Modeling Single-Line Train Operations

A. HIGGINS, L. FERREIRA, AND E. KOZAN

Scheduling of trains on a single line involves the use of train priorities for the resolution of conflicts. First, a mathematical programming model is described. The model schedules trains over a single line of track when the priority of each train in a conflict depends on an estimate of the remaining crossing and overtaking delay. This priority is used in a branch-and-bound procedure to allow the determination of optimal solutions quickly. This is demonstrated with the use of an example. Rail operations over a single-line track require the existence of a set of sidings at which trains can cross or overtake each other. Investment decisions on upgrading the numbers and locations of these sidings can have a significant impact on both customer service and rail profitability. Sidings located at insufficient positions may lead to high operating costs and congestion. Second, a model to determine the optimal position of a set of sidings on a single-track rail corridor is described. The sidings are positioned to minimize the total delay and train operating costs of a given cyclic train schedule. The key feature of the model is the allowance of nonconstant train velocities and nonuniform departure times.

This paper deals with two problems of single-line train scheduling, namely, the on-line scheduling of trains over a single-line track with multiple sidings and the optimum location of the sidings with respect to a given schedule. Part I deals with the optimum dispatching of trains on a single line of track. Trains can be dispatched from either end or from intermediate points on the track. When two trains approach each other on a single line one of them must take the siding for the safe operation of the system. Determining which train takes the siding is done by taking into account such factors as train priority, distance, lateness, and train operating costs. It is common practice for train operators to set a fixed timetable through which conflicts are resolved. A train dispatcher in a control center will act in the event of unforseen events. Because these events can cause delays to trains the dispatcher needs to continually alter the given timetable and resolve new conflicts. This is usually performed manually under strict time constraints so that the number of alternatives that can be assessed is very limited.

The operator's experience and knowledge of local conditions will continue to be used. Train dispatching decisions, which to a certain extent involve human as well as technical factors, will require human intervention to resolve problems. However, with the availability of such an optimization model, the operator is able to quickly update a schedule as unplanned events occur. The new optimal schedule offered by the model may not be fully implementable for practical reasons. However, the gap between the optimum and the practically feasible schedule can readily be assessed. The penalty for not being able to implement the optimum schedule in terms of operating cost and travel time reliability can be evaluated against the practical factors that prevent implementation of the optimum schedule.

A. Higgins and E. Kozan, School of Mathematics, Queensland University of Technology, P.O. Box 2434, QLD 4001, Australia. L. Ferreira, School of Civil Engineering, Queensland University of Technology, P.O. Box 2434, QLD 4001, Australia.

With the on-line train scheduling problem the determination of the priority of a train at a particular point on the journey involves the consideration of the initial priority, current lateness of the train, and a lower-bound estimate of possible further conflict delay. Exploiting such a lower bound in a model will act as a look-ahead function and will allow optimum schedules to be located quickly.

A second major use of the model relates to the planning of railroad operations. Such planning can be conveniently divided into two components, namely, short- to medium-term train planning and railroad infrastructure planning associated with train operations. The model can be used to evaluate the implications of changes to a timetable in terms of changed train departures, additional trains, and changes in train speeds. The optimum scheduling algorithm can be used as a simulator of proposed changes. Finally, the model can be used for long-range planning of railroad operations. In Australia, two main infrastructure planning issues are under investigation, namely, the upgrading of main line track to allow higher speeds and heavier axle loads and the need to extend sidings to allow for longer trains. The scheduling optimization model can be used to evaluate both of these investment strategies. The impact on the schedule of extending some sidings and not others can be assessed by using the model to simulate the effects of the proposed changes on future schedules. The removal of sidings has a cost in terms of flexibility and feasibility of schedules.

Part II deals with the development of a model for estimating the optimum positions of sidings on a single line of track. With high capital costs a rail line must be designed as economically as possible, and at the same time it must have enough capacity to accommodate the forecast demand. Planning for a rail line involves determining the number of sidings required, the length of each siding, its position, and the vertical and horizontal alignments for the line.

When determining the positions of sidings several variables must be considered. The sidings must be placed in order to minimize train delays and total train operating costs. If too many sidings are planned for the initial capital costs will outweigh the long-term benefits and there will be wasted capacity.

PART I: OPTIMUM TRAIN SCHEDULES

Past Research

Research involving the scheduling of trains on a single line of track is extensive, and the following highlights the major developments.

Kraft (*1*) developed a dispatching rule giving the optimal time advantage for a particular train based on train priority, track running times, and the delay penalties of each train. A similar method discussed by Sauder and Westerman (*2*) was implemented as a Decision Support System in a railway division of the United States. Those models, which assume fixed train speeds, produce train plans that minimize the weighted total travel times.

