuable tool on a more strategic level, for example in the planning process when new investments in infrastructure are analysed and compared to each other, because it makes it possible to generate a new timetable for a, partly or completely, new line, with just a small effort. Another interesting idea is to integrate it with a train traffic simulation tool as a kind of “built in dispatcher” and thereby make the simulated train traffic more realistic or “optimal”, see also Wahlborg[6].
3 Jovanovic model and algorithm

3.1 Overview of the work done by Jovanovic

This section gives an overview of the work done by Jovanovic as described in his thesis Jovanovic[4]. There he presents a detailed mathematical formulation of what he calls "the minimum tardiness cost train dispatching problem" together with two exact algorithms and one heuristic. Noteworthy is that the algorithms actually are for the subproblem with a fixed ordering and overtaking of the trains. The algorithms were tested on relatively large real-worlds problems and the heuristic was shown to give near-optimal solutions in real-time. These tests were performed on a standard work-station of the time, i.e., a very slow computer of today.

The model, covering both single and double track as well as mixed single and double track, was formulated as a non-linear mixed-integer mathematical program in order to give a concise quantitative definition of the problem, but as stated by Jovanovic as well as others, it should not directly be solved as one.

The basic algorithm developed is a traditional Branch&Bound-algorithm based on a binary search tree, where each branch represents a choice between two meetpoints for the meeting in question at that level. In order to quickly find a good initial feasible solution and upper bound, the search was designed as depth first with heuristic branching rules.

3.2 The mathematical formulation of the problem

This section presents the mathematical formulation of the "minimum tardiness train dispatching problem", but only for the part that has been implemented by us, i.e. for a single track rail line and with a fixed ordering and overtaking of the trains. The notation used follows the one used by Jovanovic[4].

3.2.1 Formulation of the problem parameters
This model is restricted to handling the train traffic along one railway line, called a lane, see figure 1. The lane starts and ends with a major station or terminal. Between those there are a number of meetpoints and other stations or terminals. The meetpoints are consecutively numbered in the outbound direction starting with the station at the inbound end of the lane. A number of trains are scheduled to run across the lane or some part of it.

![Figure 1: A (single track) Lane.](image)

### 3.2.2 Parameters and variables

- **N** number of stations and meetpoints
- **m=1..N** meetpoint-number
- **μ_m** length of the siding at meetpoint m
- **I** set of trains scheduled to run on the given lane in the outbound direction.
- **J** set of trains scheduled to run in the inbound direction.
- **σ_i** index for the first station on train i’s itinerary
- **ε_i** index for the last station on train i’s itinerary
- **S_i** the set that contains indices of all points on train i’s itinerary where it is scheduled to depart, arrive, and/or dwell for a specified time.
- **α_i^m** the scheduled arrival time for train i to station m.
- **δ_i^m** the scheduled departure time for train i from station m.
- **δ_α_i^m** minimum dwell time at a scheduled point m.
- **λ_i** length of train i.
- **τ_i^m** minimum unobstructed running time for train i from station m to station m+1.
- **a_i^m** arrival time of train i at station m (variable).
- **d_i^m** departure time of train i from station m (variable).
$e_i^m$ tardiness of train i at station m for all $m \in S_i$ (variable).

$f_i^m(\cdot)$ non-decreasing cost function of the tardiness of train i at station m for all $m \in S_i$.

$\eta_{ik}^m$ minimum headway that allows train k to follow train i through the station m without delay.

$\rho_{ik}^m$ minimum meeting headway (switching times, acceleration delay) at station m between the arrival of train i and the departure of inbound train j.

$c_{ij}^m$ binary decision variable, =1 if outbound train i runs the segment between meetpoints m and m+1 before the inbound train j. =0 otherwise.

In the same way let the attributes of inbound trains j,r be denoted by:

$\bar{\alpha}_{j}, \bar{\beta}_{j}, \bar{\delta}_{j}, \bar{\zeta}_{j}, \bar{\bar{\alpha}}_{j}, \bar{\bar{\beta}}_{j}, \bar{\bar{\delta}}_{j}, \bar{\bar{\zeta}}_{j}, \bar{\bar{\bar{\alpha}}}_{j}, \bar{\bar{\bar{\beta}}}_{j}, \bar{\bar{\bar{\delta}}}_{j}, \bar{\bar{\bar{\zeta}}}_{j}, \bar{\bar{\bar{\bar{\alpha}}}}_{j}, \bar{\bar{\bar{\bar{\beta}}}}_{j}, \bar{\bar{\bar{\bar{\delta}}}}_{j}, \bar{\bar{\bar{\bar{\zeta}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\alpha}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\beta}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\delta}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\zeta}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\bar{\alpha}}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\bar{\beta}}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\bar{\delta}}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\bar{\zeta}}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\bar{\bar{\alpha}}}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\bar{\bar{\beta}}}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\bar{\bar{\delta}}}}}}}_{j}, \bar{\bar{\bar{\bar{\bar{\bar{\bar{\zeta}}}}}}}_{j}$

### 3.2.3 Formulation of the mathematical problem

The problem formulated as a non-linear mixed integer mathematical program, where the constraints are given only for one direction.

