Railway Timetabling using Lagrangian Relaxation
U. Brännlund, P.O. Lindberg, A. Nõu, J.-E Nilsson

Henrik Fredriksson
Blekinge Institute of Technology

June 16, 2015
Scheduling trains is a complex and time consuming procedure.
Train operators places bids to gain access to rail tracks.
Based on the bids, timetables are constructed (manually) to fulfill the requests from the train operators as good as possible.
In the model, time is discretized (typically in minutes) and the train tracks are divided into blocks.
In the model, time is discretized (typically in minutes) and the train tracks are divided into blocks.

**Notations and variables**

- $x$ - *timetable*
- $x^r$ - *schedule for train* $r$
- $x^r_{it}$ - *binary variable*, 1 if train $r$ occupies block $i$ at time $t$, 0 otherwise.
- $T^r$ - *set of logically feasible schedules for train* $r$
- $v(x)$ - *total value of a given timetable* $x$
\( (P) \) Maximize \( v(x) = \sum_r v^r(x^r) \) \hspace{1cm} (1)

subject to \( \sum_r x^r_{it} \leq 1, \ \forall i, t \) \hspace{1cm} (2)

\( x^r \in T^r, \ \forall r \) \hspace{1cm} (3)

In the model, a single rail track is considered. The constraints in (2) are called the \textit{linking constraints}. 
The profit for train $r$, denoted by $v^r(x^r)$ is the profit of departure time minus unnecessary waiting time (the per minute cost is modelled as $v^r_{\text{dep}}/w^r_{\text{max}}$, where $w^r_{\text{max}}$ is the maximal tolerated waiting time for train $r$).

**Figure:** Model of profit for departure time
(P_\lambda) \quad \text{Maximize} \quad v(x) = \sum_r v^r(x^r) + \sum_{i,t} \lambda_{it} \left(1 - \sum_r x^r_{it}\right) \\
\text{subject to} \quad x^r \in T^r, \quad \forall r 

The interpretation is that if the trains \( r \), occupies block \( i \) at time \( t \), then they are charged the cost \( \lambda_{it} \).
The relaxed problem separates into independent subproblems,

\[(P^r_\lambda) \quad \text{Maximize} \quad \nu(x) = \nu^r(x^r) - \sum_{i,t} \lambda_{it} x^r_{it}\]

subject to \(x^r \in T^r\), \hspace{1cm} (5)
The relaxed problem separates into independent subproblems,

\((P^r_\lambda)\) \begin{align*}
\text{Maximize } & \quad v(x) = v^r(x^r) - \sum_{i,t} \lambda_{it}x^r_{it} \\
\text{subject to } & \quad x^r \in T^r, \quad \text{(5)}
\end{align*}

Let \(\hat{x}(\lambda)\) be an optimal solution to \((P_\lambda)\) with objective value

\[\phi(\lambda) = \sum_r v^r(\hat{x}^r(\lambda)) + \sum_{i,t} \lambda_{it} \left(1 - \sum_r \hat{x}^r_{it}(\lambda)\right).\]

This objective value is an upper bound to the optimal value of \((P)\), i.e. \(\phi(\lambda) \geq v(x)\) for \(\lambda \geq 0\).
Finding the best possible bound is the Lagrangian dual problem

$$(D) \quad \text{Minimize} \quad \phi(\lambda)$$

subject to $\lambda \geq 0.$ \hspace{1cm} (7)
Main idea

At iteration $k$, the relaxed problem ($P_{\lambda^k}$) is solved using multiplier $\lambda^k$. Then, using a heuristic approach find a feasible solution $\bar{x}_k$ to the primal problem ($P$), e.g. by adjusting $\hat{x}(\lambda^k)$. Then, update the multipliers to get closer to the solution of ($D$).

