Danmarks Tekniske Universitet

42117 Transport Optimization
Project 3: Rolling Stock Scheduling and Column Generation

Name, student number	Responsibilities
Malthe Jelstrup, s184291	Task 3, 4, 5, 6, 7, 11
Vera Nijmeijer, s213011	Task 1, 2, 8, 9, 10

Wednesday 9th November, 2022

Task 1. Defining the Problem and Solution Approach

1.

Sets	
$\mathcal{S}$	Stations
$\mathcal{T}$	Train unit types
$\mathcal{A}$	Arcs
$\mathcal{A}^r$	Ride arcs
$\mathcal{A}^s$	Stay arcs
$\mathcal{A}^{a,s}$	Artificial arcs from source to first station
$\mathcal{A}^{a,e}$	Artificial arcs from the last station to the sink
$\mathcal{P}$	Paths
$\mathcal{P}_i^s$	Paths starting at station $i \in S$
$\mathcal{P}_i^e$	Paths ending at station $i \in S$
$\mathcal{A}_p$	Set of ride arcs in path p
Parameters	
α_p^a	1 if path p contains arc $a \in A$, 0 otherwise
c_t	Cost of train units of type $t \in T$
c^p	Penalty cost of passenger that can't get a seat
l_t	Length of train units of type $t \in T$
s_t^f	Number of seats provided by train units of type $t \in T$ in first class
s_t^s	Number of seats provided by train units of type $t \in T$ in second class
d_a^f	Passenger demand for arc $a \in A^r$ for first class
d_a^s	Passenger demand for arc $a \in A^r$ for second class
u_a	Maximum train length for arc $a \in A^r$
Variables	
$y_{p,t}$	Number of times path $p \in P$ is used for train units of type $t \in T$
δ_a^f	Missing demand for arc $a \in A^r$ for first class
δ_a^s	Missing demand for arc $a \in A^r$ for second class

$$
\begin{aligned}
\text{minimize} \quad & \sum_{t\in T} c_t \sum_{p\in P} y_{p,t} + \sum_{a\in A^r} \left(\delta_a^f + \delta_a^s\right) c^p && && \text{(1a)}\\
\text{subject to} \quad & \sum_{t\in T} l_t \sum_{p\in P} \alpha_p^a y_{p,t} \le u_a && \forall a \in A^r, && \text{(1b)}\\
& \sum_{t\in T} s_t^f \sum_{p\in P} \alpha_p^a y_{p,t} + \delta_a^f \ge d_a^f && \forall a \in A^r, && \text{(1c)}\\
& \sum_{t\in T} s_t^s \sum_{p\in P} \alpha_p^a y_{p,t} + \delta_a^s \ge d_a^s && \forall a \in A^r, && \text{(1d)}\\
& \sum_{p\in P_i^s} y_{p,t} - \sum_{p\in P_i^e} y_{p,t} = 0 && \forall t \in T, i \in S, && \text{(1e)}\\
& y_{p,t}, \delta_a^f, \delta_a^s \in \mathcal{Z}^+ && \forall p \in P, t \in T && \text{(1f)}
\end{aligned}
$$

The objective function 1a minimizes the total cost of the trains being used plus the penalty cost for missing demand. Constraint 1b ensures that the length times the number of carriages used on a path is not higher than the maximum length that's allowed on that arc. Constraint 1c and 1d assign a value to δ_a^f and δ_a^s if the demand is not met. Constraint 1e ensures that the number of trains staying overnight at a station is the same every night. Constraint 1f sets the variables to non-negative integer values.

2.

For the reduced master problem $\mathcal{P}$ is replaced with a set $\mathcal{P}' \subseteq \mathcal{P}$. The variables $y_{p,t}$, δ_a^f and δ_a^s don't have to be integers in the reduced master problem, but only have to be values greater than or equal to 0.

For the pricing problem the reduced cost of a path for a train unit type is:

$$
c_t - \sum_{a\in A_p} l_t \pi_a - \sum_{a\in A_p} s_t^f \lambda_a - \sum_{a\in A_p} s_t^s \lambda_a - \mu_{t,i^s} + \mu_{t,i^t} \tag{2}
$$

The values used as input for this function are the dual variables of the reduced master problem:

- π_a: dual variable of constraint 1b
- λ_a^f and λ_a^s: dual variables of respectively constraint 1c and 1d
- μ_{t,i^s} and μ_{t,i^t}: dual variables of constraint 1e

The decision variable in the pricing problem is $y_{a,t}$, a binary variable that is 1 if there is one or more trains of type t using path a, and 0 otherwise.

$$
\begin{aligned}
\text{minimize} \quad & c_t - \sum_{a \in A_p} l_t \pi_a - \sum_{a \in A_p} s_t^f \lambda_a - \sum_{a \in A_p} s_t^s \lambda_a - \mu_{t,i^s} + \mu_{t,i^t} && \text{(3a)} \\
\text{subject to} \quad & \sum_{a \in A^{a,s}} y_{a,t} = 1 \qquad \forall t \in T, && \text{(3b)} \\
& \sum_{a \in A^{a,e}} y_{a,t} = 1 \qquad \forall t \in T, && \text{(3c)} \\
& y_{a-1,t} = y_{a,t} \qquad \forall t \in T, a \in A^r, a \in A^s, a \in A^{a,e} | a > 1, && \text{(3d)} \\
& y_{a,t} \geq 0 \qquad \forall a \in A, t \in T && \text{(3e)}
\end{aligned}
$$

The objective function 3a minimizes the reduced cost. Constraint 3b and 3c ensure that the path has one arc starting at the source and one arc ending at the sink. Constraint 3d makes sure that the y-variable for an arc a is equal to the y-variable in the previous arc. Constraint 3e sets the decision variable to non-negative values.

