A Numerical Optimization Technique for Scheduling Non-uniform Metro Trains under Time-varying Passenger Demands

Hang-fei HUANG and Ke-ping LI
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, 100044, Beijing, China

Keywords: Train scheduling problem, Non-smooth nonlinear programming, Augmented lagrangian, Pattern search.

Abstract: Urban rail transit is costly and often overloaded. Thus, the motivation of this paper is to develop specific operating strategies to reduce cost and improve service level. This paper proposed a numerical optimization technique to optimize non-uniform metro train timetables with time-varying passenger demands, where the total monetary cost of time and energy is minimized. The problem was formulated with a constrained, non-smooth, nonlinear programming model, and transformed into an unconstrained one by the proposed Augmented Lagrangian (AL) algorithm. Then, the Pattern Search (PS) algorithm was designed to search for an optimal solution in each iteration step of AL. The case study shows that the proposed technique is effective to obtain non-uniform timetables that significantly outperform uniform ones reducing cost at most 34.92%, and provides decision-making support.

Introduction

With growing metropolitan population and increasing concerns on climate change, public transportation receives considerable attention in recent decades. Urban rail transit, which has large capacity and utilizes energy more efficiently than automobiles, becomes one of the busiest transportation systems in large cities. However, current metro systems usually set uniform timetables in certain periods [1,2], while passenger demands vary from time to time [3,4]. Such uniform timetables are less efficient and economic than the non-uniform ones. In this regard, the motivation of this paper is to design a numerical technique to obtain and optimize non-uniform timetables.

To address above research gaps, this paper formulates the scheduling problem with a non-smooth nonlinear programming model, which aims to minimize both user and supplier costs, including passenger wait time, train travel time, and energy consumption. Compared to the other optimization methods in the literature [7,8,9], the main contribution of this paper is to use numerical methods to transform the constrained model into an unconstrained one, so that the direct optimization methods can be applied. The effectiveness of this technique is verified then in the case study.

Model Formulation

The metro line layouts and the basic problem descriptions are the same as in [5]. Notation used throughout this paper is listed in Table 1. The decision variables in the model are the arrival time of train i at station x ($t_{ia}$) and the departure time of train i from station x ($t_{id}$).
Table 1. Notation in this paper.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>Station index</td>
<td>i</td>
<td>Train index</td>
</tr>
<tr>
<td>t</td>
<td>Time stamp</td>
<td>h_{min}</td>
<td>Minimum headway</td>
</tr>
<tr>
<td>M_i</td>
<td>Net weight of train</td>
<td>M_{p}</td>
<td>Average weight of a passenger</td>
</tr>
<tr>
<td>N^w(i,t)</td>
<td>Number of in-train passengers</td>
<td>N^w(x,t)</td>
<td>Number of waiting passengers</td>
</tr>
<tr>
<td>N^b(i,x)</td>
<td>Number of alighting passengers</td>
<td>N^b(i,x)</td>
<td>Number of boarding passengers</td>
</tr>
<tr>
<td>P^a(t_1,t_2)</td>
<td>Number of arrival passengers</td>
<td>C_i</td>
<td>Train capacity</td>
</tr>
</tbody>
</table>

Passenger Wait Time and Train Travel Time

Passenger wait time at station x can be calculated by integrating the number of cumulative amounts of passengers with respect to time t.

\[
t^w_x = \int_{t_i}^{t_2} N^w(x,t) \, dt = \int_{t_i}^{t_2} P^c(t_2,t_i) \, dt - \sum_{x=1}^{X} N^b(i,x), \forall x \in [1,2X].
\]  

(1)

where \(i_t\) is the largest index of arrived trains at station x before time t.

Total passenger wait time is the sum of wait times at different stations:

\[
t^w = \sum_{x=1}^{X} t^w_x.
\]  

(2)

Total train travel time is the sum of differences between each train’s departure time at the last station and arrival time at the first station.

\[
t_t = \sum_{i=1}^{I} (t^d_i,x,x+1) - t^d_i(x,1).
\]  

(3)

Energy Consumption

Linear piecewise curve is used to approximate the energy per unit mass of train i in segment x:

\[
e_{i,x} = \left[ \lambda_x \cdot (t^d_i,x+1-x^d) + \mu_x \right] \cdot I_x, \forall i \in [1,1], x, x+1 \in [1,2X].
\]  

