An applied optimization based method for line planning to minimize travel time

Bull, Simon Henry; Rezanova, Natalia Jurjevna; Lusby, Richard Martin ; Larsen, Jesper

Publication date:
2016

Citation (APA):
Bull, S. H., Rezanova, N. J., Lusby, R. M., & Larsen, J. (2016). *An applied optimization based method for line planning to minimize travel time*. DTU Management Engineering.

An applied optimization based method for line planning to minimize travel time

Simon H Bull Natalia Rezanova Richard Lusby
Jesper Larsen

May 26, 2016

Abstract

The line planning problem in rail is to select a number of lines from a potential pool which provides sufficient passenger capacity and meets operational requirements, with some objective measure of solution line quality. We model the problem of minimizing the average passenger system time, including frequency-dependent estimates for switching between lines, working with the Danish rail operator DSB and data for Copenhagen commuters. We present a multi-commodity flow formulation for the problem of freely routing passengers, coupled to discrete line-frequency decisions selecting lines from a predefined pool. We show results directly applying this model to a Copenhagen commuter rail problem.

1 S-tog problem description

The S-tog network in Copenhagen is a commuter rail network serving 84 stations and between 30,000 and 40,000 passengers per hour at peak times. The S-tog network is operated by the Danish state railway operator: DSB. The trains in the network operate on published lines which each have an hourly frequency, and run according to a published timetable. We consider the lines and the frequencies, but not the exact timetable.

See Figure 1 for an example of the lines that may operate in the S-tog network. Here each coloured path refers to a different line that is operated at some frequency, and on each a train visits every station marked on the line in each direction according to that frequency. An important feature of the network is that a train may not necessarily stop at every station it passes; for example the red and orange lines (C and H) run parallel to each other in the top left of the figure, and to the same end station, but the red line stops at fewer stations and is therefore faster between stations. This is of benefit to passengers travelling past those stations, but passengers travelling to or from the skipped stations are left with fewer options. For our purposes, the possibility of different stopping patterns greatly increases the problem size.

Given the fact that lines may not stop at all stations, a route or path in the network is not sufficient to define a line. We define a line here as a sequence of tracks which a train passes, and the stations on those tracks that are stopped at and not stopped at. We refer to the sequence of tracks as a line route, and the stations stopped at and not stopped at are the line stopping pattern. Paired

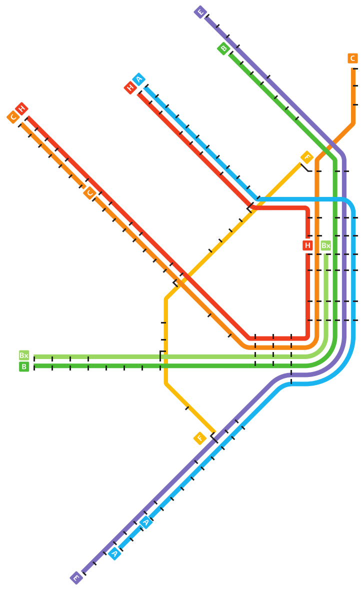

Figure 1: An example of the lines operated in the S-tog network, showing different lines in different colours. Each is identified by letter, and the presence of a dash indicates a stop.

with every line (route and stopping pattern) is an hourly frequency, and the set of lines with their frequencies defines the line plan.

Of the roughly 7,000 pairings of stations we consider non-zero demand for passengers between just over 4,600 of the pairs, which is around 65% of the possible demand pairings. As input we take a set of 174 valid lines, each with one or more valid frequencies at which the line could run; in total 350 line-frequency combinations are considered. Each line services between 11 and 39 stations with an average of 23 stations served per line, and almost all lines can operate at exactly two frequencies, while some very small number have more possible frequencies. However we also experiment with more frequencies.

Our demand data is for a specific peak period of the day, where in reality demand varies throughout the day. Real S-tog line plans have lines that operate at different frequencies at different times during the day, and operate modified line plans during the weekends. We make the following two assumptions, which are true for the S-tog problem:

- a line plan created for a peak time is valid at other off peak times, possibly operating at lower frequency;
- a line operates in both directions at the same frequency.

For the second point, we model both directions of the line as having the same frequency, where in practice it may be possible to operate each direction at a different frequency, though balancing vehicle movements may be more complicated.In practice current plans operate both directions of a line at the same in frequency, and we model the problem in that way. However, for capacity reasons the different directions of a line may not use the same rolling stock unit type; here we do not model rolling stock units but instead take a fixed capacity of the largest rolling stock unit type available. In practice there are no other differences between units that are relevant for our purposes here (for example plans are made for a fixed driving speed, not dependent on unit type) and through discussions with the operator we are confident that assuming all units are of the maximum size is a valid simplification. Assumptions on rolling stock unit are necessary because the line planning problem is solved early in a sequence of rail planning problems, usually followed by timetable creation and then rolling stock planning, and exact rolling stock details are only known at that stage. Without a timetable which itself depends on the line plan, we can only make estimates or assumptions about those details.

The presence of mixed stopping pattern lines running on the same infrastructure lines has the potential to negatively impact the timetable due to the mixed driving times of trains. This could be a source of a lack of *robustness* for the timetable and line plan. However, we do not calculate a timetable to assess this impact and make no estimate or derive any metrics for this particular feature.

The line planning problem we consider is that of selecting a set of lines from a larger pool of lines, and for each selected line, determining an hourly frequency at which the line should operate. A line is defined as a route in the infrastructure network with a stopping pattern, and several lines may share the same route but have different stopping patterns. The selected line plan must meet certain criteria, such as provide a minimum hourly service at each station, not exceed hourly limits on trains using certain track segments, visiting certain stations and turning at certain stations. These criteria are required due to the traffic contract between the operator DSB and the Danish Ministry for Transport

The line plan must also have sufficient capacity to transport all expected passengers, providing all with a good path from their origin to destination. As an objective measure we want to model the entire travel time of passengers, and include a frequency-dependent cost component to penalize occurrences of passengers switching to lines at low frequency, in favour of switching to lines at high frequency.

We take the line pool as a fixed input, and this is a subset of every possible route and stopping pattern. The lines in the pool are all assumed to be feasible alone, and are a sensible restriction of the entire pool excluding lines that are not considered to be appropriate or feasible. The pool is still sufficiently large to ensure the presence of a large range of feasible line-plan solutions.

In this paper, we present a integer program formulation for the line planning problem, formulated to allow a primary objective of total passenger travel time defined as time travelling in train and time switching between lines. We present a flow based formulation for routing passengers, and show that, despite being large, we use the formulation directly for real-world instances without requiring a path-based decomposition. However to reasonably solve the model, we present some contractions of the flow graph, present how we can aggregate passengers to reduce the problem size, and present additional constraints that improve the lower bound.