Kraay et al. (3) were the first to look at the idea of determining the cross–overtake plan and velocity profile to pace trains to conserve fuel while keeping the lateness of the trains to a minimum. Formulated constraints similar to those of Kraay et al. (3) were used in an interactive Decision Support System (SCAN) by Jovanovic and Harker (4) to develop reliable train schedules by using current schedules. Mills et al. (5) formulated a discrete network-type model by discretizing the departure and arrival time variables of this formulation.

Model Formulation

Assumptions and Inputs

The following assumptions are made with regard to the model in this section:

- The track is divided into segments that are separated by sidings.
- Crossing and overtaking can occur at any siding or double-line track segments.
- Trains can follow each other on a track segment with a minimum headway.
- For double-line track sections it is assumed that one lane will be allocated for inbound trains and one lane will be allocated for outbound trains. Usually, signal points will be set up this way.
- Scheduled stops are permitted at any intermediate siding for any train.

The model will require various pieces of information as inputs to the model. The specific information is as follows:

- An unresolved train plan to make available the number of overtake and cross interferences for each train.
- The initial priorities of each train. These are determined by several factors such as the type of train, customer contract agreements, and train load.
- The upper and lower velocity limits for each train (which are dependent on the physical characteristics of the track segment and the train).
- Segment lengths and the identification of single- and double-line track segments.
- The times of any scheduled train stops. These stops may include those for loading and unloading, refueling, and crew changes.

Definition of Variables

The set of trains is denoted by $I = \{1, 2, \ldots, m, m + 1, \ldots n\}$ for which inbound trains are from 1 to m and outbound trains are from $m + 1$ to n. The variables used in the model are listed and described in this section.

Let $P = \{P_1, P_2\}$

where P_1 is equal to a set of single-line tracks and P_2 is equal to a set of double-line tracks. The integer decision variables for determining which train traverses a section first (and which also determines the position of conflict resolution) are given by the following:

$$A_{ijp} = \begin{cases} 1 \text{ if inbound train } i \le m \text{ traverses track segment } p \in P_1 \\ \text{before inbound train } j \le m \\ 0 \text{ otherwise} \end{cases}$$

$$B_{ijp} = \begin{cases} 1 \text{ if inbound train } i \le m \text{ traverses track segment } p \in P_1 \\ \text{before outbound train } j > m \\ 0 \text{ otherwise} \end{cases}$$

$$C_{ijp} = \begin{cases} 1 \text{ if outbound train } i > m \text{ traverses track segment } p \in P_1 \\ \text{before outbound train } j > m \\ 0 \text{ otherwise} \end{cases}$$

The arrival and departure time decision variables are as follows:

X^i_{aq} = arrival time of train $i \in I$ at station $q \in Q$,
X^i_{dq} = departure time of train $i \in I$ from station $q \in Q$,
X^i_{Oi} = departure time of train $i \in I$ from its origin station, and
X^i_{Di} = arrival time of train $i \in I$ at its destination station.

The input parameters are defined as follows:

h_p = minimum headway between two trains on segment $p \in P_1$,
d_p = length of segment $p \in P$,
Y^i_{Oi} = planned departure time of train $i \in I$ from origin station,
Y^i_{Di} = planned arrival time of train $i \in I$ at destination station,
$\underline{v}^i_p$ = minimum allowable velocity of train $i \in I$ on segment $p \in P$,
$\overline{v}^i_p$ = maximum achievable average velocity of train $i \in I$ on segment $p \in P$,
W_i = initial priority of train $i \in I$ (highest for passenger trains), and
S^i_q = scheduled stop time for train $i \in I$ at station $q \in Q$.

An illustration of the ordering of a single track used for the model in this paper is given in Figure 1, in which the set of stations are represented by $Q = \{1, 2, \ldots, NS\}$ and track is represented by $(p - 2) \in P_2$.

Model Derivation

The objective function used in the model takes the following form:

$$\text{Min} \sum_i W_i \cdot (\text{delay of train } i \in I \text{ at destination}) + \text{train operating costs} \tag{1}$$

For the purposes of the solution procedure [namely, the branch-and-bound) (BB) procedure] the delay of train $i \in I$ comprises two parts. These are the current delay of train $i \in I$ at any point in time and a lower-bound estimate of remaining overtake and crossing delay from this point (6). The model is subject to various constraints to ensure safe operation, enforce speed restrictions, and permit stops. The following and overtake constraints for outbound trains $i, j \in I$, are as follows:

$$\left.\begin{aligned} M \cdot C_{ijp} + X^i_{aq+1} &\ge X^j_{aq+1} + h_p \\ M \cdot C_{ijp} + X^i_{dq} &\ge X^j_{dq} + h_p \end{aligned}\right\} \forall p \in P_1 \text{ and } i, j > m \tag{2}$$

$$\left.\begin{aligned} M \cdot (1 - C_{ijp}) + X^j_{aq+1} &\ge X^i_{aq+1} + h_p \\ M \cdot (1 - C_{ijp}) + X^j_{dq} &\ge X^i_{dq} + h_p \end{aligned}\right\} \forall p \in P_1 \text{ and } i, j > m \tag{3}$$

