Dynamic re-scheduling of trains after disruption

C. J. Goodman & R. Takagi
Electronic, Electrical and Computer Engineering,
The University of Birmingham, UK

Abstract

Railways require enormous investment in infrastructure and it is therefore imperative that means are sought to maximise the utilisation of that infrastructure. This paper reviews some of the applications of computers to the particular problem of recovering from traffic disturbances in ways which seek to minimise the consequent loss of capacity. To some extent, the processes of re-scheduling after disruption are similar to scheduling in the first place, and thus some reference will be made to methods being used for automatic time-tabling.

The management of recovery can vary between entirely manual and entirely automatic, with full predictive modelling and rigorous mathematical optimisation procedures. In practice, such systems as already exist, and those under research and development, are generally something of a compromise between these two extremes. Even so, a key issue for system designers is whether to allow the computers to control the traffic directly without reference to the line controllers, or whether they should only operate in the decision support role. Again, practical experience is leading towards systems incorporating both concepts and using one or the other depending on the severity of the disruption.
Keywords: train control, disruption, recovery.

1 Introduction

The power of modern processing and communication electronics makes the concept of using computers to take complete control of substantial areas of a railway network entirely plausible. The technological development through computer-based interlockings to automatic dispatching and traffic regulation is manifest and examples of all these levels of control exist. Not surprisingly, the

Computers in Railways IX, J. Allan, C. A. Brebbia, R. J. Hill, G. Sciutto & S. Sone (Editors)

highest levels of automation presently exist on metro-type systems, where the simple routing, identical trains and regular service patterns make it easier to achieve. Part of the reason for this is that the objectives and criteria for successfully regulating a metro are generally clear, well-understood and attract a substantial measure of agreement from different operators. For mainline networks, operators find it less easy to define clear objective criteria which encapsulate the most desirable features of a recovery strategy. This is partly due to complexity, which also causes problems with performing and implementing the control decisions in real time.

This paper collects together a few ideas for setting criteria, including a brief review of some previous work, with the intention of stimulating debate about just what we are trying to achieve with real-time regulation systems. The paper will also make some comments on the various types of control system that could be employed in an attempt to satisfy these criteria. Systems can vary from those which simply implement fixed rules when given situations occur to those which do full scale predictive simulations to find the best solution in the circumstances. For mainline networks, it may well be the case that, in practice, complexity still precludes the use of detailed modelling and searches, even with the most powerful processing currently available.

2 Types of system

2.1 Types of railway

There are a variety of specific types of railway system which have been or are currently the targets for research on dynamic control. These include: -

- Metros suffering platform delays and attempting to regain timetable or, more likely, regular headway;
- Converging junctions as part of a mainline railway where delays to a passenger train, for example, eventually lead to a need to decide whether to allow a freight train into a section ahead of it;
- Freight railways, where timetables are less important than maximising capacity utilisation. This tends to involve short-term re-scheduling, rather than actually responding to disruptions in real time, but is clearly a similar process. Constraints on finding the best solution are often due to the railway being single track with passing loops;
- Interchange of passengers at stations where incoming delays raise questions about delaying outgoing services to reduce missed connections, and;
- General networks, or sub-networks, where recovery to timetable is the usually accepted criteria. Inevitably, there has to be some limit on the size of the network considered and on the level of detail that is modelled; however, heavily used networks are known to exhibit properties where disruption can ripple through a large area and this needs attention, if at all possible.

2.2 Types of criteria

Instinctively, railway operators will state that the key criterion for successful operation is running to timetable and that, therefore, any recovery from disruption should be managed in such a way as to return to schedule as quickly as possible. Whilst this is evidently the highest objective in most situations, it is not the only possibility and some or all of the following criteria might be used in various combinations: -

- Regaining schedule as quickly as possible. Usually interpreted as recovery to timetable whilst minimising the loss of train-minutes during the recovery;
- Regaining schedule whilst minimising the loss of passenger-minutes. In some respects, this is a more basic criterion, but not easy to measure with existing technology;
- Recovery to regular headway on a metro whilst minimising accumulated headway deviation in train-seconds;
- Recovery to regular headway, whilst minimising passenger waiting times, journey times and reducing congestion in the cars, and;
- Maximising capacity utilisation over a defined period into the future (a few hours, say). Again, rationally, this should be based on tonnage or passengers moved rather than train-paths, but except on dedicated mining railways, for example, it generally resolves to train-paths.

