Optimal Reinsertion of Cancelled Train Lines

Julie Jespersen Groth, Jens Clausen

Danish State Railways (DSB), S-tog a/s, Production Planning, Kalvebod Brygge 32, 5., 1560 Copenhagen V, Denmark & Informatics and Mathematical Modelling, The Technical University of Denmark, jjespersen@s-tog.dsb.dk, jc@imm.dtu.dk

Abstract. One recovery strategy in case of a major disruption in a rail network is to cancel all trains on a specific line of the network. When the disturbance has ended, the cancelled line must be reinserted as soon as possible. In this article we present a mixed integer programming (MIP) model for calculating the best way to reinsert cancelled train lines in a rail network covered by a periodic timetable. Using a high abstraction level it has been possible to incorporate the temporal aspect in the model only relying on the information embedded in the train identification numbers of each departure. The model finds the optimal solution in an average of 0.5 CPU seconds in each test case.

1 Introduction to DSB S-tog

DSB S-tog (S-tog) is the operator of the city rail of Copenhagen, Denmark. Covering approximately 170 km double-tracks and 80 stations the city rail of Copenhagen services the inner and outer parts of Copenhagen. At a daily level the operator carries approximately 30,000 passengers. S-tog is the only user of the tracks, which are controlled by the infrastructural owner BaneDanmark (BD).

The S-tog network is formed by train lines covering the S-tog infrastructure by various compositions of routes depending on the timetable in use. Figure 1 shows the present line composition covering the network. The different parts of the network are called sections. There are 8 sections in the network; A central section, 6 fingers, and the circular rail. The lines merge in the central section as they intersect, and they de-merge as they re-enters the respective fingers according to their schedule.

The structure of the S-tog network implies that a high number of lines intersect in the central section. The trains on each line all run with a 20 minutes frequency. Given 10 lines intersecting the central section this means that within 20 minutes there is at most 2 minutes between each train in the central section i.e. there is a 2 minutes average headway in between the trains. Such low headway implies that even small delays can have a significant negative effect on a high number of trains.
1.1 Trains and Depots

Each line in the S-tog network is covered by 4 to 10 trains depending on the duration of the line circuit. Each train is covered by one or more train units. The coverage of a single train usually varies during the day in accordance with the expected passenger loads over the day.

At all times a train number is associated with each driving train. The train number is changed every time a train turns at a line terminal to run in the opposite direction. For each train there is hence a series of train numbers during the day defining the tasks of that particular train during that day. The train units used to cover the train may be changed completely during the day. The number series of a train is called the train’s train sequence. A train is defined by its train sequence and not by the train units covering it during the day. Figure 2 shows an example of a line covered by two trains where each train is covered by a train sequence and different train units during the day.

There is much information embedded in the train numbers. These are five digit numbers indicating the train line, stopping pattern, direction and time of day. The first two digits is the line identification. The specific value of the third digit represents the stopping pattern used on the line at that train number. If the digit is an even number the train is south going and if it is odd it is north
going. The fourth and fifth digits identify the time interval of day that the train number passes the central station, KH. For example, if the last two digits are 26, the train will pass KH at hour \( \frac{26}{3} \) = 8 as the third train in the 8th hour i.e. the train will pass the central station between 8:40 and 8:59. The time within the hour is decided by the last two digits in the train number modular 3, where 3 is number trains in each hour. The remainders possible are 0, 1 and 2 describing respectively the first, second and third train during each hour. Especially the information of time of day embedded in the train number becomes useful in the context of reinsertion.

Rolling stock depots are situated at the majority of line terminals and at the central station. The crew depot is located at the central station. Crew are transported from the crew depot to the rolling stock depots in situations where trains are started up, except for day start up where drivers are required to start their duties at the depot. The rolling stock depots serve as sinks and sources of the network when taking out or inserting train units.

2 Recovery Strategies

During the daily operation incidents occur that disturb the scheduled departures. These can be externally or internally caused incidents. Remedies can, to a certain extent, be employed to prevent internally caused incidents e.g. if the rolling stock is suspected to cause incidents by break downs, more frequent checks at the maintenance center can be planned.