We present our results for the Danish S-tog case study, where we show a range of different line plans and how they can be compared with different metrics.

2 Related Research

what about Goossens et al. [2004]

expand more!

There is much work in the literature on the line planning problem, with different details and objective measures. Schöbel [2012] gives an overview of line planning in public transport, and classifies different problem instance characteristics and models for addressing them.

In many line planing problems, the line and the route are interchangeable; if a line follows a certain route, then every station on the route is serviced. Goossens et al. [2006] however present several models for the train line planning problem where a route may have different stopping patterns, with a cost focus rather than a passenger focus. Other work at DSB for the S-tog problem [Rezanova, 2015] has used a version of one presented model to find low cost line plans.

There is a more recent focus on the passenger, and on minimizing the total trip time for passengers (and so modelling their moving and switching times) as we would like to do for the S-tog system. Schöbel and Scholl [2006] present a model with a station-line graph in which passengers are freely routed with both travel times and switching penalties, with continuous frequency decision variables for lines, and use a decomposition such that decisions are made in terms of path selection for every pair of stations. Nachtigall and Jerosch [2008] present a model where passengers are routed freely, measuring travel time with fixed penalties for switching lines, and with an integer decision variable per line. Borndörfer et al. [2007] similarly presents a formulation freely routing passengers (though ignoring transfers) and also dynamically generating lines, using continuous frequency decision variables. In contrast we wish to model the switching cost with a frequency dependence. Also, the line frequencies in the S-tog system are not so free that we can model them continuously as there are only discrete frequencies that are considered valid. In addition to passenger travel time, Borndörfer et al. [2007] use operator running cost and fixed line-setup costs into account in their model.

A simpler passenger-focused measure is using a direct-travellers objective, maximizing the number of direct travellers the network may transport (i.e. passengers who may travel with no transfers). Bussieck et al. [1997] present such an approach while, like many authors, selecting lines from a pool (by selecting for each line in the pool an integer frequency, which may be zero). A pool-based approach is used by many authors, and is used in this work.

Robustness of subsequent plan Goerigk et al. [2013]

3 Lines Model

We take as input a set of valid lines, $\mathcal{L}$, and for each line l there is a predefined set of discrete frequencies at which the line could operate: $\mathcal{F}_l$.

Ignoring passengers, we may simply find line plans (pairings of lines with frequencies) which satisfy all operational limits, and consider how well they serve passengers. In general such solutions do not even guarantee sufficient capacity for all passengers, though often they are very close; the minimum visits requirement per station in many areas provides more capacity than there are passengers travelling to, from or passing by the station. However, even if a solution does provide sufficient capacity, it is possibly a very poor quality solution for passengers.

We decide which lines and frequencies we will select from the line pool $\mathcal{L}$, where each has valid frequencies $\mathcal{F}_l$ (defined for each $l \in \mathcal{L}$). We let the binary decision variable $y_{lf} \in \{0, 1\}$ denote selecting line l at frequency f.

Simply selecting a valid set of lines is not in itself trivial; the selected lines must be compatible, must meet certain service levels, and must not exceed some fixed operating budget. The service level requirements can all be expressed as a minimum number of trains visiting a single station per hour, or operating on a particular track sequence per hour. Similarly, the compatibility requirements can be expressed as a maximum number of trains per hour visiting stations, turning at stations, and operating on particular tracks. Selecting a line at frequency f contributes f trains per hour towards the relevant service level constraints, and so we can enforce such constraints by summing over every line frequency decision with the frequency itself as the coefficient.

Every contractual requirement or operational limit can be expressed by determining exactly those lines which contribute toward the requirement or limit such as lines visiting the relevant station or using the relevant track sequences. Consider such a set $\mathcal{Z}$ of lines. The contractual requirement or operational limit for $\mathcal{Z}$ may have either a lower limit or an upper limit or both for the number of trains per hour. For simplicity in definition we assume both; let these be $\alpha(\mathcal{Z})$ and $\beta(\mathcal{Z})$ for the lower and upper bounds, respectively. Now, let $\mathcal{C}$ be the set of all such sets $\mathcal{Z}$; every element of $\mathcal{C}$ is a set of lines $\mathcal{Z}$ with a lower ($\alpha(\mathcal{Z})$) and upper ($\beta(\mathcal{Z})$) hourly limit.

Additionally, certain sets of lines are inherently incompatible for various reasons not explicitly related to the line plan but for other operational reasons. Let $\mathcal{I}$ be the set of all incompatible sets of lines, where any element of $\mathcal{I}$ is a set of lines from which only one line can appear in a valid line plan.

Finally, every line has a cost when operated at a particular frequency: c_{lf}, and we impose a maximum budget for the line plan $c_{\max}$. This generalized cost may not necessarily scale with frequency; selecting a line at frequency $2f$ may cost more or less than selecting the line at frequency f.

Now, the following constraints define a valid line plan, considering only the lines themselves but ignoring passengers.

$$\sum_{f \in \mathcal{F}_l} y_{lf} \leq 1 \qquad \forall l \in \mathcal{L} \tag{1}$$

$$\sum_{l \in \mathcal{L}} \sum_{f \in \mathcal{F}_l} c_{lf} \cdot y_{lf} \leq c_{\max} \tag{2}$$

$$\sum_{l \in \mathcal{Z}} \sum_{f \in \mathcal{F}_l} y_{lf} \leq 1 \qquad \forall \mathcal{Z} \in \mathcal{I} \tag{3}$$

$$\sum_{l \in \mathcal{Z}} \sum_{f \in \mathcal{F}_l} f \cdot y_{lf} \geq \alpha(\mathcal{Z}) \qquad \forall \mathcal{Z} \in \mathcal{C} \tag{4}$$

$$\sum_{l \in \mathcal{Z}} \sum_{f \in \mathcal{F}_l} f \cdot y_{lf} \leq \beta(\mathcal{Z}) \qquad \forall \mathcal{Z} \in \mathcal{C} \tag{5}$$

$$y_{lf} \in \{0, 1\} \qquad \forall l \in \mathcal{L}, \quad \forall f \in \mathcal{F}_f \tag{6}$$

Constraints (1) ensure that a given line is chosen at most once disallowing a single line at multiple frequencies (because, for example, some line might be permitted at 3, 6, or 12 times per hour but not at 9 times per hour, so combinations may not be permitted). Constraint (2) ensures that the total lines cost is no greater than the given budget. Constraints (3) permit only one line for each of the sets of incompatible lines. Similarly, constraints (4) provide minimum service levels for the same visits, turnings or track usages. Constraints (5) provide all operational constraints that can be expressed as a maximum number of trains visiting or turning at a station, or using a specific sequence of track.