It is clear from the last of these examples that the criteria to be applied are not only a function of the type of railway being considered, but also a function of the type of disruption, and especially its severity. It has been recognised for some time in the metro platform delay case, for example, that, for all but very small delays (where the make-up available by dispensing with coasting could be used), it is impossible and undesirable to try and recover to absolute timetable. Instead, it is better to move towards regular headway. This also highlights issues of how the regulation is to be implemented, as this again depends on the severity of the disruption.

2.3 Types of implementation

Ideally, any computer-based regulator would control the trains directly, without reference to the line controllers. This presumes that control is available in terms of setting signal aspects and/or providing instructions to depart from platforms, etc. Increasingly, especially on metros, it will be possible to have virtually continuous control of all trains, which would enable a better solution to be applied. There is an issue, however, of whether the computer literally drives the trains, or whether it advises the drivers. From the view-point of conducting theoretical studies, it has to be assumed that the trains will actually do what the computer requests them to do, and in that sense, the detail of who or what drives the train is immaterial.

Computers in Railways IX, J. Allan, C. A. Brebbia, R. J. Hill, G. Sciutto & S. Sone (Editors)

A similar situation pertains in the control centre, where it is technically possible for the computer, having calculated an optimal recovery strategy, to directly issue commands to the trains or signals. Alternatively, it can advise the line controllers what it has calculated as the best strategy and leave to them the final decision as to whether to accept it or implement a manual solution. In this latter role, the computer-based optimisation system would be referred to as a “Decision Support System”. Observation of existing systems suggests that for small disruptions (for example, platform delays on metro or commuter services of up to a few minutes), operators are happy to let the computer take direct control. For larger disturbances, however, especially where trains need to be taken out of service, it is universal to require manual authority, albeit often just to accept a re-scheduling that has been worked out in real time by the computer.

2.4 Types of modelling and solution methods

The most basic type of controller essentially implements a set of rules when it recognises a particular situation has occurred. These rules are derived form operational experience and, in some measure, already exist in the ‘rule books’ or in the acquired skills of the control centre staff. This type of controller is generally referred to as an ‘expert system’ or sometimes an ‘heuristic system’. The Automatic Route Setting (ARS) systems available in modern control centres often effectively incorporate such “expert systems’. The detail of the rules they contain is worked out by consultation with the infrastructure controller and the train operators using the relevant area of the railway. The rules can include fairly sophisticated junction control, wherein the operators can specify their own values of the weighting factors applied in equations which will allocate junction priorities according to lateness and significance of particular trains.

At the opposite extreme of complexity, the control system can consist of a dynamic model of the railway operation, a mathematically formulated objective function and some sort of search procedure to find the set of operating instructions that optimises the objective function. This makes a number of assumptions:

- It is possible to define a suitable objective function
- It is possible to derive a computer model of the railway in sufficient detail to represent the key interactions which characterise the behaviour under disruption
- Controls exist that will impact on the disruption
- A search procedure can be found that will move reliably towards the optimum

If these assumptions can be satisfied and, furthermore, the computational search can be achieved much faster than real time, then the optimising controller can operate on-line and in real time. This requires the optimum solution to be found within only a few seconds of computing time, when looking ahead to optimise the railway operation over a forward period of a few hours, say. This

Computers in Railways IX, J. Allan, C. A. Brebbia, R. J. Hill, G. Sciutto & S. Sone (Editors)

period is referred to as the ‘optimisation period’, or sometimes ‘optimisation window’. Intrinsically, this requires repeating the simulation of the train operation over the look-ahead period for thousands, or even tens of thousands, of times with small test changes in the control inputs (starting times, for example) until the best solution is found. The organisation of this search requires the use of some established optimisation procedure; the choice must recognise that this particular problem is highly non-linear and heavily constrained. If at all possible, it is preferable to use gradient-based search methods. In these methods, partial derivatives of the objective function with respect to the control inputs are calculated about some initial set of values. These partial derivatives are then used to produce new values of the control inputs by stepping in the direction and by an amount indicated by the vector of derivatives.

Alternatively, the search can be guided by methods such as the Tabu search or by Genetic Algorithms which use concepts of eliminating ‘unfit’ candidate solutions. These methods are inevitably slower than the gradient-based methods. The other factor that affects the overall speed of the solution is the speed of the model itself. Various levels of model are used, including at least:

- Self-written sets of logic and time-of-flight statements that reflect routings and train paths available through a network
- Similar conceptual approach, but implemented via some standardised software such as Petri Nets
- Full train movement simulation, with dynamic mechanical calculations (i.e. not pre-stored trajectories) and full representation of the signalling system and the interactions between the trains.