The external incidents are harder to compensate for in advance. There are, though, the possibility of constructing all schedules of departures, crew an rolling stock with included buffer times at relevant places in the schedules and according to the expectations of delays. It is not necessarily evident where in the schedules it is optimal to locate the buffers. This information can be derived by e.g. simula-
tion studies based on real observations of the schedules, c.f. e.g. Hofman et al. [5].

Even though precautions are taken to minimize the effects of incidents, it will not be possible to avoid delays completely. Therefore, different recovery strategies have been developed for recovering the schedules of departures, crew and rolling stock.

Examples of recovery strategies are *Trains turned earlier on their route* (see figure 3), *Train set to stop at all stations can be made drive-through train* or *Entire train lines can be cancelled*. A drive-through train is a train that does not stop on specific smaller stations.

![Route without early turn around](image1)

**Fig. 3. Illustration of an early turn around**

Managing recovery is a joint task for the infrastructural owner and the operator using the tracks. In the S-tog network the infrastructural owner has the responsibility of the departures being processed and therefore the decision authority on this issue. The rail operator has at its disposal the resources crew and rolling stock and has therefore the responsibility of matching these resources to the demand defined by the departures.

When larger disturbances occur on the DSB S-tog network, the disturbed situation is often managed by taking out an entire S-tog train line i.e. all departures on a train line are cancelled. By taking out a train line more slack is created in the timetable, i.e. the headways are increased between train lines, which are time-adjacent according to the timetable. In this way buffer times are increased in the timetable and more room is created for absorbing the delays.

A take-out is carried out by shunting the rolling stock to depot tracks as the trains arrive at rolling stock depots. In the process of take-out it is normally not allowed to drive “backwards” in the network. A train can only drive forward to the next depot to be taken out. Therefore, the trains in the circuit of the line in
question ends up being distributed among the depots along the line according to where they were in the network when the decision of cancelling the line was made. It is crucial to realize that train units that are taken out at a depot are not necessarily used to cover the same trains when reinserted. Recall that a train is defined by its train sequence and not by the train units covering it.

3 The Reinsertion Problem

When an adequate level of regularity has been re-established in the operation, the cancelled train lines are reinserted according to schedule. The status of operation is evaluated by a train controller from the infrastructure operator. After the decision of initiating reinsertion has been made, the reinsertion should be carried out as quickly as possible under certain consideration regarding keeping order of trains.

When a train is reinserted it is transported as empty stock from the depot tracks to the platform. A train driver arrives on a running train from the crew depot at the central station, KH, according to a scheduled arrival. The train to be reinserted departs according to a scheduled departure on the relevant train line.

The reinsertion scheme is calculated by a rolling stock dispatcher from DSB S-tog. The reinsertion is presently scheduled for one train line at a time. It is necessary to decide which trains already in operation can transport train drivers to the rolling stock depots, where trains are inserted. The number of trains to be inserted from each depot is determined by the dispatcher, however, it is not given which train units at the rolling stock depots should be inserted to cover which trains in the schedule of the train line.

For the majority of lines, intermediate rolling stock depots exist along the line’s route. As for the terminal depots, it is determined how many trains must be inserted from the intermediate depots in total. There is, however, for each intermediate depot a possibility of inserting trains in both directions. Inserting in both directions decreases the finishing time of the reinsertion process.

The problem is now to decide when the reinsertion shall start on each rolling stock depot. This choice must be made taking into account the order of trains on several levels.

Firstly, if a reinsertion has begun from a certain rolling stock depot, the remaining trains to be inserted from that depot must be inserted in order according to frequency, so that there at no time occurs a frequency-interval with an uncovered departure. For example, at DSB S-tog the frequency is 20 minutes on all train lines. If 3 trains must be reinserted from Farum rolling stock depot and the first reinserted train departs at 15:18, then the remaining 2 trains must be reinserted and depart at respectively 15:38 and 15:58. Inserting the remain-
ing two trains at 15:58 and 16:18 would mean a vacant frequency interval at 15:38 i.e. order would not have been kept and that would be an illegal solution. Secondly, the order with respect to frequency must also be kept across rolling stock depots. After the initiation of reinsertion, the time between two adjacent departures on any station in the network must always be the frequency of 20 minutes.