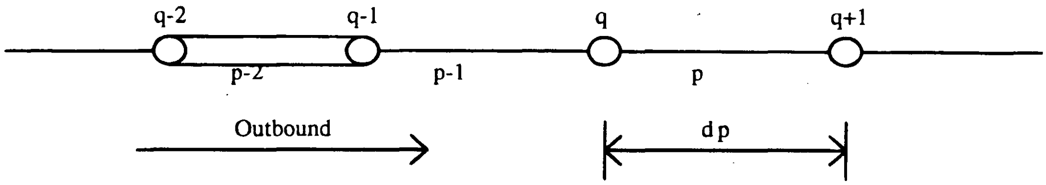

FIGURE 1 Sample of a network showing the single and double track segments.

The following and overtake constraints for inbound trains $i, j \in I$, are as follows:

$$\left.\begin{array}{l} M\cdot A_{ijp} + X^i_{aq} \geq X^j_{aq} + h_p \\ M\cdot A_{ijp} + X^i_{dq+1} \geq X^j_{dq+1} + h_p \end{array}\right\} \forall\, p \in P_1 \text{ and } i, j \leq m \tag{4}$$

$$\left.\begin{array}{l} M\cdot (1 - A_{ijp}) + X^j_{dq} \geq X^i_{aq} + h_p \\ M\cdot (1 - A_{ijp}) + X^j_{dq+1} \geq X^i_{dq+1} + h_p \end{array}\right\} \forall\, p \in P_1 \text{ and } i, j \leq m \tag{5}$$

Equation 2 implies that if train $j \in I$ goes first then train $i \in I$ must depart station $q \in Q$ after train $j \in I$ plus the minimum headway and arrive at station $(q + 1) \in Q$ after train $j \in I$ plus the headway. Equation 3 is similar except train $i \in I$ goes first. Equations 4 and 5 are the same as Equations 2 and 3, respectively, but for inbound trains. The constraints for the case when two trains approach each other are

$$\left.\begin{array}{l} h_p + X^j_{aq+1} \leq X^i_{dq+1} + M \cdot B_{ijp} \\ h_p + X^i_{aq} \leq X^j_{dq} + M \cdot (1 - B_{ijp}) \end{array}\right\} \forall p \in P_1\ (i \leq m, j > m) \tag{6}$$

Equation 6 implies that if outbound train $j \in I$ goes first then inbound train $i \in I$ must depart station $q \in Q$ after train $j \in I$ arrives plus a safety headway. The constant M is chosen to be large enough so that both equations in each crossing and overtake constraint are satisfied. Given the upper and lower velocities for each train on each segment, the upper and lower limits for traversal time of train $i \in I$ on segment $p \in P_l$ are given by

$$\frac{d_p}{\bar{v}^i_p} \leq X^i_{aq+1} - X^i_{dq} \leq \frac{d_p}{v^i_p}, \quad i > m, p \in P$$

$$\frac{d_p}{\bar{v}^i_p} \leq X^i_{aq} - X^i_{dq+1} \leq \frac{d_p}{v^i_p}, \quad i \leq m, p \in P \tag{7}$$

To stop trains from departing before their scheduled times and trains departing intermediate stations before they arrive, the following constraints are included:

$$\left.\begin{array}{l} X^i_{Oi} \geq Y^i_{Oi} \\ X^i_{aq} + S^i_q \leq X^i_{dq} \end{array}\right\} \forall\, i \in I, q \in Q \tag{8}$$

The objective is to minimize Equation 1 subject to constraints given by Equations 2 through 8.

Solution Procedure

The solution procedure described in this section is based on the BB procedure and uses the depth first search for the resolution of conflicts. Each node in the BB tree represents a partially resolved schedule that is calculated by solving a nonliner program i.e., solve objective function (Equation 1 subject to Equations 7 and 8) and the appropriate overtake or crossing constraints from Equations 2 through 6). The lower bound to the conflict delay costs of the remaining conflicts is calculated after the partial schedule is determined and is added to the cost of the partially resolved schedule. The BB technique used is described in full detail by Higgins et al. (*6*).

Model Testing

The exact algorithms of Sections 3 and 4 are implemented in FORTRAN on a 80486 personal computer. To solve the nonlinear programs GAMS/MINOS 5.2 (*7*) is accessed from the FORTRAN program. The model was tested on train schedules varying from 9 trains to 49 trains and was compared with a BB procedure with a lower bound calculated by relaxing the remaining conflict constraints. The method in this paper was able to find the optimal solution with up to 30 times fewer evaluations of the nonlinear program for most problems. For most problems the first upper bound was the optimal solution. The method of using a lower bound calculated by relaxing the remaining conflicts required from a few hundred to several thousand evaluations. This is very important for a real-life scheduling system because a solution would be required within a set time limit. The problem represented in Figure 2 contains 30 trains (53 conflicts) and was solved with 13 times fewer evaluations when the improved lower-bound estimate was used.