For the purposes of optimising recovery from disruption, it is essential to model abnormal running and signal interaction, which means that the third type of model above is perhaps the only approach which can satisfy the requirements. Some models have been used which try to predict the various possible modes of disruption and have stored trajectories ready to deal with particular cases. This can result in considerably faster modelling, but is limited when something that was not anticipated occurs.

Eventually, within a few seconds of computer time, an optimum solution should be achieved which gives the values of the control variables to be applied to the railway over the optimisation period as it now evolves in real time. In practice, this is not actually done as it takes no account of new disturbances arriving into the system or of the inevitable inaccuracies that will exist in the solution due to modelling and measurement noise. A technique referred as the ‘sliding window’ can be used to reduce these problems. Here, only the first part of the calculated solution is actually applied, and a short time later, the whole process is repeated. The repeat interval would be of the order of 5 or 10 seconds. An unfortunate aspect of this technique is that it significantly increases the amount of calculation required as most of each generated solution is in fact thrown away. Some of this effect can be mitigated by starting each new optimisation from the optimum values calculated previously (recursive solution).

Computers in Railways IX, J. Allan, C. A. Brebbia, R. J. Hill, G. Sciutto & S. Sone (Editors)

Operated in this manner, the controller takes on some of the properties of a true closed-loop controller.

It may well be the case, that even with the most carefully constructed model that provides the necessary functionality most efficiently and with the fastest search procedures, it is still not possible to perform the optimisation calculations fast enough for them to be embedded in a real-time controller. An approach mid-way between the two extremes so far discussed is to do the full optimisations off-line and encapsulate the results in some form of expert system rules which can then be applied in real time to the actual railway. The encapsulation could be achieved by neural networks learning the patterns that presumably exist in the control actions made in response to particular disruptions. The main problem is the same as with the partly pre-stored simulations; all the possible disruptions have to have been anticipated in the off-line data preparation.

3 Some examples of existing systems and current research activity

3.1 Freight railways

Freight traffic, as compared to passenger trains, has the following characteristics:

- The required punctuality is generally much lower, especially when transporting bulk goods, and;
- There is less necessity of interconnection between trains.

These conditions make it far easier to optimise operation of freight railways than passenger railways, and therefore a number of implementations of highly automated decision support systems already exist.

Some earlier examples date back to the beginning of the 1990s, in which either real time locomotive distribution [1] or service planning [2] were automated by software tools. These incorporate online optimisation, in which costs of operating the locomotives to carry given tonnage of freight are minimised.

3.2 Scheduling techniques for passenger railway networks using simplified models

Generally speaking, it is more difficult to deal with re-scheduling than with scheduling. This is because re-scheduling requires decision-making in a very limited amount of time, in response to a situation which is generally unknown, as discussed in section 2.4. However, because of the inherent similarity in scheduling and re-scheduling, both will be reviewed.

Optimisation requires some sort of evaluation of the schedule created. However, the evaluation of the schedule across the whole rail network is not always easy, and requires careful analysis. A group of researchers in Holland [3],[4],[5] has developed a system incorporating an evaluation algorithm based

Computers in Railways IX, J. Allan, C. A. Brebbia, R. J. Hill, G. Sciutto & S. Sone (Editors)

on "Max-plus algebra", which is a mathematical technique which makes comparisons between large numbers of variables easy. Take the example in Figure 1, in which services for Route 1 and 2 arrive at the interchange at times t_1 and t_2, respectively, and a service for Route 3 is to depart at t_3. The required connection time between Routes 1 and 3 is given as c_{13}, and between Routes 2 and 3 as c_{23}. For passengers on both Routes 1 and 2 to have connection to Route 3, the following equation must hold:

$$t_3 = \max\left(t_1 + c_{13}, t_2 + c_{23}\right)$$

Here, define "max" as $\oplus$, and the normal addition (+) as $\otimes$. Then we get:

