OBJECTIVE FUNCTIONS FOR ADAPTIVE RAILWAY TRAFFIC MANAGEMENT BY APPROXIMATE DYNAMIC PROGRAMMING

Taha Ghasempour[a], Benjamin Heydecker[a], Taku Fujiyama[a]
[a] *Centre for Transport Studies, University College London*

BACKGROUND

Regulation of railway traffic aims at ensuring safe, punctual and energy-efficient train operations. The usual basis for this is the timetable (the plan of train departures, passing times and train arrivals). In the presence of disturbances caused initially by delays, train dispatchers rely on their own experience to modify the timetable to create what is known as a *working timetable*, while the efficiency of this is often unknown. In such cases deploying automatic rescheduling tools have been shown to improve performance by limiting delays* and returning operation to the planned timetable as quickly as possible (Cacchiani et al., 2014).

In practical implementations of rescheduling systems, important matters are quickly to compute an appropriate dispatching measure to be disseminated to affected trains and junction controllers). The objective of the present research is therefore to develop a control framework that responds to disruptions in a way that is adapted according to the current operating conditions. The approach developed here is based on approximate dynamic programming (ADP), which can accommodate the complexity of railway operations and hence can optimise railway traffic quickly and reliably.

The effectiveness of an ADP formulation depends on specification of the objective function, and within this the choice of variables to describe the state of the railway system. In the present case of railway rescheduling, there are numerous candidates. This choice is dependent on the requirements and constraints of railway operators and infrastructure managers, where in many cases these include mutual inconsistencies. Therefore, we focus our evaluation on the performance of the ADP formulation and the effects of different objective functions according to a range of measurements including and exceeding their primary objective.

ADAPTIVE RAILWAY TRAFFIC CONTROL

We consider sequence controls at the time each train enters a designated control area: each such event corresponds to a stage of the dynamic optimisation. Let $s_n \in S$ be a vector of state variables of the system at stage n, $u_n \in U$ the decision variable for stage n, and g the cost during a single stage. Given the initial state s_0 and a sequence of decisions u, a future-discounted dynamic programme is:

$$\min_{u \in U} E \left\{ \sum_n \alpha_n g(s_n, u_n, s_{n+1}) \mid s_0 \right\} \tag{1}$$

where $s_{n+1} = f(s_n, u_n, v_n)$, v_n represents arrivals during stage n, and $f(.)$ represents the state transition (plant) function.

* *Initial delays that arise in traffic, e.g. longer than planned dwell times, are defined as primary delays and delays that are propagated to other trains from the primary delays are defined as consecutive delays. The combination of these delays is defined as total delays.*

This is addressed by calculating recursively:

$$J(s_n) = \min_{u_n \in U} E\,\{g(s_n, u_n, s_{n+1}) + \alpha_{n+1}\, J(s_{n+1}\,)\} \tag{2}$$

where J is the value of future discounted costs that depends on the state and, implicitly, optimal future decisions, α is the discount factor that here is calculated as $\alpha_n = e^{(-n\theta)}$, and E denotes the expectation over stochastic variation. In the present case of railway traffic control, the state s_n is a combination of traffic state and control state at stage n and the controls are decisions for the sequence of mutually incompatible trains.

Although equation (2) represents an elegant way of optimising railway traffic, it can be computationally expensive and therefore is not practical for operational use. The reason for this is the need to consider many future states to evaluate the optimal decision at each stage. Furthermore, states are increasingly uncertain into the future. Together, these lead to increasing volumes of calculation that has decreasing certainty and relevance to the current decision with time into the future. Therefore, we reduce the number of states considered by replacing the true value $J(.)$ with an approximate value function $\hat{J}(.)$ to make the computational requirement polynomial to the number of state variables, rather than being exponential to the size of state space. We employ a separable linear approximation function, which can be expressed as:

$$\hat{J}(s, r) = r^T . \phi(s) \tag{3}$$

where $r = (r_1, r_2, \ldots r_k)^T$ is a column vector of coefficients and ϕ is a features extraction function (or basis function) defined on the state space S and so maps the state to a feature vector for which r is the associated weight vector. Hence at each stage n we calculate:

$$u_n^* = \underset{u \in U}{argmin}\, E \left\{ \sum_{i=1}^{N} \alpha_i g(s_{n+i}, u_{n+i}, s_{n+i+1}) + \alpha_{N+1}\, \hat{J}(s_{n+N+1}, r_n\,) \right\} \tag{4}$$

where N is the number of stages to be considered explicitly.