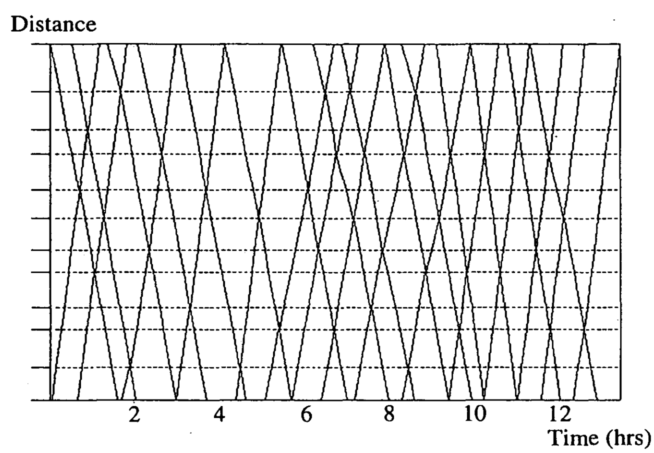

FIGURE 2 Optimal solution of 30 train problem.

PART II: OPTIMUM LOCATION OF SIDINGS

Past Research

Since most of the work done to determine the best positions of sidings uses simulation, optimal strategies are not usually found. The limited literature that does consider the optimality of siding positions only assumes simple train movements.

Petersen and Taylor (*8*) investigated an analytical model to determine the required numbers and lengths of sidings for a schedule of passenger trains. The determination of the length of the siding was to obtain the maximum benefit of the acceleration and deceleration characteristics of the trains.

An approach was taken by Kraft (*9*) to derive an analytical equation for determining the best position between two yards to put a siding. To construct the model free running time between sidings, average running speed (including delays), and the number of trains per unit of time was considered. The model cannot consider multiple sidings simultaneously. The equation may, however, be useful as an initial estimate. Mills et al. (*10*) use simulation, analytic, and heuristic techniques to investigate the line capacity of a mine-to-port track system. The analytic model, which determines the optimal number of equally spaced sidings, is based on the expected crossing delay to a train.

None of the studies described earlier considers solving for the optimal positions of sidings without the assumption of constant velocities and equally spaced sidings. The remainder of this paper considers the model formulation and solution to this problem.

Model Development

In this section the analytical model is formulated. The main feature of the model is the treatment of the track segment lengths (or siding positions) as a variable. Some sidings will be located at fixed positions and are not considered as variables in the model. This occurs when the siding is already existing or if it is to serve another purpose besides resolving conflicts. Scheduled stops will be permitted only at fixed sidings (stations).

Assumptions

The following assumptions are made specifically in conjunction with the siding location model:

- Double-line track sections are allowed and can be solved for optimum length, but their positions in a string of track segments cannot be moved.
- Generally, only one train can occupy a siding at one time (unless specified) except for the origin and destination stations, which are assumed to have infinite capacity.
- Scheduled stops are permitted only on fixed sidings.
- The train schedule is a cyclic schedule that is repeated on a daily or weekly basis.

The following information (in addition to the information required in the first part of this paper) is required by the user:

- The upper velocities of each train at 1-km intervals of the track. These are used to approximate the upper velocities of trains on track segments because they are considered as a variable during the calculation procedure.
- Any cost parameters such as cost of lateness per time and train operating costs.
- Initial positions of the sidings. A good initial solution will ensure fast convergence.

Definition of Variables

The variables used in the siding location problem are the same as those in the first part of this paper except for the following differences. The sidings are represented by the set $Q = \{1, 2, \ldots, NS\}$, where NS is the total number of sidings in the track system. Let Q_1 represent the set of fixed stations (sidings), and let Q_2 represent the set of variable sidings.

If the train schedule considered consists of daily and weekly trains then the cycle will be 1 week (i.e., the schedule considered is one cycle). The scheduled stop time is defined as

S_q^i = scheduled stop time for train $i \in I$ at station $q \in Q_1$

The upper velocity of a train on a discrete interval of track (used when calculating the upper velocities on a track segment) is given by $\overline{vel}_g^i$, which is equal to the upper velocity of train $i \in I$ at distance interval g on the track. Assume that the minimum headway is given by h. It does not, however, have to be constant for all trains and track segments since the minimum headway may be train dependent or may be determined by signal points.

Formulation and Constraints

The objective function will generally take the form of minimizing train delay costs and train operating costs. A dynamically prioritized delay criterion that allows the priority of each train to change from origin to destination is discussed in the model given in Part I. Objectives involving minimizing the destination lateness of trains are found in reports by Kraay et al. (*3*) and Mills et al. (*5*), whereas Petersen et al. (*11*) minimize total traveling times. Although the model in this paper does not depend on the objective used, it is important for the objective function to be convex to avoid the location of local optima. The overtake, crossing, upper velocity, and scheduled stop constraints are the same as those given in Part I of this paper.