Minimise

$$z = \sum_{i \in I} \sum_{m \in S_i} f_i^m(e_i^m) + \sum_{j \in J} \sum_{m \in S_j} f_j^m(\bar{e}_j^m)$$  \hspace{1cm} (1)

subject to:

**Schedule constraints**

$$d_i^m \geq \delta_i^m \quad \forall i \in I, m \in S_i$$  \hspace{1cm} (2)

$$a_i^m \leq \alpha_i^m + e_i^m \quad \forall i \in I, m \in S_i$$  \hspace{1cm} (3)

Preventing early departure of trains from their scheduled points and ensuring that the tardiness variables are equal to the amount of lateness at a given scheduled point.

**Travel time constraints**

$$a_i^{m+1} \geq d_i^m + \tau_i^m \quad \forall i \in I, m \in \sigma_i, \ldots, \varepsilon_i - 1$$  \hspace{1cm} (4)

Restricting the minimum meetpoint-to-meetpoint travel time to the time corresponding to the performance of each train.

**Minimum dwell time constraints**

$$d_i^m \geq a_i^m + \delta_i^m - \alpha_i^m \quad \forall i \in I, m \in S_i$$  \hspace{1cm} (5)

Ensuring that trains spend at least the scheduled amount of time at points where they perform something.

**Travel time continuum constraints**

$$d_i^m \geq a_i^m \quad \forall i \in I, m \in \{\sigma_i, \ldots, \varepsilon_i\} \setminus S_i$$  \hspace{1cm} (6)
Ensuring that trains do not depart before they arrive, at not scheduled points.

**Following constraints**

\( \pi_{ai}^m \) and \( \pi_{di}^m \) is the index of the immediately preceding train entering and the immediately preceding train leaving a meetpoint, respectively. The following constraints ensure that the minimum headway is retained between two trains following each other.

\[
a_i^m \geq d_i^m + n_{ci}^m \quad \forall i \in I, m \in S_i \backslash \{ m : \pi_{ai}^m \neq \pi_{di}^m \}, \sigma_i < m < \epsilon_i \quad (7)
\]

Maintaining minimum headway between two succeeding trains entering a scheduled point or a point where train \( i \) is overtaking another train.

\[
a_i^m \geq d_i^m + n_{ci}^m \quad \forall i \in I, m \in \{ \sigma_i, \ldots, \epsilon_i \} \backslash \{ \pi_{ai}^m \} \quad (8)
\]

Applying to sidings where train \( i \) is not involved in overtaking. It ensures that a minimum headway has elapsed after the previous train has left the siding before train \( i \) can enter that siding.

\[
a_i^m \geq a_i^m + n_{ci}^m \quad \forall i \in I, m \in \{ m : \pi_{ai}^m \neq \pi_{di}^m \}, \sigma_i < m < \epsilon_i \quad (9)
\]

Maintaining minimum headway at scheduled points and points where train \( i \) is overtaking another train.

**Meeting constraints**

The purpose of these constraints is to prohibit meeting at any point where one of the trains is planned to overtake some train.

\[
c_{ij}^m = c_{ij}^{m-1} \quad \forall i \in I, j \in J, m \in \{ m : m \notin S_i \cup S_j, \mu^m < \lambda_i, \mu^m < \lambda_j \} \backslash \{ m : \pi_{ai}^m \neq \pi_{dj}^m, \pi_{di}^m \neq \pi_{dj}^m \} \quad (10)
\]

Preventing meetings at sidings that are not long enough to accommodate at least one of the trains involved in the meet, or infeasible for any other reason.

\[
d_{ij}^m \geq a_{ij}^m + n_{ci}^m - \Psi c_{ij}^m \quad \forall i \in I, j \in J, m \in \{ \sigma_i, \ldots, \epsilon_i \} \quad (11)
\]

Meeting is allowed only at the meetpoints.

\[
c_{ij}^m = c_{ij}^{m-1} \quad \forall i \in I, j \in J, m \in \{ m : m \notin S_i \cup S_j, \mu^m < \lambda_i, \mu^m < \lambda_j \} \quad (12)
\]

Preventing meetings at sidings that are not long enough to accommodate at least one of the trains involved in the meet, or infeasible for any other reason.

\[
c_{ij}^m \leq c_{ij}^{m-1} \quad \forall i \in I, j \in J, m \in \{ \sigma_i + 1, \ldots, \epsilon_i \} \quad (13)
\]

A given pair of trains can meet only once.

\[
c_i^m \geq 0 \quad \forall i \in I, m \in S_i \quad (14)
\]
3.2.3 The algorithm
The algorithm has been designed and implemented as a minor prototype system in order to make it possible to evaluate the model.

Our implementation of the algorithm follows only to a certain extent the greatly simplified description made by Jovanovic, as it has not been possible to get hold of the one implemented by him.