One of the advantages of the reinsertion model is the solution time of the model compared to manual calculations. Also, it is possible to calculate a reinsertion plan immediately when the distribution of trains among depots is known after the take out. As the timetable is periodic the reinsertion scheme calculated will in principle be the same except for the exact train numbers that must be inserted. This might lead to some advantages with respect to coordinating the train driver schedules according to the reinsertion, thereby preventing reinsertion schedules being discarded because of the lack of drivers.

4 References

No literature has been found that resembles the exact problem of reinserting train lines. The problem is, however, similar to that of assigning resources to tasks and of making lines of work (also called pairings) for each member in a set of resources. The objectives differ in these problems. In the assignment of tasks to resources, the objective is most often to minimize the total cost of resources, see for example Gamache et al. in [4]. In the reinsertion problem we minimize the time of the final task performed.

Typically, general crew scheduling problems are exceedingly larger than the reinsertion problem. A well-documented approach for solving these large problems is by formulating the problem using a column generation model. This is done in a substantial number of papers. We refer to Jaumard et al. [5] and Eveborn and Rönquist [3].

The structure of the crew scheduling problems are also found in the vehicle routing problems, where routes among a known set of customers are found to be for a fleet of vehicles. The column generation approach for solving the vehicle routing problem with time windows was introduced by Desrosiers et al. [2] in 1984.

Resource scheduling problems are also often solved heuristically. In Cai and Li [1] and Wassan and Osman [8] different heuristics procedures are presented. Mason et al. [7] presents a method for personnel rostering that integrates integer programming procedures with heuristics and simulation.
5 Productions-data Regarding Train Numbers and Timetables

Because of the structure of the reinsertion problem, this can be solved given relatively little information. There are two types of data necessary in the model. The model must be built with background data based on the long term planning of timetable and rolling stock. When a reinsertion plan is needed in the operation, certain information of the real time situation is necessary.

From the long term planning it is necessary to know for each rolling stock depot how many trains (on the cancelled train line) can depart from the relevant rolling stock depot from the departure time of the first train that can transport drivers from the crew depot station to rolling stock depots, until the departure time of the first train that can be inserted from the relevant rolling stock depot. Furthermore, it is necessary to know the number of trains on the train line.

As mentioned in Section 1.1, each train in the train line covers a series of train numbers collectively forming a train sequence for each train. For the mathematical model it is only necessary to be able to differ between the trains. It is sufficient to make one calculation for each distribution of trains over depots. This is due to the periodic format of the timetable which implies that the solution to the reinsertion problem is generic (in that the structure is independent of the specific times given in the timetable). As there is only limited number of distribution of trains among depots, all solutions can easily be generated in advance and updated according to time of day in the real time situation.

In real time it is necessary to know for each rolling stock depot how many trains must be inserted from each depot. By comparing this information with the information of the first driver-carrying train, the rolling stock dispatcher can easily lookup the relevant solution.

For the rolling stock dispatcher it is also necessary to be able to identify the specific train numbers on the train line from each rolling stock depot. Additionally, the train numbers that will be used to transport the train drivers to the rolling stock depot must be identified. That is, the train numbers are synonymous with knowing the time of day of insertion. All train numbers relevant in the reinsertion process can be calculated using the train number of the first train that can transport train drivers to rolling stock depots.

The solution looked up by the rolling stock dispatcher is used to find the train numbers of respectively the trains to be reinserted and the trains to transport drivers.

6 A real life example

Two lines, H and H+, run on the route between Frederikssund (FS) and Farum (FM). When large disturbances occur involving the sections of this route, the H+ line is typically taken out. The 10 trains forming the line H+ circuit are
taken out on the terminal rolling stock depots, FS and FM, and on the intermediate depots of Ballerup (BA) and KH. Recall that the crew depot is at KH.