Task 2. Implementing the Column Generation Procedure

3.

The time-space graph for the problem can be seen in Figure 1 below. The graph contains 955 nodes and 1358 edges.

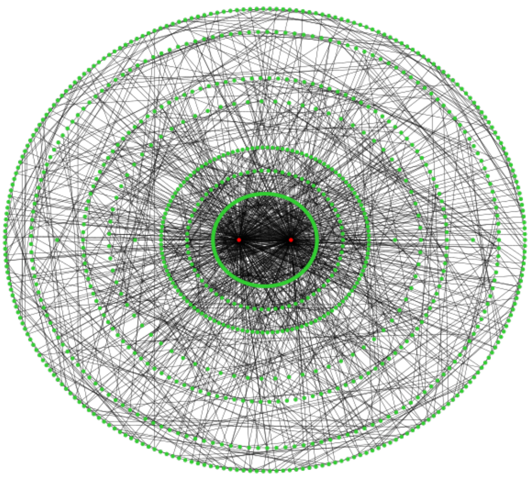

Figure 1: Time-space graph for problem

4.

The reduced master problem as described in Equation 1a-1f has been implemented in Julia, which can be found on Github [1]. The number of variables depends on the size of the input,but at the initial step with an input of size $1358x1$ (a column of zeros) the model has 27188 variables and 6810 constraints.

5.

The first three iterations of the column generation procedure was executed and the resulting information can be found in table 1. It is apparent that the column generation procedure doesn't reduce the cost in the first iteration, but the addition of train 621 allows for a reduction in cost on the second iteration. The third iteration also sees a cost reduction.

Table 1: Results of first three iterations of the column generation procedure

		Trains Added			
Iteration	Columns Found	Type 1	Type 2	Cost	Reduced Cost
init	0	-	-	78261.5	-
1	2	621	621	78261.5	0
2	2	622	618	76890.1	1731.4
3	2	625	625	76305.1	585

6.

A local optimum for the column generation procedure was achieved after 47 iterations. The resulting number of trains used and passengers that are not offered a seat in is found in Table 2. The optimal solution uses 86.56 trains of type 1 and 13.88 trains of type 2. $17,226.26$ first class passengers and $74,182.64$ second class passengers were not offered a seat. Table 3 shows the number of paths containing a number of arcs per train type. There is a noticeably larger number of paths with 4 arcs. Given that you pay a fixed cost for a train and not a cost per arc this outcome makes sense, as running a train on those particular paths would be more cost effective.

Table 2: Number of used trains of each type and the number of passengers that are not offered a seat at local optimum

# Trains		# Passengers Without Seat	
Type 1	Type 2	First Class	Second Class
86.56	13.88	17226.26	74182.64

Table 3: Number of paths containing a number of arcs per train type

# Paths with # Arcs for Type 1			# Paths with # Arcs for Type 2		
2 Arcs	3 Arcs	4 Arcs	2 Arcs	3 Arcs	4 Arcs
2 Paths	4 Paths	41 Paths	0 Paths	0 Paths	11 Paths

7.

Using the columns found in the column generation procedure to find a integer solution results in an upper bound of the optimal solution of 60405. Table 4 shows a slight increase in the use of type 1 trains and a slight decrease in the use of type 2 trains. This results in a reduction of first class passengers and increase in second class passengers without seat.

Table 4: Number of used trains of each type and the number of passengers that are not offered a seat found with integer solution

# Trains		# Passengers Without Seat	
Type 1	Type 2	First Class	Second Class
94	12	17098	75412

Task 3. Adjusting your Prototype Based on Feedback from NS

8.

The initial and reduced master problem used are the same model as model 1 with the addition of one constraint, shown in equation 4, with Ut referring to the station Utrecht.

$$\sum_{p \in P_i^s | i = Ut} \sum_{t \in T} y_{p,t} \leq 8 \tag{4}$$

9.

When running the model with the new constraint the optimal solution has the same objective value as it had without the constraint. The fact that it didn't affect the objective value is most likely due to a small error in the model. The cost of the restriction for NS would otherwise be the difference between the original objective value and the new objective value.

10.

Because the limit of 8 trains being allowed to be parked at night in Utrecht did not affect the objective value, it is hard to perform a good sensitivity analysis. The objective value is calculated for a limit of trains ranging from 1 to 8 in Utrecht. The only value for which the objective function was lower than in the original value is when the maximum number of overnight trains in Utrecht is set to 1. The cost for that problem is 59380.66.

11.

Branch-and-Price could be used to find the optimal integer solution by branching on every non-integral decision variable and solving the linear relaxation at each node through performing column generation and continuing this until no non-integral variables remain. An example of a branching decision one might make when using such an approach could be as follows: Let's assume that that the solved relaxation of RMP, when no more columns with negative reduced cost can be generated, includes 3.5 type 1 trains on path 1. We branch on this type-train-path combination setting it to 3 and 4 respectively. For each branch we now solve the relaxation of the RMP again, because the change in this particular type-train-path combination could have enabled the generation of another negative reduced cost column. We continue branching on each non-integer decision variable like this until non remain. This way the optimal integer solution can be found.

References

[1] M. Jelstrup and V. Nijmeijer, "42117 Transport Optimization Project 3." https://github.com/Stinth/Transport-Optimization.