If block $i$ is occupied at time $t$ by more than one train, i.e.,

$$\sum_r \hat{x}_{it}^r(\lambda^k) > 1,$$

then the price is raised such that $\lambda_{it}^{k+1} > \lambda_{it}^k$, and thereby making it less attractive to use. On the other hand, if $\lambda_{it}^k > 0$ and no train is using block $i$ at time $t$ we lower the price such that $\lambda_{it}^{k+1} < \lambda_{it}^k$. 
Let
\[
\phi^k_{\text{best}} = \min_{j \in \{1, \ldots, k\}} \phi(\lambda_j) \quad \text{(upper bound)}
\]
and let
\[
v^k_{\text{best}} = \max_{j \in \{1, \ldots, k\}} \sum_r v^r(\bar{x}^r_j) \quad \text{(lower bound)}
\]

denote the so far best objective value of \((D)\) and \((P)\) respectively.
Let
\[ \phi_{\text{best}}^k = \min_{j \in \{1, \ldots, k\}} \phi(\lambda^j) \] (upper bound)
and let
\[ v_{\text{best}}^k = \max_{j \in \{1, \ldots, k\}} \sum_r v_r(\bar{x}_r^j) \] (lower bound)
denote the so far best objective value of \((D)\) and \((P)\) respectively.

If \( \Delta^k = \phi_{\text{best}}^k - v_{\text{best}}^k \) is sufficiently small, we terminate.
Let
\[ \phi^k_{\text{best}} = \min_{j \in \{1, ..., k\}} \phi(\lambda^j) \] (upper bound)
and let
\[ \nu^k_{\text{best}} = \max_{j \in \{1, ..., k\}} \sum_r \nu^r(\bar{x}_j^r) \] (lower bound)
denote the so far best objective value of \((D)\) and \((P)\) respectively.

If \(\Delta^k = \phi^k_{\text{best}} - \nu^k_{\text{best}}\) is sufficiently small, we terminate.

Duality gap or the heuristic may cause that \(\Delta^k\) do not go to zero. If true optimum is required, the approach could be incorporated in a branch-and-bound algorithm.
Algorithm 1: A dual iteration scheme

Step 1. Set $k = 1$ and initialize $\lambda^1$ (e.g. $\lambda^k_{it} = 0, \forall i, t$). Set a limit, $k_{max}$, of numbers of iterations.

Step 2. Solve $(P_{\lambda^k})$ to obtain $\phi(\lambda^k)$ and $\hat{x}(\lambda^k)$.

Step 3. Find feasible solution, $\bar{x}_k$ to $(P)$ with objective value $\sum_r v^r(\bar{x}^r_k)$.

Step 4. Stop if $\Delta^k$ is small enough, or $k = k_{max}$. Otherwise, update $\lambda^{k+1}$, set $k = k + 1$ and return to Step 1.
Solving the relaxed problem

For each train \( r \), the subproblems \( (P^r) \) can be viewed as a shortest path problem in a space-time network. A path in this network corresponds to a technically feasible train schedule.
Finding primal feasible solutions

Normally, it cannot be expected to get feasible solutions by just adjusting the prices as in Algorithm 1. Instead of adjusting $\hat{x}(\lambda^k)$, the following heuristic approach is used.

1. Set a priority for the train types:
   1. Intercity trains
   2. Passenger trains
   3. Freight trains

2. The trains are scheduled one-by-one according to the priority list by the shortest-path problem procedure.

3. The highest priority train will get its most desired schedule with cost $\lambda^k$.

4. The network is updated by removing the occupied arcs.

5. A feasible schedule is found by finding a profit maximizing path in the reduced network.
Finding primal feasible solutions

Normally, it cannot be expected to get feasible solutions by just adjusting the prices as in Algorithm 1. Instead of adjusting $\widehat{x}(\lambda^k)$, the following heuristic approach is used.

1. Set a priority for the train types:
   1. Intercity trains
   2. Passenger trains
   3. Freight trains

2. The trains are scheduled one-by-one according the priority list by the shortest-path problem procedure.

Henrik Fredriksson
Railway Timetabling using Lagrangian Relaxation
Finding primal feasible solutions

Normally, it cannot be expected to get feasible solutions by just adjusting the prices as in Algorithm 1. Instead of adjusting $\hat{x}(\lambda^k)$, the following heuristic approach is used.

1. Set a priority for the train types:
   1. Intercity trains
   2. Passenger trains
   3. Freight trains

2. The trains are scheduled one-by-one according the priority list by the shortest-path problem procedure.

3. The highest priority train will get its most desired schedule with cost $\lambda^k$. 
Finding primal feasible solutions

Normally, it cannot be expected to get feasible solutions by just adjusting the prices as in Algorithm 1. Instead of adjusting $\hat{x}(\lambda^k)$, the following heuristic approach is used.