(4)

where \(\lambda_x\) and \(\mu_x\) are two parameters, and \(I_x\) is the length of segment x. The values of \(\lambda_x\) and \(\mu_x\) can be estimated by the mean squared error (MSE) method. To obtain the actual energy, the train mass, the average weight of a passenger, and the number of in-train passengers are needed:

\[
E_{i,x} = N^w(i,t^d_i,x) \cdot M_i + M_p \cdot N^b(i,t^d_i,x), \forall i \in [1,1], x, x+1 \in [1,2X].
\]  

(5)

Finally, total energy consumed by all trains in a period is:

\[
E = \sum_{i=1}^{I} \sum_{x=1}^{2X-1} E_{i,x}.
\]  

(6)

Objective Model

Denote the unit costs for passenger wait time, train travel time and energy as \(c_w\) (yuan/hour), \(c_t\) (yuan/hour) and \(c_e\) (yuan/Kwh). The optimization model is formulated as follows.

Minimize:
\[ T_i = c_w \cdot t_u + c_t \cdot t_t + c_e \cdot E. \] \hfill (7)

Subject to:

\[ t^d_{x,i} - t^d_{x,i+1} \geq h_{\text{min}}, \forall i, i+1 \in [1, I], x \in [1, 2X]. \] \hfill (8)

\[ t^d_{x,i} - t^d_{x,i-1} = t_{\text{turn}}, \forall i \in [1, I]. \] \hfill (9)

\[ N^y(i, t) \leq C_i, \forall i \in [1, I], t \in [t_0, t_{\text{end}}]. \] \hfill (10)

\[ t^d_{x,i}, t^d_{x,i} \in \mathbb{N}, \forall i \in [1, I], x \in [1, 2X]. \] \hfill (11)

Eq. 8 imposes the minimum headway constraint. Eq. 9 requires a fixed turnaround time \( t_{\text{turn}} \). Eq. 10 is the train capacity constraint. Eq. 11 defines that the decision variables are positive integers. The model is a nonlinear programming one. Since the train capacity is limited, it is also non-smooth.

### Numerical Optimization Technique

In practice, travel times in segments are the same, and thus the differences of train travel times are affected by dwell times, which are dependent on train arrivals. In this regard, the decision variables \((t^d_{x,i} \text{ and } t^d_{x,i})\) are determined by the arrival headways at the first station and the segment travel times. Based on the cost functions, the segment travel times here are optimized in similar ways to the study in [10], and separately from the arrival headways. The headways are optimized as follows.

The continuous time is divided into seconds. The Determining process of Passenger Boarding and Alighting (D-PBA) is used to obtain dwell times, and is shown as follows.

<table>
<thead>
<tr>
<th>Algorithm 1: D-PBA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1:</strong> For train ( i ) at station ( x ) with remaining capacity ( C_i ), parameter ( \mu_y ) indicates the ratio between the number of waiting passengers with destination ( y ) and total waiting passengers.</td>
</tr>
</tbody>
</table>
| **Step 2:** Determine boarding passenger for train \( i \) at station \( x \):

\[ N^y_{i,x} = \min(c^i_{x}, N^y_{i,x}), \forall i \in [1, I], x \in [1, 2X]. \] \hfill (12)

**Step 3:** Determine the number of boarding passengers with destination \( y \):

\[ N^y_{i,x,y} = N^y_{i,x} \cdot \mu_y, \forall i \in [1, I], 1 \leq x < y \leq 2X. \] \hfill (13)

**Step 4:** Obtain the number of alighting passengers:

\[ N^a_{i,x} = \sum_{y=2}^{X} N^y_{i,x,y}. \] \hfill (14)

For model constraints, Eq. 9 can be satisfied by setting the travel time of segment \( X \) as a constant, Eq. 10 and Eq. 11 are considered in the D-PBA. Thus, only Eq. 8 should be integrated by the AL. Considering Eq. 8, the objective function (Eq. 7) is rewritten as follows.

\[ \min T_c = c_w \cdot t_u + c_t \cdot t_t + c_e \cdot E + \frac{1}{2\sigma} \sum_{j=1}^{J-1} \{ \max(0, \omega_j - \sigma \cdot g_j) \}^2 - \omega_j^2. \] \hfill (15)
where \( \sigma \) is a sufficiently large penalty factor, \( \omega_j \) is a Lagrangian multiplier, \( g_j \) represents a headway constraint. There are \( I-1 \) constraints in total. The AL is designed as follows.

**Algorithm 2: AL**

**Step 1:** Given a feasible point \( H^{(0)} = (h_1, h_2, ..., h_{I-1}) \). Set initial multipliers \( \omega^{(0)} = (\omega_j^{(0)}) \), parameter \( \sigma \), allowable error \( \varepsilon \), factors \( \alpha \geq 1 \) and \( \beta \in (0, 1) \). Set \( k = 1 \).