The formulation (1)–(6) defines a valid line plan. It completely ignores passengers; some feasible solutions to the formulation will fail to provide sufficient capacity for all passengers in the network, and those that do provide sufficient capacity may nevertheless provide a poor solution for many passengers. However, solving the formulation will find a line plan with some capacity that services all stations, so it can be assessed to determine whether or not it does provide sufficient capacity. If so, we can assess how well it serves passengers, and if not we can identify where capacity is lacking.

4 Passengers

4.1 Graph

We model passenger travel as a movement in a graph, where the existence of components of the graph depends on the presence of a line in the solution. We could model each line-frequency pair as a completely distinct component of the graph. However, this leads to a very large graph, especially if we want to experiment with many frequencies for each line, and much of the information depends on the line itself and not its frequency.

Consider Figure 2 showing the structure of a single line at a single frequency visiting three stations. For each station (1, 2, 3) there are three vertices; a source vertex, a sink vertex, and a platform vertex (s_i^1, s_o^1, p^1 for station 1, respectively). All passenger paths originate from some source vertex, and terminate at some other sink vertex, travelling on dashed line travel edges or switching lines using a platform vertex. To capture the information we want about frequency-dependent aspects of the line, we could simply duplicate this structure for every frequency at which the particular line operates. That is, we would have a parallel structure to l_1, l_2, ... vertices representing the same line with route and stopping pattern, but operating at a different frequency. However, much of the information would be redundant, and when experimenting with large numbers of frequencies per line the graph becomes very large. Alternatively we could simply have one such structure that represents every frequency, except that then the cost of a particular path could have no dependence on frequency of lines used. In our problem we want to penalise switching to low frequency lines more than high frequency lines. However capacities of edges, though dependent on frequency, can still be maintained even with a single structure by summing the capacities of the frequency-line decisions that would contribute toward them. This suggests that it is possible to partially aggregate the line-frequencies into simply lines, being careful to accommodate the frequency-dependent switching cost between lines.

The aggregated graph then contains three types of node:

- a source and sink for every station;

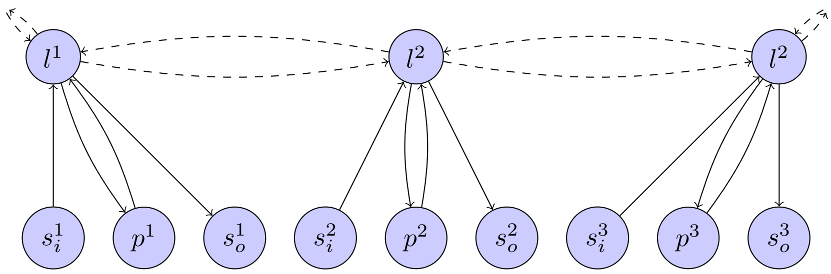

Figure 2: The structure for a single line at one frequency visiting multiple stations.

- a platform for every station;
- a station-line for every station a line visits, for every line.

The graph also contains several types of edges:

- A travel edge between every adjacent pair of station-line edges for every line, in each direction;
- A get-off edge from every station-line to every station sink;
- A get-off edge from every station-line to every station platform;
- A get-on edge from every station source to every station-line;
- A get-on edge from every station platform to every station-line at every frequency.

Note that this is an aggregation of the line/frequency combinations, though without being able to aggregate those boarding frequency edges. It means the capacity of an edge (the station-line to station-line edges) is dependent on a summation over all frequency decisions for that line.

The graph structure is similar to the stop-and-go graph described by Schöbel and Scholl [2006]. See Figure 3 for the structure of the problem graph for passengers. The figure shows a single station with its three nodes, and two distinct lines which visit the station at two frequencies each. Here depicted as a multi-graph, the graph can be made simple with auxiliary nodes and edges. For each passenger, a path through the graph from their origin station s_i^1 vertex to their destination station s_o^2 vertex must be found, which incurs the travelling time (on dashed edges) and switching time costs (on red edges). Differing from the Schöbel and Scholl [2006] problem structure, in our case the discrete frequencies a line may operate at are an important feature, and we want to model different passenger time costs for switching to lines at different frequencies, so our graph has additional station structure.

4.2 Flow decisions

The problem can be represented as a multi-commodity flow problem, with one commodity per OD pair, with additional constraints linking flows to line presence and capacity. However, with the roughly 4,600 OD pairings in our regular problem

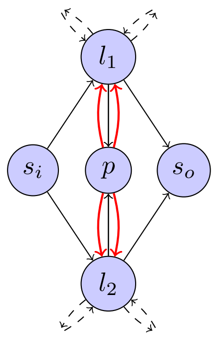

Figure 3: The graph structure for two lines at a station. Each station has three vertices, s_i and s_o, which representing either entering or exiting the system at this station, and p, which represents switching lines at the platform. Vertices l_1 and l_2 represent the two lines visiting the station. The solid black edges have zero cost, while the dashed black edges cost the travel time to the next station along a line. The thick red edges represent the costly switching from one line to another, and depend on the frequency of the boarded line (one edge per frequency the line may operate at).

instance and the relatively large graph we describe above, the problem would be very large. Let us refer to such a model, which we do not formulate here, as a per-OD arc flow model, where for every OD we would select a proportion of passengers who use every edge in the graph such that every OD has one path from origin to destination, and those edges used correspond to selected lines and frequencies. We have tested the per-OD arc flow model for small instances (such as with only the *lines* of a known feasible solution, but unspecified frequency still to be determined) and, though solvable, the model is very large and would not scale to having very many lines.

The proposed per-OD arc flow model would have one flow variable for every OD combination, for every edge in the graph, and we would require one path with capacity sufficient for that OD demand for every OD combination, respecting every other OD path. As an aggregation, we can combine flows that have the same origin (or alternatively the same destination), and instead have one type of flow for every origin. The number of flow decisions is then lower by a factor of $|\mathcal{O}|$. Instead of requiring one path per OD, we will require the aggregation of those paths scaled by passenger counts; that is, we will require a network flow from each origin which supplies the sum of passengers from the origin to every destination, and each destination from that origin consumes just the passenger demand from the origin to that destination. The flow variables are $x_o^e \geq 0$; the number of passengers from origin o using edge e. Note that we do *not* require integer flows, and we do not require a single path between every origin and destination. In fact we see for currently used plans that it is infeasible for every OD pair to use a single path, as there is insufficient capacity. Instead some proportion of passengers on some OD trips are forced to take less favourable paths than the best available due to limited capacity on their most attractive path. Here, we note that the capacity we take for an operating line is assuming the largest possible rolling stock unit is operating the line, while in reality DSB operates different units with different capacities. In discussions with DSB we determined that this simplification was appropriate and, while it potentially over-estimates the true capacity achievable for a line plan (as DSB has too few of the largest units to use them everywhere), we do not see solutions where this capacity is required on every line simultaneously.