Since the track segments are of variable positions and lengths (during the solution procedure), the upper velocities must be approximated. To estimate the upper velocities (maximum achievable velocities) it will be assumed that the upper velocities on each 1-km interval of the track corridor are known (or calculated by using a train movement simulator). If 1-km intervals are too fine then larger intervals may be used. Since the problem will be solved iteratively, the upper velocity of a train on a track segment is calculated by taking the average upper velocity of the intervals that lie in the track segment of the current solution. The upper velocity of train $i \in I$ on track segment $p \in P$ is calculated by the following equations:

$$\overline{v}_p^i = \frac{\sum_{g=dl}^{dh} \overline{vel}_g^i}{(dh - dl) + 1} \tag{9}$$

where dl is equal to the integer part of $\left(\sum_{k=1}^{p-1} d_k\right) + 1$, and dh is equal to the integer part of $\left(\sum_{k=1}^{p} d_k\right)$

The expected arrival times at intermediate sidings are also dependent on the positions of sidings and are calculated by first determining the planned velocities on the track segments. The planned velocities are calculated as follows:

$$RA_i = \frac{Y_{Di}^i - Y_{Oi}^i}{\sum_g \frac{1}{\overline{vel}_g^i}}$$

where RA_i is the ratio of the fastest journey time to the expected journey time, and

$$PV_p^i = \frac{\overline{v}_p^i}{RA_i}$$

where PV_p^i is the planned velocity of train $i \in I$ on track segment $p \in P$. From the planned velocity the expected departures from each of the intermediate stations are calculated by using Equation 10. The expected arrival times will be the same as the expected departure times unless there are scheduled stops.

$$Y_{dq}^i = Y_{Oi}^i + \sum_{k=1}^{q-1} \frac{d_k}{PV_k^i}, \text{ train } i \in I \text{ is outbound}$$

$$Y_{dq}^i = Y_{Oi}^i + \sum_{k=q}^{TRP} \frac{d_k}{PV_k^i}, \text{ train } i \in I \text{ is inbound} \tag{10}$$

where TRP is the number of track segments on the rail corridor. The fastest times that the trains travel from origin to destination are assumed to not be affected by the siding positions, so the expected arrivals and departures at these sidings do not change. The last constraint is to ensure that the sum of the length of the track segments is equal to the length of the entire track corridor, that is

$$\sum_{k=1}^{i-1} d_k = TLEN_i, \ i \in Q_1 \tag{11}$$

where $TLEN_i$ is the length of the track system from the origin to the fixed siding $i \in Q_1$.

Solving the Model

In this section a decomposition procedure is presented to obtain a solution to the formulation presented earlier. Solving the problem as formulated can be difficult because of the requirement that three sets of variables be solved (track segment lengths, arrival and departure times, and binary conflict resolution variables). The binary conflict resolution variables are solved by a BB type of procedure (or a heuristic procedure) and require the sidings to be at fixed positions. The problem must be decomposed so that solutions can be obtained for the three sets of variables.

The decomposition procedure proposed here is different from the Generalised Benders Decomposition (GBD) by Geofferon (*12*). The GBD partitions the model via the set of continuous variables and the set of integer variables. A more efficient way would be to partition the problem so that the structure of the problem could be exploited. This will allow a more efficient means of solving the subproblems to be used. The model here will be decomposed into two submodels: one that is solved for track segment lengths and arrival and departure times and the other that is solved for the optimal train schedule given the track segment lengths. The process will iterate between the two subproblems until there is no more improvement. This type of decomposition procedure is popular when solving complicated routing and scheduling problems. When one set of variables is fixed the problem can sometimes be reduced to a well-known form that can be easily solved by common procedures or heuristics. Two good examples are found in papers by Koskosidis et al. (*13*), which looks at the soft time window constraints for the vehicle routing problem, and Sklar et al. (*14*), which considers the aircraft scheduling problem.

The complete model for this paper can be stated by Equation 12:

$$\text{Min } Z = f(d_k \, \forall k, X_{dq}^i \, \forall i, q, X_{aq}^i \, \forall i, q, A_{ijp} B_{ijp} C_{ijp} \, \forall i, j, p) \tag{12}$$

which is subject to the constraints in Equations 2 through 11 and where $f(\cdot)$ represents the nonlinear (or linear) objective function of the variables defined in the first part of this paper. The model is decomposed to form Models Z_1 and Z_2. The Model Z_1, which is represented by Equation 13, is solved to obtain the optimum track segment lengths subject to fixed conflict resolution variables A_{ijp}, B_{ijp}, and C_{ijp} (i.e., fixed schedule). Model Z_2 is solved to obtain the optimum schedule subject to fixed track segment lengths (i.e., normal train scheduling problem). Each model is solved by using the output from the other model as initial values.

$$\text{Min } Z_1 = f(d_k \, \forall k, X_{dq}^i \, \forall i, q, X_{aq}^i \, \forall i, q) \tag{13}$$

which is subject to the constraints in Equations 7 through 11, and

$$\text{Min } Z_2 = f(X_{dq}^i \, \forall i, q, X_{aq}^i \, \forall i, q, A_{ijp} B_{ijp} C_{ijp} \, \forall i, j, p) \tag{14}$$

which is subject to the constraints in Equations 2 through 9 and 11.