**Step 2:** Set \( H^{(k-1)} \) as the initial point, and obtain the solution \( H^{(k)} \) by PS.

**Step 3:** If \( \|g(H^{(k)})\| < \varepsilon \), then terminate the process; Otherwise, go to Step 4.

**Step 4:** If \( \|g(H^{(k)})\|/\|g(H^{(k-1)})\| \geq \beta \), then set \( \sigma = \alpha \cdot \sigma \), and go to Step 5.

**Step 5:** Update the multipliers with the following equation:

\[
\omega_j^{(k+1)} = \max(0, \omega_j^{(k)} - \sigma \cdot g_j), \forall j = 1, 2, ..., I-1.
\]

Set \( k = k + 1 \), and return to Step 2.

In Step 2 of the AL, the unconstrained problem is solved by the PS. Here a step can be deemed sufficiently small when its size is smaller than 1 second. Let \( e_j = (0, ..., 1, ..., 0) \) be a basis vector, and \( f(H) \) be the value of objective function. The PS is designed as follows.

**Algorithm 3: PS**

**Step 1:** Set an initial point \( H^{(1)} \), basis vectors \( e_1, e_2, ..., e_j \), initial step size \( \delta \), speeding factor \( \alpha \), shrinking factor \( \beta \in (0, 1) \), allowable error \( \varepsilon = 1 \). \( Y^{(1)} = H^{(1)}, k=1 \) and \( j=1 \).

**Step 2:** If \( f(Y^{(j)} + \delta \cdot e_j) < f(Y^{(j)}) \), then set \( Y^{(j+1)} = Y^{(j)} + \delta \cdot e_j \), and go to Step 4; Otherwise, go to Step 3.

**Step 3:** If \( f(Y^{(j)} - \delta \cdot e_j) < f(Y^{(j)}) \), set \( Y^{(j+1)} = Y^{(j)} - \delta \cdot e_j \); Otherwise, set \( Y^{(j+1)} = Y^{(j)} \).

Go to Step 4.

**Step 4:** If \( j=n \), then set \( j=j+1 \), and return to Step 2; Otherwise, go to Step 5.

**Step 5:** If \( f(Y^{(n+1)}) < f(H^{(k)}) \), then go to Step 6; Otherwise, go to Step 7.

**Step 6:** Set \( H^{(k+1)} = H^{(n+1)}, Y^{(1)} = H^{(k+1)} + \alpha \cdot (H^{(k+1)} - H^{(k)}) \). Set \( k=k+1 \), return to Step 2.

**Step 7:** If \( \delta \leq \varepsilon \), then terminate the process, and \( H^{(k)} \) is the solution; Otherwise, set \( \delta = \beta \cdot \delta \), \( Y^{(1)} = H^{(k)}, H^{(k+1)} = H^{(k)}, k=k+1, j=1 \), and return to Step 2.

**Case Study**

To test the effectiveness, the technique is performed on a metro line with 4 stations. The segment lengths are 1800m, 1600m and 2000m. The AL+PS and the uniform results are shown in Table 2. Table 3 shows the reducing ratios between the non-uniform and uniform timetables, where \( \mu_w \) represents the ratio of passenger wait time, \( \mu_r \) the ratio of train travel time, \( \mu_e \) the energy, and \( \mu_c \) the total cost. Figure 1 shows the convergences of AL+PS with different numbers of trains.