In the graph described in Section 4.1, let $\mathcal{V}$ be the the set of vertices and $\mathcal{E}$ be the set of edges. For every vertex in the graph, we define a demand for passengers for every origin station in the network. Let the demand for passengers at vertex v who originate from station s be a_v^s.

$$a_v^{s_1} = \begin{cases} d_{s_1 s_2} & \text{if vertex } v \text{ is a sink vertex for station } s_2 \\ -1 \cdot \sum_{s_2} d_{s_1 s_2} & \text{if } v \text{ is the source vertex for station } s_1 \\ 0 & \text{otherwise} \end{cases}$$

We also require constraints that link the flow variables to the line decision variables y_{lf}, ensuring both that if any flow uses a line, then the line is present, and that every connection of the line has sufficient capacity for all flows which use it. Further, we will require that the edges corresponding to the frequency-dependent boarding of a line are only used if the line is present at the correct frequency. Constraints linking the flow variables to the capacity of the selected lines are in fact sufficient, and it is not necessary to impose additional constraints to link simply a usage of a line to a line decision variable. To do this, let $\mathcal{E}^l$ be the set of all edges in the graph that depend on the presence of line l at undetermined frequency. Let $\mathcal{E}^l_f$ be the set of all edges in the graph that depend on the presence of line l at exactly frequency f.

We impose the following constraints:

$$\sum_{(u,v)\in\mathcal{E}} x_s^{(u,v)} - \sum_{(v,w)\in\mathcal{E}} x_s^{(v,w)} = a_v^s \qquad \forall s \in \mathcal{O} \quad \forall v \in \mathcal{V} \tag{7}$$

$$\sum_{o\in O} x_o^e \leq \sum_{f\in\mathcal{F}_l} P_f y_{lf} \qquad \forall l \in \mathcal{L}, \forall e \in \mathcal{E}^l \tag{8}$$

$$\sum_{o\in O} x_o^e \leq P_f y_{lf} \qquad \forall l \in \mathcal{L}, \forall f \in \mathcal{F}_f, \forall e \in \mathcal{E}^l_f \tag{9}$$

Constraints (7) ensure that flow is conserved in the graph, for every origin defining a network flow moving the required number of passengers from each origin to every destination. Constraints (8) ensure that for every edge on a line ($\mathcal{E}^l$), the edge provides sufficient capacity for all flows using it. The constant P_f is the capacity of any line at frequency f, which we take as a constant. An obvious simple extension is to have a line-specific capacity, but that is not present in our data. Finally, constraints (9) ensure that for those boarding edges from a platform that are frequency dependent (edges $\mathcal{E}^l_f$ for line l at frequency f), again sufficient capacity must be present. In effect, the difference between constraints (8) and (9) is that (8) are for the aggregated frequency edges and therefore we sum the y_{lf} variables for all frequencies.

These constraints (7)–(9) define the flows and link them to the line decision variables. The full formulation then is constraints (1)–(6), and (7)–(9).

5 Objective functions

There are a number of possible objective functions we could use, either related to passengers, to the operator, or to some combination. Above with constraints (1)–(6) we gave the operator cost as a constraint with a fixed budget, and our primary goal is to minimize a passenger-based objective. However, other measures are possible. Here we give our primary measure and some other alternative measures we use.

5.1 Passenger travel time

As already stated, we are interested in penalising switching time for passengers with emphasis on discouraging switching to lines operating at low frequency. As we do not know the timetable in advance, we can't know the exact time required for a switch. In the ideal case, for every switching occurrence, both trains would arrive at a station at the same time and the station layout would permit passengers to switch from either train to the other, losing no time. However, generally this is impossible in the S-tog system. The best case at most stations is that one train arrives shortly before another, in such a way that passengers may switch from the earlier train to the later time with minimal waiting time, but then passengers switching in the opposite direction have almost a worst-case wait time for their next train.

Overall, we consider passenger travel time to be the most appropriate measure. In tests, if we ignore switching time and minimize only moving travel time, we find solutions with many undesirable switches required. Conversely if we ignore travel time and consider only minimizing some measure of switch cost, we find solutions which do not use "fast" lines appropriately and have higher overall average total travel time. Travel time therefore includes both the moving travel time on train lines and an additional estimate of the wait time. However, in addition to this we include a separate expression for the "unpleasantness" of switching lines which we express as a time, in effect calculating a weighted sum of estimated travel time and the number of switches.

For every edge in the graph, we assign some cost to the passenger. Let t_e be the time cost to one passenger for using edge e. For every travel edge on a line (the dashed lines in Figure 3), the edge time cost is the exact, known, travel time for trains between the two stations. However, for the frequency-dependent switching edges (the red edges on Figure 3), the edge time cost includes an estimate of the waiting time and the penalized fixed cost of switching. For such edges e at frequency f, let $t_e = p_{\text{fixed}} + \lambda \frac{1}{f}$, where $\lambda \in [0, 1]$. That is, the time cost is a fixed term with a fraction of the worst case wait time (where for example in the worst case, a line operating twice per hour has a worst case switch time of $\frac{1}{2}$ an hour). We take, as a parameter, a fixed penalty of six minutes and $\lambda = 0.5$, or an average case wait time estimate. For any other edges let $t_e = 0$. Now, we can define our objective function as simply:

$$\sum_{e \in \mathcal{E}} \sum_{o \in \mathcal{O}} t_e x_o^e \tag{10}$$

5.2 Minimum lines cost objective

We are not only interested in line plans which have minimal passenger cost, but also those that have low operating cost while still providing a good passenger service. An operating cost objective can easily be used with the formulation (2)–(6), iteratively finding many low cost solutions and assessing their passenger cost. However it is likely that such plans are not feasible; minimal line cost solutions likely provide low passenger capacity. We can use the following as our objective function, which expresses operating line cost:

$$\sum_{l \in \mathcal{L}} \sum_{f \in \mathcal{F}_l} c_{lf} \cdot y_{lf} \leq c_{\max} \tag{11}$$

We observe that there are many similar cost solutions with the same lines but with different frequencies, where the aggregated sum of frequencies is the same (or similar). If we iteratively find solutions and forbid them (which we can do with constraints described in a later section), we see many similar solutions. Also, we observe that if a low-cost solution of lines with some frequencies is infeasible for passengers, then a different low-cost solution with the same lines at different frequencies is also likely to be infeasible for passengers. However there is also likely to be other solutions with the same lines at different frequencies that *are* feasible, but not low-cost. We can also find these similar solutions by taking a limited line pool consisting only of the lines present in a low cost solution, but at any frequency, and solve a problem with passenger cost as the objective. A feasible solution is guaranteed to be present as every line and frequency of the provided solution is present, but we may find there are better feasible solutions with respect to the passenger objective.