The upper velocities of Model Z_1 will be those of the latest solution of model Z_2. This is reasonable, since to have the upper velocities as a function of track segment lengths d_k (which is variable in Model Z_1) would require nonlinear constraints. This may cause the solution to Model Z_1 to be slightly inaccurate for the first couple of iterations if there is a large change in siding positions. The results generated in the next section have indicated little effect on the convergence.

The following variables will be defined for the decomposition algorithm to resemble the current stage of solution.

d_k^t = length of track segment $k \in P$ after the tth iteration using model Z_1,

$X_{dq}^{i,t,1}$ = departure time of train $i \in I$ from station $q \in Q$ after the tth iteration using model Z_1,

$X_{aq}^{i,t,l}$ = arrival time of train $i \in I$ at station $q \in Q$ after the tth iteration using model Z_1,

$\left.\begin{matrix} B_{ijp}^t \\ A_{ijp}^t \\ C_{ijp}^t \end{matrix}\right\}$ conflict resolution decision variables after the tth iteration of model Z_2,

$X_{dq}^{i,t,2}$ = departure time of train $i \in I$ from station $q \in Q$ after the tth iteration using model Z_2,
$X_{aq}^{i,t,2}$ = arrival time of train $i \in I$ at station $q \in Q$ after the tth iteration using model Z_2, and
$\overline{v}_p^{i,t}$ = upper velocity of train $i \in I$ on segment $p \in P$ for the tth iteration.

The expected departure times are calculated by using Equation 10, and these constraints will be linear since the planned velocities are constant. This is because the upper velocities from Model Z_2 are used in the current iteration of Model Z_1. The initial track segment lengths d_k^0 can be estimated by simulation techniques or by a simple inspection to see where the conflicts occur. Another method is to just assume equal track segment lengths for the initial estimates. If the purpose is to upgrade an existing track corridor, then the current positions of some existing sidings may be used for the initial estimate.

The optimum siding positions are calculated by the following decomposition procedure:

1. Given initial values d_k^0 $\forall k$ solve the Model Z_2 to obtain $X_{dq}^{i,1,2}$, $X_{aq}^{i,1,2}$, B_{ijp}^1, A_{ijp}^1, and C_{ijp}^1 $\forall i, j, p$. Solving Model Z_2 is exactly the same as solving the normal train scheduling problem (*3,6*). Let t equal 1.
2. Given B_{ijp}^t, A_{ijp}^t, and C_{ijp}^t solve the nonlinear program Z_1 for d_k^t, $X_{dq}^{i,t,1}$ and $X_{aq}^{i,t,1}$. This part is not a computational burden, but the objective function is more complex because d_k^t is variable. The form of this model makes it suitable for solution by using a simplicial decomposition procedure (*15*).
3. Let t equal $t + 1$. Solve the problem Z_2 given d_k^{t-1} for $X_{dq}^{i,t,2}$, $X_{aq}^{i,t,2}$, B_{ijp}^t, A_{ijp}^t, and C_{ijp}^t by using $X_{dq}^{i,t,1}$, $X_{aq}^{i,t,1}$, B_{ijp}^{t-1}, A_{ijp}^{t-1}, and C_{ijp}^{t-1} as initial values. The procedure terminates when the conflict resolution strategy does not change from iteration $t - 1$ to t. It is a major computa-tional burden to solve for the integer variables by a BB procedure. It is required for the initial solution of Step 1, but if only a couple of conflict resolutions change as the positions of the sidings converge, then a much more efficient method of updating the conflict resolution strategy is necessary. A heuristic for this is described in a paper by Higgins et al. (*16*). Go to Step 2.

Model Testing

The examples considered here contain seven trains and six sidings, four of which are movable. The examples were chosen to illustrate the time savings of having sidings at their optimal positions compared with their current positions. The objective function chosen for examples is minimum tardiness plus fuel cost. The fuel consumption function is the same as that used by Mills et al. (*5*). For the two examples presented here the restriction of one train per siding is relaxed. The lateness at destinations for the initial and optimal solution (both examples) are given in Table 1(*a*), with the track segment lengths given in Table 1(*b*). Figure 3(*a*) represents the initial resolved train graph for the first example. By inspection of this train graph it appears that the sidings are at quite reasonable positions with respect to the conflicts. The only real indication is that Siding 2 could be closer to the inbound origin station. When the sidings are at optimum positions, as shown in Figure 3(*b*), considerable time savings are obtained for the trains and they are kept closer to schedule throughout the journey. The second and fourth columns of Table 1(*a*) indicate the time saved for trains when sidings are at optimal positions. More than an hour of delay has been cut for all trains.