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>( t_e ) (s)</th>
<th>( t_i ) (s)</th>
<th>E (Kwh)</th>
<th>( T_e ) (yuan)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform 7 trains</td>
<td>4.0491\times10^7</td>
<td>14277</td>
<td>2.2167\times10^4</td>
<td>2.2987\times10^4</td>
</tr>
<tr>
<td>Non-uniform 7 trains</td>
<td>3.3598\times10^7</td>
<td>14277</td>
<td>2.2355\times10^4</td>
<td>1.9159\times10^4</td>
</tr>
<tr>
<td>Uniform 10 trains</td>
<td>2.5770\times10^7</td>
<td>19902</td>
<td>3.1117\times10^4</td>
<td>1.5005\times10^4</td>
</tr>
<tr>
<td>Non-uniform 10 trains</td>
<td>2.1403\times10^7</td>
<td>19902</td>
<td>3.1139\times10^4</td>
<td>1.2579\times10^4</td>
</tr>
<tr>
<td>Uniform 12 trains</td>
<td>2.1528\times10^7</td>
<td>23380</td>
<td>3.6676\times10^4</td>
<td>1.2769\times10^4</td>
</tr>
<tr>
<td>Non-uniform 12 trains</td>
<td>1.5219\times10^7</td>
<td>23380</td>
<td>3.6634\times10^4</td>
<td>9.2573\times10^4</td>
</tr>
<tr>
<td>Uniform 15 trains</td>
<td>1.5348\times10^7</td>
<td>28463</td>
<td>4.4671\times10^4</td>
<td>9.5119\times10^4</td>
</tr>
<tr>
<td>Non-uniform 15 trains</td>
<td>9.3779\times10^6</td>
<td>28243</td>
<td>4.4671\times10^4</td>
<td>6.1905\times10^4</td>
</tr>
<tr>
<td>Uniform 20 trains</td>
<td>8.3617\times10^6</td>
<td>37318</td>
<td>5.7782\times10^4</td>
<td>5.9311\times10^4</td>
</tr>
<tr>
<td>Non-uniform 20 trains</td>
<td>6.1363\times10^6</td>
<td>36966</td>
<td>5.7782\times10^4</td>
<td>4.6870\times10^4</td>
</tr>
</tbody>
</table>
Table 3. Comparison between uniform and non-uniform timetables.

<table>
<thead>
<tr>
<th>Indicators</th>
<th>7 trains</th>
<th>10 trains</th>
<th>12 trains</th>
<th>15 trains</th>
<th>20 trains</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_w )</td>
<td>17.02%</td>
<td>16.95%</td>
<td>29.31%</td>
<td>38.90%</td>
<td>26.61%</td>
</tr>
<tr>
<td>( \mu_t )</td>
<td>0</td>
<td>0</td>
<td>1.28%</td>
<td>0.77%</td>
<td>0.94%</td>
</tr>
<tr>
<td>( \mu_e )</td>
<td>-0.84%</td>
<td>-0.07%</td>
<td>0.11%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \mu_c )</td>
<td>16.65%</td>
<td>16.17%</td>
<td>27.50%</td>
<td>34.92%</td>
<td>20.98%</td>
</tr>
</tbody>
</table>

Figure 1. Convergences of AL+PS with different number of trains

Table 2 shows that the non-uniform timetables outperform the uniform ones due to more flexible headways. Their relations are further demonstrated in Table 3. The indicator \( \mu_w \) points that the passenger wait time is considerably reduced in each scenario, with the reduction ratio at least 16.95% and at most 38.90%. The indicator \( \mu_f \) demonstrates that as the number of trains increases, flexible headways reduce dwell times. The indicator \( \mu_e \) shows that the numbers of served passengers are nearly the same in uniform and non-uniform cases. The indicator \( \mu_c \) demonstrates that the total costs of non-uniform timetables obtained by the AL+PS are greatly reduced compared to uniform ones.

In addition, the total cost decreases as the number of trains increases, which demonstrates that the passenger demands are large. It is consistent with practical experiences that the rail transit operators schedule more trains in peak hours to improve the service level.

Finally, in Figure 1, all cases converge to the minimum values. It should be noted that in each iteration of PS, the algorithm terminates when the step size is smaller than 1s, which considers a tradeoff between solution quality and computation efficiency. If a more accurate solution is required, the step size can be set smaller. The operators can determine the desired time and accuracy by setting the values of parameters in the AL+PS method. It also can be seen that in some scenarios a sharp descending direction is found after the 30th iteration. Combining the numerical results, to obtain a good solution, the minimum iteration size should be at least 60 (for this short metro line, could vary with different cases). In this study, they all terminate before the iteration reaches 100.

All in all, the effectiveness of the proposed AL+PS method is demonstrated in the case study.

**Conclusion**

This paper formulates the non-uniform train scheduling problem with a constrained, non-smooth, nonlinear programming model, which is transformed into an unconstrained one and solved by the proposed numerical optimization technique. The technique consists of augmented Lagrangian (AL) and pattern search (PS) algorithm, with results outperforming the uniform timetables and consistent with practical experiences. The numerical results verify the effectiveness of the technique. This technique can be further developed as a decision-making tool for operators. Future studies can be focused on a more complex scheduling problem integrated with line planning and vehicle scheduling.
Acknowledgements

This research was financially supported by the National Natural Science Foundation of China (Grant Nos. U1434209, 71621001), the China national 13th Five-Year Plan (2016YFB1200402), and the Fundamental Research Funds for the Central Universities (2015JBZ006).

References