5.3 Direct travellers objective

In model (2)–(6) we can instead use a (pseudo) direct travellers objective, seeking to maximize the number of direct travellers. This will not be entirely correct as it will count the number of passengers who have a direct path in the line plan, but without the guarantee that all such passengers may take their direct path (due to limited capacity).

To measure the direct travellers we introduce a new binary variable: let $z_{od} \in \{0, 1\}$ denote whether the line plan provides a direct connection between stations o and d. Then, let $\mathcal{L}_{od}$ be the set of lines which provide a direct connection between stations o and d, which can easily be precomputed. We then maximize the objective function:

$$\sum_{(o,d) \in \mathcal{O}} d_{od} \cdot z_{od} \tag{12}$$

where d_{od} is the demand between stations o and d.

We require the following additional constraints:

$$z_{od} \leq \sum_{l \in \mathcal{L}_{od}} \sum_{f \in \mathcal{F}_l} y_{lf} (o, d) \in \mathcal{O} \tag{13}$$

That is, z_{od} may have value 1 if there is one line providing a direct connection between o and d in the line plan.

As mentioned in the previous section, we may iteratively find solutions by forbidding those we do discover with appropriate constraints. We then assess the feasibility and travel time cost of the line plans. The results are remarkably poor; the direct traveller objective is not a good approximation of our true travel time estimate objective. For the S-tog problem we assessed the first three thousand solutions; all had very high line operating cost, and almost all had high travel time cost as well. None was a noteworthy solution; all discovered solutions are dominated by other solutions, even (some of) those found with minimal operator cost in the previous section.

Examining the discovered solutions, there are two explanations for the poor quality of the solutions favouring direct travellers. Firstly, lines that skip stations in our network provide a small travel time benefit to potentially many passengers, but provide fewer direct connections than lines with the same route stopping at all stations. If measuring only direct travellers then these faster "skipping" lines are less valued, and almost none appear in any solutions we discover. Secondly, the solutions have a very wide selection of lines, with more lines being present overall and those lines that are present all terminate at extreme end stations of the network, not intermediate end stations. This then has the problem that the line frequencies must all be low to accommodate many lines turning at each of the extreme end stations, and then by our travel time estimate, those passengers that do *not* have a direct connection are then required to switch onto lines operating at low frequency, incurring higher cost. We see that, indeed, fewer passengers are required to make connections (roughly 30% fewer connections in total than plans targeting total travel time), but that is still then a large number of passengers and most are switching to low frequency lines.

The direct travellers objective is used by other authors, and in fact has been noted as being a source of a lack of delay robustness for a timetable [Goerigk et al., 2013]. However that is attributed by the authors to the presence of very long lines in the solution which provide many direct connections. In the S-tog network there is not the possiblity for particularly long lines, although a direct traveller objective does discourage using the "intermediate" end stations on the fingers and instead using the end depots at greater capacity. However a direct travellers objective in the S-tog case discourages skipped stations which leads to more homogeneity in driving speed, which might be expected to provide better delay robustness.

6 Other modeling considerations

6.1 Additional linking constraints

In the previous section we defined a formulation where the linkage between edge flows and line presence is imposed only by the capacity of a line and the sum of all usages of every element of the line.

In general, in our problem, the demand between some particular pair of stations is lower than the capacity of a line operating at only the lowest frequency. In a non-integer solution only a small fractional line decision variable (y_{lf}) is required to provide capacity for some OD pair to make use of the line, if no other OD pair uses that line. Suppose for an OD based arc flow model there are variables x^e_{od} deciding the flow on edge e for flow from o to d. In addition to summing all such flows for every OD pair for the usage of the line y_{lf} which contains e to constrain the capacity of the edge, the following constraint could be used:

$$x_{od}^e \leq d_{od} \cdot \sum_{f \in \mathcal{F}_l} y_{lf}$$

where d_{od} is the demand for pair (o, d). This would provide a tighter linkage between the flow variables and the line variables, at the expense of requiring very many constraints, though it is not necessarily required that such constraints are included for every edge of a line.

Unfortunately, we do not have individual flow variables x_{od}^e as we have aggregated the variables by origin; we have only x_o^e. Unlike previously, where the maximum flow on any edge for one (o, d) flow was d_{od}, now the maximum demand of flow on any edge from one origin is $\sum_d d_{od}$, which is *not* in general significantly smaller than the capacity provided by one line; in fact it can often be that any one line capacity is less than the aggregated demand. The analogous constraint is much weaker:

$$x_o^e \leq \sum_{f \in \mathcal{F}_l} y_{lf} \cdot \sum_d d_{od}$$

In general, the flow originating from an origin has much higher edge usage than any single d_{od}, and close to the origin itself it may in fact be as high as $\sum_d d_{od}$. However the flow from that origin terminates at many destinations and at those destinations the flow is much lower; exactly d_{od} flow terminates at a particular destination from some origin, and that corresponds to usage of an edge in our graph that belongs to a specific line and is *only* used by flow terminating at that destination. That can be seen on Figure 3, where edges to s_o can each be associated with a single line, and where there are no edges out of s_o. Therefore, for every such specific edge e we can include the following constraint:

$$x_o^e \leq d_{od} \cdot \sum_{f \in \mathcal{F}_l} y_{lf}$$

Let t_{ld} be the terminating edge of line l at destination station d (if there is destination d on line l). Then, we impose the following for every line and for every (o, d):

$$x_o^{t_{ld}} \leq d_{od} \cdot \sum_{f \in \mathcal{F}_l} y_{lf} \tag{14}$$

Given our tight operational constraints, as well as the budget constraint (constraints (2) and (5)), such constraints improve the bound given by solving the LP relaxation of the model, as in general the forcing of some line variables to have higher value must cause a decrease in others, and then some passengers must use less favourable lines. However this comes with the addition of many new constraints; one for every OD pair, for every line that visits the destination of the pair (up to $|\mathcal{L} \times \mathcal{O} \times \mathcal{O}|$ constraints).