The first example required only two iterations (terminated at t = 2) of the decomposition procedure to achieve the optimal solution. Only one conflict required changing from the first iteration to the second. The original track segment lengths d_k^0 indicate that a good initial solution will ensure fast convergence. The initial positions of the track segments in the second example are a lot poorer than those

TABLE 1 Comparison of (*a*) Lateness at Destinations and (*b*) Original and Optimal Track Segment Lengths.

	Destination lateness for current solution (hrs)		Destination lateness for optimal solution (hrs)	
Train	Example 1	Example 2	Example 1	Example 2
1	0.01	0.00	0.00	0.00
2	1.62	0.46	1.01	0.46
3	1.86	0.38	1.58	0.22
4	0.00	0.21	0.00	0.00
5	0.38	0.42	0.33	0.55
6	1.15	1.07	1.08	0.21
7	1.09	0.40	1.07	0.40
All Trains	6.11	2.94	5.07	1.84

(*a*)

Track segment k	Original length (d_k^0 km)	Optimal length (km)	
		Example 1	Example 2
1	20.24	8.18	23.23
2	28.86	35.40	20.75
3	28.47	35.40	22.25
4	43.14	43.52	24.50
5	25.26	23.46	55.23

(*b*)

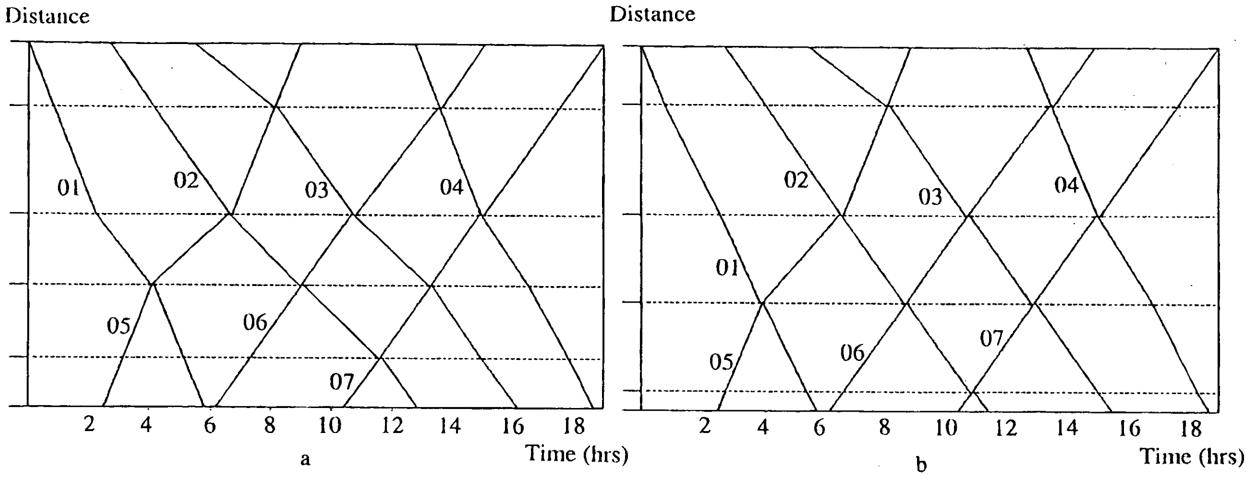

FIGURE 3 Optimal schedule given current siding positions (*a*) and given optimal siding positions (*b*).

in the first example. This example was set up so that most of the train interactions are toward the midpoint of the journey, where there are fewer sidings. The outbound trains suffer heavy delays because of this, and the optimal solution relocates the sidings toward the middle of the train graph. Referring to the third and fifth columns of Table 1(*a*), there has been a reduction in delay for many trains, with the average delay being significantly reduced.

MODEL LIMITATIONS

The models in both parts of this paper have some limitations as far as real-life applications are concerned. Although the emphasis of Part I was to allow optimal solutions to real-life problems to be found, it does not allow random delay events. Instead, the risk delay of a given resolved schedule can be calculated separately by a model described by Higgins et al. (*17*). The model takes into account the risk delay due to terminal and stoppage delays, train-related delays, and track-related delays.

Some trains will have different characteristics such as the number of wagons on a given day and the number of locomotives used. The upper achievable velocity is easily adjusted to cater to such train differences. At this stage the models have been tested by using a real-life problem that consists of a single-line track of 120 km with 13 sidings and a daily density of 30 trains. Most types of objective functions that have been proposed in the past can be accommodated by using the models. However, the inclusion of other variables such as delay risk would not be possible.