An example of distribution of the H+ trains over depots is that 2 trains are taken out on each of the terminal depots and 3 on each of the intermediate depots. One scenario of reinsertion is then that two trains must be reinserted from each terminal depot and three trains must be reinserted from each intermediate depot where insertion is possible in both directions.

Figure 4 illustrates the driver-carrying trains. Line a and d are the first trains going respectively south and north that can bring out drivers to depots. As the first driver-carrying train (in each direction) passes each depot the reinsertion at the depots can be initiated.

![Figure 4](image).

**Fig. 4.** The straight lines a, b and c illustrates driver-carrying trains going south and the lines d, e and f those going north. The lines initiate reinsertion at the different depot as they pass them.

In the reinsertion model the initiation time of reinsertion is counted in integral time slots. It is counted how many trains on the train line in question was planned to leave the depot from the decision of reinsertion until the first driver-carrying train reaches the depot. In Figure 5 there are 2 trains originally planned to leave the FS depot before reinsertion can begin.

The exact approach of a reinsertion is illustrated in Figure 6. For each of the figures a) - d) trains are inserted from a depot. Observe that order is kept at all time. There is no vacant frequency intervals at depots and there are no stations where passengers experience vacant frequency intervals. Illustrated in red on a), b) and d) is the driver-carrying trains transporting drivers for the reinsertion.
Fig. 5. Illustration of the reinsertion start time at the FS depot. From the decision of reinsertion until the reinsertion can begin at the FS depot, there are two scheduled trains that can not be reinserted as the driver-carrying train has not yet reached the depot.

7 The Mathematical Model

The goal of the model is to decide which train, \( i \in I \) should be inserted from which depot, \( k \in K \). Each originally scheduled train \( i \) (before take out) must be covered with train units and hence reinserted in operation according to schedule. Also it must be decided for each train in which time slot \( j \in J \) the reinsertion will be carried out.

The model decides which trains will run but it does not consider which train units to use to cover the trains. It is assumed that the information of distribution of train units across depots is provided as input and thereby sufficient in number to cover the trains.

The variables representing which train to be inserted from which depot and when are binary:

\[
x_{i,j,k} = \begin{cases} 
1 & \text{if train } i \text{ is inserted in time slot } j \text{ from depot } k \\
0 & \text{otherwise}
\end{cases}
\]

If we consider the trains to be inserted as tasks and the depot as a resource, the problem strongly resembles an assignment problem of assigning crew to tasks. Each task in this case must be covered once. This is guaranteed by the partitioning constraints (1):

\[
\sum_{j,k} x_{i,j,k} = 1, \quad \forall \ i \in I
\]  

(1)

Equations (2) are included so that no time slot for a depot or train is covered more than once:
In the model it is known how many trains are to be inserted from each depot. Therefore, binding constraints exist for each depot. They differ for respectively terminal and intermediate depots. As the trains are inserted only in one direction at the terminal depots, \( k \in K_T \), the binding constraints for these depots are:

\[
\sum_{i,j} x_{i,j,k} \leq 1, \quad \forall j \in J, k \in K
\]

(2)

As mentioned earlier the speed of insertion is increased when insertion on intermediate depots are made in both directions. In the model it is chosen for each intermediate depot to insert half of the trains in one direction and the other half in the other direction. When an odd number of trains is to be inserted from a depot, the model decides how many must be inserted in each direction. This is handled in the model by including two depots for each intermediate depot. The set of intermediate depots is denoted \( K^I \). It is constructed by sets of two depots together denoting one intermediate depot where reinsertion can be carried out in \( l \) directions, \( K^I = K_1^I \cup \ldots \cup K_l^I \), where \( l \in L \). \( L \) is the set of directions, which in the S-tog network for all depots is north or south. The total set of depots is \( K = K^T \cup K^I \).