The formulation is constraints (1)–(6), (7)–(9) and (14).

Figure 4 shows how the LP lower bound differs with and without constraints (14). Here the problem itself is the S-tog problem, and more precise details are given later. The effect differs with the budget limit: without the constraints only a very strict budget limit has any effect on the LP bound, while with the constraints the bound is affected more with a wider range of budget limit levels. The difference between the two bounds is most pronounced at lower budget limits than higher, but in all cases is less than half way from the worse lower bound to a best IP solution (though the IP solutions were only obtained with a time limit of 1 hour and so not all are optimal). Note that for the IP solutions, the actual lines cost may not correspond exactly to one of the cost constraints applied, as with the applied constraint the best solution found may have a lower lines cost.

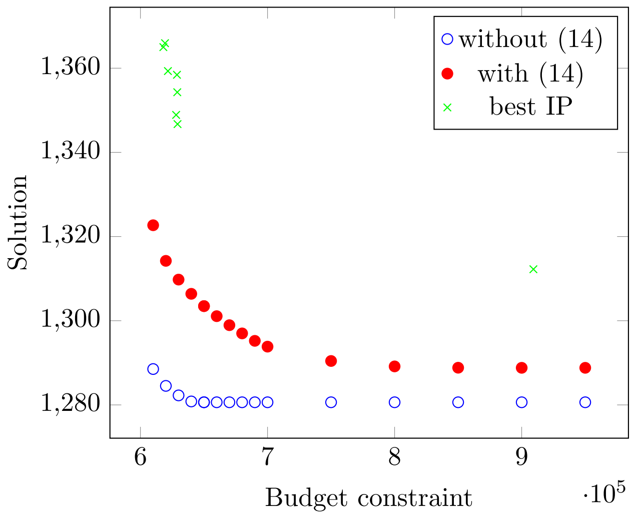

Figure 4: The LP relaxation lower bound found for different cost constraints with (blue open circle) and without (red filled circle) constraints (14). Green cross-marks show some best integer solutions found with a one hour time limit for a range of cost constraints.

6.2 Forbidding a solution

Let $\mathcal{S}$ be a set of lines in a particular solution, and let f_l be the frequency of line l in $\mathcal{S}$. The following constraint forbids this exact solution:

$$\sum_{l \in \mathcal{S}} y_{l,f_l} \leq |\mathcal{S}| - 1 \tag{15}$$

This has the potential problem that it does not forbid a solution containing only some of the lines in $\mathcal{S}$, or conversely that it does forbid solutions which contain the lines of $\mathcal{S}$ and additional lines. As an alternative, a solution with the given solution lines at *any* frequency can be forbidden with the following constraint:

$$\sum_{l \in \mathcal{S}} \sum_{f \in \mathcal{F}_f} y_{lf} \leq |\mathcal{S}| - 1 \tag{16}$$

As before, this does not forbid a solution containing only some of the lines in $\mathcal{S}$, and does forbid solutions with additional lines. However, we find that to be acceptable for our problems.

Forbidding solutions in this way is used to remove solutions known to be infeasible for later planning; and iteratively, to find a large pool of solutions by repeatedly solving the model, forbidding the last found solution, and solving again.

6.3 OD grouping

Stations can be grouped together if they are served similarly by all lines. Consider two adjacent stations, s_1 and s_2, which lie on a track sequence, and all lines in the line pool stop at either *both* s_1 and s_2 or *neither* s_1 nor s_2. That is, they are served identically by all lines. Then, if there is a third station s_3 with demand to both stations s_1 and s_2 we can treat the two demands as a single combined demand to (say) s_1. Any demand for travelling directly from s_1 to s_2, or vice versa, which would be discarded, can be reserved by requiring sufficient aggregated extra capacity on the lines visiting both stations. This may under-reserve capacity on the connection between s_1 and s_2 if s_1 is closer to s_3 than s_2 is, or over-reserve capacity if s_1 is further from s_3. We optionally apply a pessimistic grouping strategy which reduces the problem size (by reducing the total number of OD pairs), but given the potential error in under-use or overuse of some connections, we only consider low magnitude OD pairs and always assess solutions found using all OD pairs.

With reference to Figure 1, the S-tog network has a central region and several "fingers". In the presented plan, on each finger, one line serves every station and the other line serves only some stations, but this is not necessarily true for any solution plan. However it is in the fingers where grouping is most relevant; for example the final few stations on every finger may not be skipped by any line, and are therefore visited identically by every line and may be grouped.

7 Instance size

The OD data is not necessarily well-suited to all applications in that the hourly demand between many station pairs is non-integral (being the output of a demand model). Table 1 shows the total passenger demand for different rounding strategies of the individual OD demands between pairs of stations.

Table 1: Total OD demand for different roundings of individual OD demands

method	passenger number
rounded up	+6.52%
rounded down	-6.53%
rounded	-0.02%

Always rounding down results in a loss of over 6% of passengers, while rounding up adds more than an additional 6% passengers in total. Some solutions (including real solutions) are very tight in capacity in certain sections, and the additional passengers in the round-up scenario can make such solutions infeasible. Rounding down has a similar problem, in that solutions may be found which have very limited capacity and insufficient capacity for the true total number of passengers. We instead round all demands to the closest integer number of passengers which results in a total passenger demand very close to our raw input data, and is less likely to have such capacity problems.

Now that all components of the problem are defined, we can more explicitly define the standard S-tog instance size. Table 2 shows the size of various sets defined earlier.

Table 2: Sizes of problem instance parameters, and some derived values.

parameter	explanation	size
$\lvert\mathcal{L}\rvert$	Number of lines	174
$\sum_{l\in\mathcal{L}} \lvert\mathcal{F}_l\rvert$	Number of line-frequency decisions	350
$\lvert\mathcal{I}\rvert$	Number of marked incompatible lines	258
$\lvert\mathcal{C}\rvert$	Number of operational requirements	189
$\lvert\mathcal{OD}\rvert$	Number of (o, d) pairs	4645

Figure 5 shows the cumulative number of passengers for (o, d) pairs, when sorted by the number of passengers demanding a route for each (o, d). A small proportion of the (o, d) pairs account for the majority of the passengers. A potential simplification may then be to simply ignore some proportion of (o, d) pairs with small demands; however experiments solving reduced problems and then assessing solution quality considering all ODs gave poor results, as those low-demand ODs cover a diverse range of station pairings that are given insufficient consideration.