$$t_3 = t_1 \otimes c_{13} \oplus t_2 \otimes c_{23}$$

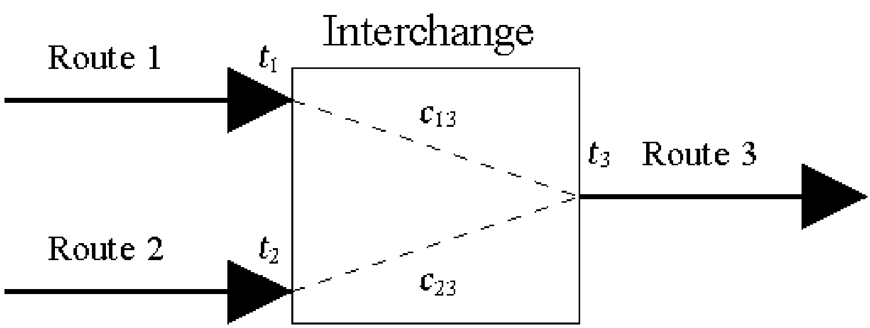

Figure 1: Connection at an interchange.

This algebra, combined with a graphical representation showing connections between routes that form an entire network, can be used to evaluate the minimum possible cycle in which the whole network can be operated with ideal connection at transport hubs; the minimum is defined by the maximum of "loop travel time", in which a passenger can travel around a small loop in the network.

In another example, Isaai *et al* [6],[7],[8]. conducted their work on a simplified model, based on the Iranian railway network, which is mostly single-track. Hence, among the constraints was a requirement for one block section between adjacent stations. The research group proposed to first use the algorithm to search for solutions that satisfy the constraints, then do the optimisation, using algorithms such as Simulated Annealing or Tabu Search.

The works discussed above use simplified models for train movement, which is necessary from the point of view of handling the whole, or at least a large part, of the network in the systems being considered. As pointed out in section 2.4, this could result in an inability to find the true optimum when events occur which were not anticipated in the design of the model.

Computers in Railways IX, J. Allan, C. A. Brebbia, R. J. Hill, G. Sciutto & S. Sone (Editors)

3.3 Re-scheduling

As stated in section 3.2, re-scheduling requires rapid decision-making in response to any incidents that occur during operation. The systems currently in use, including ARS (Automatic Route Setting) [9] in the UK, have integrated train regulatory functions. In the case of ARS, these include (1) detection and resolution of train conflicts, and (2) determination of train priority using predictive methods. The latter includes consideration of:

- Minimising additional train delay;
- Train class weighting;
- Tidal flow weighting;
- Timetable order margin, and;
- Maximum waiting time

As an example of purely academic research, a research group at the University of Tokyo has looked at re-scheduling of a double-track railway, without diverging or converging routes, by modelling the schedule using graph theory [10]. This fast algorithm enables the prediction of train movements once the topology of the schedule on the time-distance graph has been determined. An evaluation based on a simplified cost function, which derives from delay-minutes and passenger numbers on the train, is done to compare different topologies to finalise the proposed regulation against one particular disruption incident.

'Bunching' is commonly seen on metro-type railways, and the control algorithm to deal with it has been well studied, and even integrated into some control centres. Goodman and Murata [11] took a fresh approach to the problem by combining the passenger delay minutes and energy consumption into a new evaluation function to be optimised.

Junction optimisation is another problem that has been studied. Ho *et al.* [12] studied optimum regulation of trains on two converging routes passing through a junction as quickly as possible, by using methods such as dynamic programming or genetic algorithms.

Maintaining connection at interchanges is identified as a problem in section 2.1. Takagi *et al.* [13],[14] proposed a simplified model using coloured Petri Nets that can be used to model passenger flows at these interchanges.

4 The future - tracking individual passengers?

In situations where the timetable has been disrupted, the current policy of many railway undertakings is to "maintain connections as advertised", and keep delays to a minimum across the network. From the passengers' perspective, it generally suffices if they can reach the final destination with the minimum of delay, and as long as appropriate information is given to them, they would not mind any changes of routing. Many Japanese rail operators have passenger services sections, which are integrated as part of their train control centres. They deal with passengers (generally speaking, long-distance ticket holders with reservations) who have lost, or are going to lose connections, because of disruptive incidents and subsequent line control actions. However, individual passengers are only imperfectly connected to this system in that they can only get information from members of staff and discrete information displays at stations.

Advances in microelectronics and wireless communication will make it possible for the system to provide each passenger with individually tailored guidance. Such a proposal, based on technologies such as smart cards, PDAs (personal digital assistants) and mobile phones, has been made by the research group at the University of Tokyo [15],[16]. The system can effectively track a passenger, from the start to the end of his/her trip, which means information as to where each passenger is and where and when s/he came from and is heading to will be available any time and in detail. These sorts of technologies will enable collection of statistical and other data on passengers in real time. The system will then be able to conduct train controlling based on accumulated passenger delay-minutes, rather than train delay-minutes as currently happens.