For features ϕ in equation (3), we extracted the scheduled running time of remaining trains in the control area (ϕ_1) and their expected entrance delays (ϕ_2) to approximate value functions $\hat{J}(s, r)$. We considered three objective functions to evaluate the cost during a single stage g in equation (4): 1) Minimising total consecutive delays, 2) minimising total delays, and 3) minimising total running times inside the control area.

$$Objective\ 1 = \sum_n [d_n]_+ \tag{5}$$

$$Objective\ 2 = \sum_n L_n^e + \sum_n d_n \tag{6}$$

$$Objective\ 3 = \sum_n T_n + \sum_n [d_n]_+ \tag{7}$$

Where d_n is the accumulated consecutive delay inside the control area per stage, L_n^e is delay on entry to the control area, and T_n is the scheduled running time inside the control area.

To achieve convergence between the true value function J and the approximated value function $\hat{J}$, the vector of weights r is updated according to information as it becomes available from the operating environment. There are several different methods for updating r (Powell, 2011), among them Reinforcement Learning techniques have received considerable attention in the literature and have been adopted successfully to ADP approaches for adaptive traffic management (Cai et al., 2009). In this investigation, we use the least-squares temporal difference (LSTD) learning technique as presented by Bradtke and Barto (1996).

RESULTS AND DISCUSSION

To evaluate the present ADP approach, infrastructure and operational data for a section of Great Britain's East Coast Main Line (around 20 km between Stevenage Station to Hatfield Station) was used to simulate disturbances and compare First-Come-First-Served (FCFS) sequences against those calculated using ADP. We used 2018 timetable for the area for the southbound morning peak hours between 7:00 and 10:45 am which includes 30 trains. In total, we randomly generated 60,000 individual train delays (2000 morning peaks) resulting in a mean entry delay of 9 minutes 56 seconds.

We found that the ADP approach achieves superior performance compared to FCFS using all objective functions, though performance varies among the objective functions. Table 1 presents comparison of performance based on delays on the exit of our test case network for each of the objective functions, and hence provides a combined view of performance. It is apparent that Objective 2 does not perform well: the performance under ADP with this objective is worse than with the others, even in terms of its own measure. Noticeably, Objective 1 and Objective 3 both perform better in terms of Objective 2's primary objective i.e. minimising total delays.

	Mean total consecutive delay (s) - Obj.1	Mean total delay (s) - Obj.2	Mean running time (s) - Obj.3
FCFS	155.21	16409.96	285.11
ADP-Obj.1	62.48	16291.24	281.06
ADP-Obj.2	96.38	16304.51	281.51
ADP-Obj.3	73.32	16278.07	280.61

Table 1 - Comparison of objective functions against all performance measures

Fig. 1 presents total consecutive delays for all simulated morning peaks (all delay scenarios). It is evident from this graph that Objective 1 performs the best in all percentiles shown on the graph. Objective 2 has more variance after the 95th percentile compared to others. The maximum for Objective 2 is higher than the maximum for FCFS. Fig. 2 presents total delays for all delay scenarios. Improvements in total delays are small compared to FCFS. The reason for this is the limited opportunity for bringing trains back to their planned schedules. All objective functions improve total delays with Objective 3 performing the best. Fig. 3 presents train running times for all trains. All objectives perform favourably compared to FCFS in this measure up to the 95th percentile. Higher maximums in Fig. 3 for our objectives indicate trains that were held at the junction to allow other trains through. This action is beneficial as is evident from the improved overall performance.