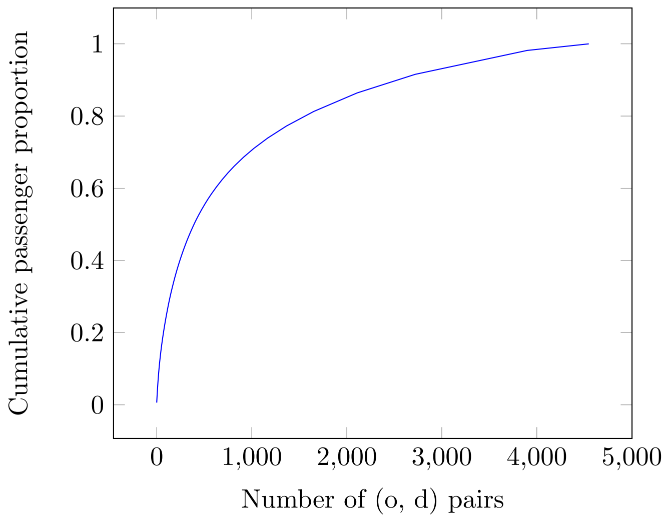

Figure 5: The cumulative proportion of passengers when considering (o, d) pairs sorted in decreasing order by the size of their passenger demand.

8 Method

Primarily, we use the formulation as described, with a constraint on the operator cost which we vary to find different results, and the total passenger time estimate as the objective function. Solving the resulting mixed integer program (MIP) directly can often provide good quality solutions, but we propose two additional methods for finding solutions.

In the following, we will refer to two real solutions: R1 and R2. These are both real historic line plans and frequencies as operated by DSB.

All testing is implemented with Gurobi 5.6 as the MIP solver on a machine with 8 GB of memory, and a four-core 2.5 GHz i7 processor.

8.1 Refining a given solution with a limited pool

From a given solution, we can create a limited pool of lines to use as the input to the model, which may or may not contain all the lines of the given solution and possibly also has the original solution as the optimal solution. The advantage of using a limited pool is that if the pool is sufficiently small, the model is solvable to optimality in reasonable time.

Suppose a solution is given, and let $\mathcal{S}$ be the set of lines present in the solution, with unspecified frequency. A simple restricted instance is to solve the model with only the lines $\mathcal{S}$, but with the definition of $\mathcal{F}_l$ for each line l in $\mathcal{S}$ unchanged. This is then a very small line pool (with generally one or two frequencies per line), but there is guaranteed to be at least one feasible solution present and likely to be others. We see that solving this limited model for any given solution can quickly find very similar but better solutions, especially in the case of real past line plans which were generally not planned with the same objective we use, and likely respect additional requirements. The similarity of any solution found to a real past solution is potentially useful, avoiding solutions that are significantly different to a plan that is not only feasible but known to operate in practice, which we can not in general guarantee.

To determine a wider but limited line pool, consider a set of lines from the entire line pool that are similar to a given line; let $N(l)$ be a set of "neighbouring" lines to line l which only differ in some small way to l. Now, we may use the following as a limited line pool:

$$\bigcup_{l \in \mathcal{S}} N(l)$$

We assume that $l \in N(l)$, and therefore $\mathcal{S}$ is a subset of this limited line pool.

The definition of $N(l)$ has a large effect on the problem size and solution quality. For example if $N(l) = \{l\}$ for all lines in $\mathcal{S}$, then this limited line pool is the same as simply taking the given solution lines. Alternatively, if $N(l)$ is very large then the resulting problem may have every line in the entire pool $\mathcal{L}$, and the line pool would not be "limited" at all.

Table 3 shows the problem sizes if we apply these two options, either taking lines but unspecified frequencies, or expanding with neighbouring lines, to the two real line plans R1 and R2. Here we indicate the solution with fixed frequencies as R1, the problem with the lines of R1 but open frequencies as R1$^+$, and the problem with all neighbouring lines to R1 as R1* (and the equivalent for R2). For each, we report the number of considered line-frequency combinations, and for the expanded problems report the solve time and the percentage improvement in the moving time, train switching time, and line cost. Note that the line cost is not considered in the objective, and as expected it increases, while we see modest improvements in the components of the total travel time.

Table 3: Solve times and objective improvements for different limited line pools. Cost improvements are the improvement in total moving time and switching time, and the improvement in operator line cost (which in fact becomes worse as operator cost may increase if it below the budget constraint).

	Lines			Cost improvements		
Problem	$\lvert\mathcal{L}\rvert$	$\sum_{l\in\mathcal{L}} \lvert\mathcal{F}_l\rvert$	Solve time (s)	Moving	Switching	Line
R1	9	9	-	-	-	-
R1$^+$	9	19	1	0.1%	9.7%	-3.3%
R1*	29	59	395	0.1%	9.7%	-3.3%
R2	8	8	-	-	-	-
R2$^+$	8	17	1	0.2%	5.5%	-2.4%
R2*	25	52	175	1.5%	4.5%	-5.3%

Solve times for the R1* and R2* instances are much greater than the R1$^+$ and R2$^+$ instances, and in the case of R1* no improvement is seen over R1$^+$. However, we can see that we can relatively quickly find line plans that "neighbour" a given plan, and here we can improve over these real line plans with very similar line plans (in the case of R1$^+$ and R2$^+$ the found solutions modify only the line frequencies). We can see here that there is more scope for improvement in switching time than travel time, given these reduced problems. Note, however, that here we report the percentage improvement for each individually, but the magnitude of those changes is very different, and in our problem a small relative improvement in travel time can be more significant than a large relative improvement in switching time.

When we solved both R1$^+$ and R2$^+$, the moving time improved even though we had exactly the same lines and can alter only frequencies. This may seem impossible, as the travel time between any pair of stations is unchanged. However the reason that we see small improvement is that either some passengers did not take their quickest (moving time) route due to a costly low frequency connection, or because some passengers could not take their fastest (moving time) route due to a lack of capacity, but with higher frequency (and therefore capacity) can now take that route.

8.2 LP heuristic

We propose the following as a simple heuristic for finding solutions. Initially, we solve the LP relaxation of the model and consider exactly those y_{lf} variables that have non-zero value. Then, we restrict the problem to only those lines present but find the optimal integer solution with the restricted problem. We compare the value of the optimal integer solution to the initial lower bound to the LP that we had, and, if we wish, we can re-introduce the missing line-frequency decision variables and allow the solver to tighten that lower bound and potentially discover better integer solutions. The advantage we see is that it is much faster to solve to optimality when there is a restricted pool of lines, and so we can relatively quickly find good solutions. In fact, in some experiments the solutions are optimal or near optimal. We hope that the smaller resulting problems have acceptable solve times but still have solutions of good quality.