5 Concluding remarks

This paper has reviewed the general concepts and procedures being applied to the control of train services following disruption. Systems vary from simply encapsulating known rules for recovery to using full scale simulation and optimisation embedded in an on-line real-time controller. The computer-generated control inputs can be applied directly to the railway or used to advise the human line controllers. Systems employing stored algorithms are already in operation, often taking direct control for small disturbances but handing back to the line controllers when the disturbances exceed a set level.

The authors believe that this trend will continue and greater use will be made of embedded controllers as the only real way of coping with all manner of disruptions. Further developments in computing and communication capabilities will indeed eventually make it possible for such on-line control systems to optimise the recovery from disruption in terms of the passengers' own benefits, as opposed to just running a tidy timetable.

References

[1] Chih, K.C., *et al.*: Implementation of a Real Time Locomotive Distribution System, *Proc. Second International Conference on Computer Aided Design, Manufacture and Operation in the Railway and other Advanced Mass Transit Systems (COMPRAIL 90)*, Rome, Italy (1990).

[2] Van Dyke, C. and Davis, L.: Software Tools for Railway Operations / Service Planning: the Service Planning Model Family of Software, *ibid.*

[3] Goverde, R. M. P., The max-plus algebra approach to railway timetable design, *Computers in Railways VI (Proc. COMPRAIL 98)*, Lisbon, Portugal (1998).

Computers in Railways IX, J. Allan, C. A. Brebbia, R. J. Hill, G. Sciutto & S. Sone (Editors)

[4] Middelkoop, D., *et al.*, Train network simulator for support of network wide planning of infrastructure and timetables, *Computers in Railways VII (Proc. COMPRAIL 2000)*, Bologna, Italy (2000).
[5] Soto y Koelemeijer, G., *et al.*: PETER, a performance evaluator for railway timetables, *ibid.*
[6] Isaai, M. T., *et al.*: Hybrid Applications of Constraint Satisfaction and Meta-Heuristics to Railway Timetabling: A Comparative Study, *IEEE Trans. Syst., Man, Cybern. C,* vol. 31, pp. 87-95 (2001).
[7] Isaai, M. T., *et al.*: Predictive and Reactive Approaches to the Train-Scheduling Problem: A Knowledge Management Perspective, *IEEE Trans. Syst., Man, Cybern. C,* vol. 31, pp. 476-484 (2001).
[8] Isaai, M. T., *et al.*: An Object-Oriented, Constraint-Based Heuristic for a Class of Passenger-Train Scheduling Problems, *IEEE Trans. Syst., Man, Cybern. C,* vol. 30, pp. 12-21 (2000).
[9] AEA Technology: Automatic Routesetting Concepts and Application, *presented at IRSE Spring Technical Visit* (2003).
[10] Takano, M., *et al.*: Computer Assisting System to Propose and Evaluate Train-Rescheduling with a function of Passenger-Path Allocation (in Japanese), *Proc. of 10th Joint Railway Technology Symposium 2003*, 1412, Kawasaki, Japan (2003).
[11] Goodman, C. J. and Murata, S.: Metro traffic regulation from the passenger perspective, *Proc. Instn. Mech. Engrs,* vol. 215, Part F, pp. 137-147 (2001).
[12] Ho, T. K., *et al.*: Railway junction traffic control by heuristic methods, *IEE Proc. – Electr. Power Appl.,* vol. 148, pp. 77-84 (2001).
[13] Takagi, R., *et al.*: "Modelling Passenger Flows at a Transport Interchange using Petri Nets", *Proc. Instn. Mech. Engrs,* Vol. 217, Part F, pp. 125-134 (2003).
[14] Takagi, R., *et al.*: Optimising Departure Times at a Transport Interchange to Improve Connections When Services are Disrupted, *COMPRAIL 2004* (to appear).
[15] Takagi, R., *et al.*: IPASS (Intelligent Passenger Assistance for Public Transport) (in Japanese), *IEE-Japan Technical Meeting on Transportation and Electric Railways*, TER-97-50, Osaka, Japan (1997).
[16] Takagi, R., *et al.*: IPASS – Intelligent Passenger Assistance System for Public Transport, *presented at International Conference on "Developments in Mass Transit Systems"*, London, UK (1998).

Computers in Railways IX, J. Allan, C. A. Brebbia, R. J. Hill, G. Sciutto & S. Sone (Editors)