4 The pros and cons of Jovanovic’ model

This section includes a discussion concerning the actual model and algorithm created by Jovanovic, based on our experiences from studying, implementing, and working with it.

4.1 Pros
The major advantages of the model is that it is close to the corresponding real system and gives a concise formulation of the problem, at the same time as the level of detail is well adapted to the purpose of the model, i.e., to be able to create a fairly robust timetable for a railway line in real-time. Some related factors worth mentioning here, are for example that:

• meetpoints are treated as points.
• it handles both single and double track as well as mixed single and double track.
• incorporates both arrival and departure times for trains.
• the basic algorithm is a standard Branch-and-Bound-algorithm, although there are several other sub-algorithms is involved as well.

Another important factor is that the heuristic developed is proved to be quite efficient in time, finding a near-optimal solution.

Finally, on the low level there are some important parameters included, e.g., a slack, a tardiness, and a cost function parameter for every scheduled station of a train, and the minimum following headway as well as minimum meeting headway times for every track section and every pair of trains.

4.2 Cons

• There are three major disadvantages with the model and they are:
• the absence of the formulation of a well-functioning heuristic solving the general scheduling problem, and not just the subproblem with fixed ordering and overtaking.
• that there are no more than two tracks at any meetpoint except at the scheduled points.
• that all stations having a scheduled arrival or departure time for a train becomes a station with an unlimited number of tracks.

The last two is of course related, so a solution for one of them will probably eliminate the other as well.

Some of the minor disadvantages are that the cost function has to be non-decreasing, that it does not cover networks at all, that it is not known which track is used by which train at a station, that the running times between stations does not explicitly include the additional time necessary for acceleration and deceleration in case of train stops, that the description of a station at the micro level is obscure, and that it is not known how different constraints added will affect its time performance.

6 Test cases, an example

We have created a sample of test cases, based on real-world data. The test cases are: 1. a speed restriction to 40 km/h on a track section, 2. a large delay to one train, 3. several trains with minor delays, 4. a locomotive engine breakdown, 5. a track section closure. We have, with good results, tested to adapt the model to handle the types of disturbances and faults that these cases represents. Here we present our results for case 1. The results obtained from the other cases will be published later. The scenario is in all cases, that one or more faults or other disturbances has come about, making the original plan (timetable) more or less unfeasible. So, finally at a certain point of time the dispatcher asks the assumed CATD-system to reschedule the line, with midnight as the upper limit of the time horizon. We are using a linear cost function, with different minute-costs for different train types in the model.
The real-world data is from a single track main line in the middle of Sweden, with mixed operation of freight and passenger trains with different speeds. The data was obtained from Banverket (the Swedish National Rail Administration) and SJ (the Swedish State Railways). The original train graph is shown in figure 2 below.

**Example Case: Speed restriction to 40 km/h**

A speed restriction to 40 km/h on track section HÅG-VIA for 2 h starting at 5pm, due to a signal system fault. The model has to be extended in order to handle a case like this. One solution is to add the following variables and constraints.

- \( g^m \) binary decision variable, =1 if there is a speed restriction on the segment between \( m \) and \( m+1 \) between \( t_0 \) and \( t_1 \). =0 otherwise.
- \( h^m_{1i}, h^m_{2i} \) two auxiliary binary variables handling time intervals for train \( i \).
- \( \tau_{40i}, t_0^i, t_1^i \) Running time with speed 40 from \( m \) to \( m+1 \) for train \( i \), starting time and stopping time for the speed restriction, respectively.

\[
\begin{align*}
g^m & \geq d_i^m + \tau_i^m - \Psi g^m h^m_{1i} \\
d_i^{m+1} & \geq d_i^m + \tau_{40i}^m - \Psi (1 - g^m h^m_{1i})
\end{align*}
\]

Figur 2: Original train graph
The algorithm gives three feasible plans in this case. The optimal plan is shown in figure 3.

![Figure 3: Optimal plan](image)

### 6 Conclusions

The model discussed so far has the potential to be part of a real-time CATD-system working under so-called normally disturbed conditions. This paper has shown that it is possible to adapt the optimisation model created by Jovanovic to handle some of the more important frequently occurring faults and disturbances in the daily train traffic control. We have also implemented a minor prototype system in order to, off-line, further evaluate and develop the model, and especially important, to be able to do that in co-operation with train dispatchers and planners.
The main issues that will be studied in the next step are: An evaluation of system proposed schedules through dispatchers and planners opinions and the implications on cost functions, models, and algorithms from the decision criteria actually used in Sweden. Future work will also concern solving the main deficiencies reported here, i.e., the train ordering and overtaking problem, the rigidity in number of tracks per meetpoint etc.

The final conclusion, based on our experiences, is that this model, even if implemented as in our simple prototype system, represents a rather complex system that, like for example a train traffic simulation system, has to be tested and put into operation in an iterative process. Hopefully it is a long lane that has no turnings.

References