The sum of trains inserted in both directions should equal the total number of trains to be inserted from the intermediate depot. Current practise is that half of the trains are inserted from the intermediate depot in one direction and
the other half in the opposite direction. Equations (4) ensure that the number of trains inserted in each direction is the total number of trains to be inserted divided by 2. If an odd number of trains is to be inserted, the result is rounded up or down to nearest integer depending on which is more favorable to the model. See equations (5) and (6).

$$\sum_{i,j,k} x_{i,j,k} = \sum_k D_k, \quad \forall \quad l \in L, k \in K^l$$  \hspace{1cm} (4)  

$$\sum_{i,j} x_{i,j,k} = D^I_k, \quad \forall \quad k \in K^I$$  \hspace{1cm} (5)  

$$D^I_k \geq \lceil \frac{D_k}{2} \rceil \quad D^I_k \leq \lfloor \frac{D_k}{2} \rfloor$$  \hspace{1cm} (6)  

It is crucial that certain orders are kept as the trains are inserted. As mentioned in Section 3 order should be kept within depots and between depots. Also, reinsertion must not begin on a depot before a train driver can arrive from the crew depot to drive the train to be reinserted.

To assure that each train is inserted only once, it is necessary to take into consideration the train sequences of each train describing in which time slot each train is at the different depots. To handle this a constant is introduced, \(in_{i,j,k}\).

\[ in_{i,j,k} = \begin{cases} 
1 & \text{if train } i \text{ may depart from depot } k \text{ in time slot } j \\
0 & \text{otherwise} 
\end{cases} \]

It is not possible to insert a train from a depot, if it is not there at that specific time slot. We refer to this as the order between stations and it is assured by equation (7).

$$x_{i,j,k} \leq in_{i,j,k}, \quad \forall \quad i \in I, j \in J, k \in K$$  \hspace{1cm} (7)

To model the order within stations we introduce two sets of integer variables, \(start_k\) and \(end_k\). Also, we introduce equations (8) to (11). Equations (8) connect the start and end variables. Equations (9) assure that reinsertion is not begun before the first driver can arrive at the depot. For the purpose of determining this a constant, \(C_k\), is given. The constant indicates how many trains has been scheduled at depot \(k\) from the time of the decision of reinsertion until drivers are able to reach the depot, cf. Figure 5. Equations (10) and (11) ensure that when a reinsertion has begun on depot, it is carried out continuously in adjacent time slots.

$$start_k + \sum_{i,j} x_{i,j,k} - 1 = end_k, \quad \forall \quad k \in K$$  \hspace{1cm} (8)

$$start_k \geq C_k + 1, \quad \forall \quad k \in K$$  \hspace{1cm} (9)

$$start_k \leq j + M \cdot (1 - x_{i,j,k}), \quad \forall \quad i \in I, j \in J, k \in K$$  \hspace{1cm} (10)
\[ end_k \geq j - M \cdot \left(1 - x_{i,j,k}\right), \quad \forall i \in I, j \in J, k \in K \] (11)

Much of the information of the timetable and departures is embedded in the train numbers. The periodic form of the timetable supports this formulation. The train numbers to be inserted when \( x_{i,j,k} \) is 1 is calculated from an initial train number on a train able to carry train drivers to the depots and some constant describing the relationship between the train numbers on the driver-carrying line and the line to be reinserted. The train number is adjusted according to the time slot in which it is to be inserted. See equation (12).

\[
TrainNumber_{i,j,k} = \left(InitialTrain + \text{TrainConst}_k + j\right) \cdot x_{i,j,k}, \quad \forall i \in I, j \in J, k \in K
\] (12)

The objective of the model is to reinsert as quickly as possible. This is assured by an objective function of minimizing the maximum inserted train number, \(\text{MaxTrainNumber} \). As information of time of day is embedded in the train numbers this equivalent with minimizing the latest time of reinsertion. This is achieved by minimizing the maximum value of the last two digit number in the train number.

\[
\text{Minimize } \text{MaxTrainNumber}
\] (13)

\[
\text{MaxTrainNumber} \geq \text{TrainNumber}_{i,j,k}, \quad \forall i \in I, j \in J, k \in K
\] (14)

8 Computational Experience

The running time of the model is not relevant in real time as reinsertion schemes are generated in advance and looked up at the relevant time. Test results do, however, show that the running time of the model on average is only approximately 0.5 CPU seconds, i.e. the model solves the problem in real time for the relevant problem instances. The real-time approach is not chosen partly due to software license issues, partly due to the generic nature of the reinsertion schemes.