1. Set a priority for the train types:
   - Intercity trains
   - Passenger trains
   - Freight trains

2. The trains are scheduled one-by-one according the priority list by the shortest-path problem procedure.

3. The highest priority train will get its most desired schedule with cost $\lambda^k$.

4. The network is updated by removing the occupied arcs.
Finding primal feasible solutions

Normally, it cannot be expected to get feasible solutions by just adjusting the prices as in Algorithm 1. Instead of adjusting $\hat{x}(\lambda^k)$, the following heuristic approach is used.

1. Set a priority for the train types:
   1. Intercity trains
   2. Passenger trains
   3. Freight trains

2. The trains are scheduled one-by-one according to the priority list by the shortest-path problem procedure.

3. The highest priority train will get its most desired schedule with cost $\lambda^k$.

4. The network is updated by removing the occupied arcs.

5. A feasible schedule is found by finding a profit maximizing path in the reduced network.
Four different methods have been used to update the multipliers. Let $g(\lambda) = 1 - \sum_r \hat{x}^r(\lambda)$ be a subgradient of $\phi$ at $\lambda$. Then two of the updating procedures are as follows:
Four different methods have been used to update the multipliers. Let 
\( g(\lambda) = 1 - \sum_r \hat{x}_r(\lambda) \) be a subgradient of \( \phi \) at \( \lambda \). Then two of the updating procedures are as follows:

1. **Relaxation step technique or Polyak II**

\[
\lambda^{k+1} = \max\{0, \lambda^k - \frac{\phi(\lambda^k) - \bar{\phi}(\lambda^k)}{\|g(\lambda^k)\|^2} g(\lambda^k)\}
\]

where \( \bar{\phi}(\lambda^k) \) is some target value (convex combination of the best dual and primal objective values so far found).
Four different methods have been used to update the multipliers. Let \( g(\lambda) = 1 - \sum_r \hat{x}^r(\lambda) \) be a subgradient of \( \phi \) at \( \lambda \). Then two of the updating procedures are as follows:

1. **Relaxation step technique or Polyak II**

   \[
   \lambda^{k+1} = \max\{0, \lambda^k - \frac{\phi(\lambda^k) - \overline{\phi}(\lambda^k)}{\|g(\lambda^k)\|^2} g(\lambda^k)\}
   \]

   where \( \overline{\phi}(\lambda^k) \) is some target value (convex combination of the best dual and primal objective values so far found).

2. Let \( \lambda^{k+1} \) be solution to the projection problem

   \[
   \min\{\|\lambda - \lambda^k\|^2 : \overline{\phi}(\lambda^k) \geq \phi(\lambda^k) + g(\lambda^k)^T (\lambda - \lambda^k), \lambda \geq 0\}
   \]
Four different methods have been used to update the multipliers. Let \( g(\lambda) = 1 - \sum_r \hat{x}^r(\lambda) \) be a subgradient of \( \phi \) at \( \lambda \). Then two of the updating procedures are as follows:

1. **Relaxation step technique or Polyak II**

   \[
   \lambda^{k+1} = \max\{0, \lambda^k - \frac{\phi(\lambda^k) - \bar{\phi}(\lambda^k)}{\|g(\lambda^k)\|^2}g(\lambda^k)\}
   \]

   where \( \bar{\phi}(\lambda^k) \) is some target value (convex combination of the best dual and primal objective values so far found).

2. Let \( \lambda^{k+1} \) be solution to the projection problem

   \[
   \min\{\|\lambda - \lambda^k\|^2 : \bar{\phi}(\lambda^k) \geq \phi(\lambda^k) + g(\lambda^k)^T(\lambda - \lambda^k), \lambda \geq 0\}
   \]

In all four methods, the target value is updated according to

\[
\bar{\phi}^k = (1 - \gamma^k)\phi^k_{\text{best}} + \gamma^k v^k_{\text{best}}
\]

where \( \gamma^k \) is initially one, and then halved if \( \phi^k_{\text{best}} \) has not improved has not improved for the last \( K_{\text{max}} \) iterations.
A train $r$ can be cancelled ($x_{it}^r = 0, \forall i, t$), possibly at the price of a penalty.

The authors suggests various extension where a network of single track lines are considered.

The profit function probably needs to be modeled more accurately.
Thank you for listening!