CONCLUSIONS

This paper has presented an on-line model for the scheduling of trains on a single-line track and a planning tool for determining the optimal positioning of sidings. The on-line model allows the priority of a train to change from origin to station, resulting in a more reliable system. This is a more realistic interpretation of how the train dispatcher would consider the rail network. Conflicts are resolved on the basis of their current priorities, which are dependent on the future delays for each train.

The on-line model will be useful to train dispatchers for generating more reliable train schedules. Optimum schedules will be generated quickly, and trains will be kept on schedule with respect to future delays. The results have demonstrated significant computation time improvements, especially for larger problems that involve tight schedules.

For the siding location problem a decomposition procedure was used iteratively to solve for the best siding positions and corresponding resolved schedule. Results of the model have shown much improvement in delays to trains when the sidings are at optimal positions. If using this model to determine the positions of sidings only reduces the overall delay by a small percentage (while keeping the train costs uniform) then the long-term benefits may be large.

Positioning of sidings is one aspect of trying to optimize freight rail transport. A larger concern, however, is the upgrading of existing track. It is important to know the effects of lateness and the reliability of schedules when upgrading the existing track corridor. Besides the delay that occurs when a train waits at a siding for another train to pass, other delays must be considered. These delays are categorized as risk delays and are caused by maintenance, any train failure, or environmental problems. A prime interest for upgrading existing track is the knowledge of this risk delay while the track is in its current state, and an estimate of this delay if certain upgrading was carried out. The authors are continuing research on the development of a model to estimate the risk delay to a system and to identify which sections of track contribute the most risk.

REFERENCES

1. Kraft, E. R. A Branch and Bound Procedure for Optimal Train Dispatching. *Journal of the Transportation Research Forum,* Vol. 28, 1987, pp. 263–276.
2. Sauder, R. L., and W. M. Westerman. Computer Aided Train Dispatching: Decision Support Through Optimisation. *Interfaces,* Vol. 13, 1983, pp. 24–27.
3. Kraay, D., P. Harker, and B. Chen. Optimal Pacing of Trains in Freight Railroads. *Operations Research,* Vol. 39, 1991, pp. 82–99.

4. Jovanovic, D., and P. Harker. Tactical Scheduling of Rail Operations: The SCAN I System. *Transportation Science,* Vol. 25, 1991, pp. 46–64.
5. Mills, R. G., S. E. Perkins, and P. J. Pudney. Dynamic Rescheduling of Long Haul Trains for Improved Timekeeping and Energy. *Asia-Pacific Journal Operational Research,* Vol. 8, 1991, pp. 146–165.
6. Higgins, A., E. Kozan, and L. Ferreira. An Improved Exact Solution Procedure for the Online Scheduling of Trains. *Physical Infrastructure Research Report 36–94.* Queensland University of Technology, Brisbane, Australia, 1995.
7. Brooks, A., D. Kendrick, and A. Meeraus. *GAMS: A User's Guide.* Scientific Press, 1988.
8. Petersen, E. R., and A. J. Taylor. Design of a Single-Track Rail Line for High Speed Trains. *Transportation Research A,* Vol. 21, 1987, pp. 47–57.
9. Kraft, E. R. Analytical Models for Rail Line Capacity Analysis. *Transportation Research Forum Proceedings,* Vol. 23, 1983, pp. 461–471.
10. Mills, G., S. Perkins, and D. Sier. Capacity Planning for a Railway Corridor. *Proc., 12th National Conference of the Australian Society of Operations Research,* 1983, pp. 470–487.
11. Petersen, E. R., A. J. Taylor, and C. D. Martland. An Introduction to Computer Assisted Train Dispatching. *Journal of Advanced Transportation,* Vol. 20, 1986, pp. 63–72.
12. Geofferon, A. M. Generalised Benders Decomposition. *Journal of Optimisation Theory and Applications,* Vol. 10, 1972, pp. 237–263.
13. Koskosidis, Y. A., W. B. Powell, and M. M. Solomon. An Optimisation-Based Heuristic for Vehicle Routing and Scheduling with Soft Time Window Constraints. *Transportation Science,* Vol. 26, 1992, pp. 69–85.
14. Sklar, M. G., R. D. Armstrong, and S. Samn. Heuristics for Scheduling Aircraft and Crew During Airlift Operations. *Transportation Science,* Vol. 24, 1990, pp. 63–76.
15. Hohenbalken, B. V. Simplicial Decomposition in Non-Linear Programming Algorithms. *Mathematical Programming,* Vol. 13, 1974, pp. 49–68.
16. Higgins, A., E. Kozan, and L. Ferreira. Development of a Model to Optimise the Position of Crossing and Overtaking Loops on a Single Line Railway. *Physical Infrastructure Research Report 16–94.* Queensland University of Technology, Brisbane, Australia, 1994.
17. Higgins, A., L. Ferreira, and E. Kozan. Modelling Delay Risks Associated with Train Schedules. *Transportation Planning and Technology,* in press.

Publication of this paper sponsored by Committee on Electrification and Train Control Systems for Guided Ground Transportation Systems.