9 An Improved Planning Process

The initial request for a tool for calculating reinsertion was made by the rolling stock dispatchers themselves. They regarded the problem of creating reinsertion schedules by hand as complicated and time demanding. First a tool was made that was not based on the principles of MIP. It was merely a spreadsheet calculating the reinsertion plan from basic knowledge of the distribution of trains,
the first driver-carrying train and a large set of if-then-else-loops. The project of creating an optimization model for calculating the reinsertion was started mainly due to the quite complicated task of updating the initial reinsertion tool. The mathematical model of the reinsertion problem has been implemented in Gams and solved in Cplex. The initial reinsertion model is still in use. It will be replaced by the MIP model during the autumn 2006.

Solutions are generated with the MIP model for all possible scenarios of distributing trains over rolling stock depots. The solutions are then stored and the rolling stock dispatcher can look up the solutions via a spreadsheet.

The reinsertion model has increased the level of service offered to the passengers. Earlier, when the rolling stock dispatcher had to make the calculation of reinsertion by hand, the solutions where either not generated because it took too long time to calculate a solution, or a solution was generated with a longer total reinsertion time than the optimal solution. In the first case the train lines would remain cancelled for the remainder of the day.

Besides the passenger service improvement, the reinsertion model decreases the level of stress for the rolling stock dispatchers. Solutions can be generated immediately to satisfy the demand of the train controllers in charge of the reinsertion decision. This has rendered a more efficient planning process with resources left for other tasks.

Also the maintenance of the model has been eased to the satisfaction of the analyst updating it. The MIP model is easy to update according to a new periodic timetable. This is done simply by changing the set of constants presently used in the model and generating a new set of solutions.

10 Further Developments

Presently the reinsertion model is used only for scenarios where a cancelled line needs to be inserted into a running operation, in which running trains can transport drivers to rolling stock depots. Future developments on the model will enable complete startups where trains can be inserted as the first on their route. This requires that train numbers are calculated from virtual train numbers.

The number of trains to be reinserted on each depot is input to the model. Occasionally the number of trains available for operation on the depots is larger than the number of trains that has been taken out. It seems very relevant to change the model to account for this fact in such a way that the model decides the exact number of trains to be inserted from each depot, ensuring that the total number of trains reinserted is the number of trains needed to cover the line.
The present model works with a distribution of trains reinserted in each direction on intermediate depots of half reinserted in one direction and the other half reinserted in the other direction. Further developments on the model involves changing the constraints (6) to enable solutions where the number of trains reinserted in each direction on each intermediate depot are decision variables.

At some of the routes of the S-tog network more than two lines cover a route simultaneously. A relevant recovery scenario is that more than one line is cancelled along the route. It would be relevant to adjust the reinsertion model so that it can coordinate and give the results for the reinsertion of more than one line at a time. In the model this would mainly mean modification of input.

The quality of solutions generated by the reinsertion model strongly depends on the distribution of trains on depots. At the time of disruption, take out is carried out with no regards to a later reinsertion. There is no time for rearranging trains at depots so that the best possible reinsertion is possible at a later time. There will though be the possibility of making small changes in the take out plan. For example, it might be possible to drive forward to next-closest depot instead of taking it out of operation at the closest depot. The evaluation of which is best can be made by the reinsertion model.

11 Conclusion

Because of the information embedded in train numbers the MIP model has been made at a high level of abstraction. Regardless of this, the reinsertion model can be used in operation in its current form.

The MIP model is easy to update according to a new periodic timetable. This has decreased the possibility of the rolling stock dispatchers having to wait for an updated version.

The solutions generated with the reinsertion model always generates optimal solutions with respect to the latest inserted train. This was not often achieved when the reinsertion was calculated by hand. Frequently, the train line remained cancelled for the remainder of the day.
References