Considering all findings (Fig. 1 - Fig. 3), it is evident that performance of the ADP approach is sensitive to the choice of objective function and unsuitable choice can result in poor performance. We have shown that, in the case of railway rescheduling, objective functions may be outperformed by others in terms of their primary objective. In any case, the ADP approach outperforms FCFS. The benefits of the ADP approach include improved management of perturbations to the operation of trains, resulting in improved resilience of operation and capability to run trains at reduced headways, and hence increased capacity.

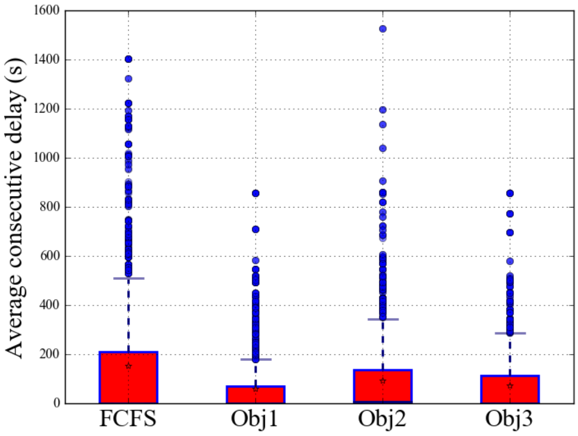

Fig. 1 – Average consecutive delay per scenario

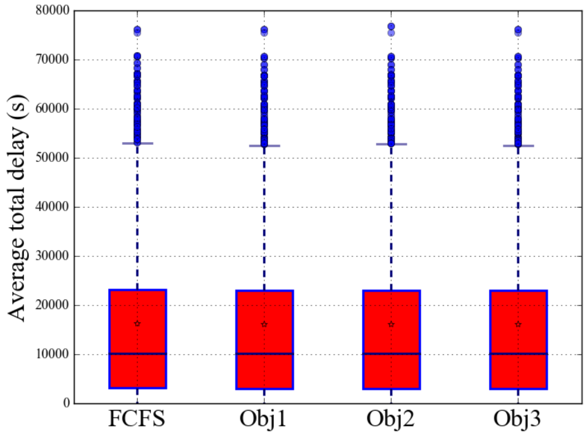

Fig. 2 – Average total delay per scenario

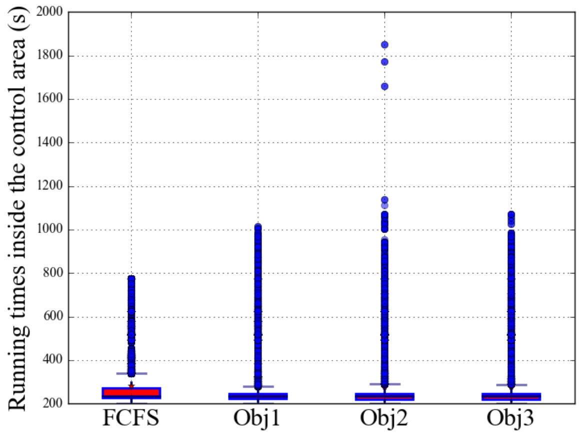

Fig. 3 – Train running times inside the control area

References

Bradtke, S.J., Barto, A.G., 1996. Linear Least-Squares algorithms for temporal difference learning. Machine Learning, 22, 33–57. https://doi.org/10.1007/BF00114723

Cacchiani, V., Huisman, D., Kidd, M., Kroon, L., Toth, P., Veelenturf, L., Wagenaar, J., 2014. An overview of recovery models and algorithms for real-time railway rescheduling. Transportation. Research (Part B, Methodological), 63B, 15–37. https://doi.org/10.1016/j.trb.2014.01.009

Cai, C., Wong, C.K., Heydecker, B.G., 2009. Adaptive traffic signal control using approximate dynamic programming. Transportation Research (Part C, Emerging Technology) 17C, 456–74. https://doi.org/10.1016/j.trc.2009.04.005

Powell, W.B., 2011. Approximate Dynamic Programming: Solving the Curses of Dimensionality. John Wiley & Sons, Inc., Hoboken, New Jersey.