Also, as a possibility, we can expand the lines in the LP solution using the ideas from Section 8.1.

We compare four different formulations, summarized in Table 4. The formulations differ in the presence of the additional constraints (14), and in whether or not the grouping of ODs from Section 6.3 is applied.

Table 4: Four different formulations, differing in the presence of additional constraints and their grouping of OD pairs.

	Without cons. (14)	With cons. (14)
No grouping	M1	M3
Grouping	M2	M4

We try solving the problem with the LP heuristic for each of the four methods. The solve times are summarized in Table 5, referring to firstly the solve time for the LP, and then the additional solve time to reach (potentially) the optimal solution. However it can be seen that only formulation M4 can solve both the LP and the subsequent IP to optimality in reasonable time. The others can all provide the LP solution but none can prove optimality for a solution in reasonable time. We allowed 5000 seconds for attempting to solve the resulting reduced IP for each model. After this time, neither M1 nor M2 both had an incumbent solution, whereas M3 had nearly found and proven an optimal solution. In fact, the line-frequencies solution provided by M3 was exactly the same as that provided by M4.

Table 5: LP and IP solve times for different formulations for an LP non-zero heuristic, where the IP is solved only considering non-zero variables in the LP solution. Termination with no incumbent marked 1; termination with a 0.8% gap marked 2.

	Times (s)	
Formulation	LP	IP
M1	325	5000^1
M2	42	5000^1
M3	4728	5000^2
M4	178	1086

We see that M1 and M2 do not solve to optimality or even find any feasible solutions in reasonable time. However, the comparison is potentially misleading because each is solved using the lines and frequencies found to be non-zero at LP optimality, and as we might expect (due to M1 and M2 lacking the additional constraints (14)) their LP solutions potentially have more non-zeros than the LP solutions for M1 and M2.

Consider Table 6 where we show the number of line-frequency combinations in the LP solutions to M1 and M3, and then consider expanding those as we did with integer solutions in Section 8.1.

From the table, it can be seen that the additional constraints (14) reduce the number of non-zero variables in an LP solution significantly. Also, in contrast

Table 6: The problem size in line-frequency combinations given by taking non-zero elements of an LP, or expanding that with additional frequencies (denoted $^+$), or with neighbouring lines (denoted *)

Problem	$\sum_{l\in\mathcal{L}} \lvert\mathcal{F}_l\rvert$
M1 LP	172
M1 LP$^+$	324
M1 LP*	350
M3 LP	59
M3 LP$^+$	62
M3 LP*	156

to the LP solution without constraints (14), the solution with the constraints features the majority of lines at every valid frequency (due to there being only an increase from 59 to 62 line-frequency variables when the missing frequencies are included). Without the constraints, however, lines tend to occur at only a single frequency but a far greater variety of lines is present. For M1 the expansions of the problem are unlikely to give the benefit we would like; rather than resulting in a still small problem, it is expanded to almost every line, and so we would gain little over attempting to solve the entire problem.

As can be seen from the table, the LP solution when using M1 has many more non-zero elements than the LP solution of M3. The same relative difference occurs comparing M2 and M4. As the heuristic method was to then take only those non-zero elements, it is perhaps not surprising that it was difficult to solve M1 with the 172 line-frequency combinations to IP optimality, as it was to solve M4 to IP optimality with its 59 line-frequency decisions. However, as a further experiment, we instead solved M2 to LP optimality, discarded the non-zero line-frequency elements of the problem not in the LP solution, and then added constraints (14). In this case, with a 5000 second time limit, we were able to solve the model to within 1.3% of optimal. This solution was in fact 2.8% *worse* than the solution found with the 59 non-zero line-frequencies of the M3 LP solution. This reveals a weakness of the LP heuristic method, in that it may be possible that there are no good integer solutions given the restricted problem, and possibly no feasible integer solutions at all.

Finally, we simply attempted to solved M4 as a MIP with all lines and frequencies, which, given 5000 seconds found a solution 1.7% worse than the solution provided by M4 with the LP heuristic. Allowing significantly more time, the solver can show that the M4 LP heuristic solution is within 1% of optimal, and it finds the exact same solution itself, but not better solutions.

9 Results

Though we assess line plans with total passenger travel time and by line cost to the operator, there are other metrics we may use to distinguish between different line plans. We note that to be implemented in practice a real line plan must meet other requirements we have not captured, such as facilitating good timetables. We see that with our formulation of line plan requirements, examples of real line plans are not optimal for either passenger travel time or lines operator cost, or for any weighted sum of those measures. However, we might expect that a usable line plan is one which has an appropriate trade-off between those two measures and falls within certain bounds for many other metrics.

Here, we consider the following metrics for line plans described in Table 7.

Table 7: The metrics used for line plan comparisons

Metric	Description
Line cost	The estimated cost to the operator, per hour, for running the line plan.
Hourly turn time	The minimum time per hour spent by trains dwelling/turning in end stations.
Hourly run time	The total moving time of trains in the line plan, per hour.
Hourly skipped stations	The number of stops that see a train pass by but not stop, per hour, in the line plan.
Unit requirement	The minimum number of units that would be required to operate the line plan (calculated using the minimum circulation time for a unit operating on the line).
Mean station visits	The average number of trains that visit a station every hour (or the total number of hourly station visits in the line plan divided by the number of stations).
Low-service stations	The number of stations which see no more than some threshold number of trains per hour. Here we use a threshold of 6.
Direct trips	The number of direct connection trips possible in the network, or the number of pairs of stations which feature together on at least one line.
Passenger cost	The total cost to the passenger, where cost is travel time with switching estimate and switching penalty.
Average travel time	The average per-person traveling component of the passenger cost, excluding switching.
Average switch cost	The average per-passenger switching cost consisting of the switching estimate and the switching fixed cost penalty.
Average switching wait	The average per-passenger time spent switching trains in the line plan.
Average switch penalty	The average per-passenger fixed switching penalty in the line plan.
Total transfers	The number of passengers who must transfer trains in the line plan.

To explore the range of values we might see for these metrics, we generate a set of solutions that we assess as being “interesting” due to either their acceptable trade-off between the costs to the operator or the passenger, or for being particularly good for some other measure. These are primarily generated iteratively, by solving the problem with different operator cost constraints and with a passenger cost objective, and storing incumbent solutions found by the solver. Good solutions were also refined as described in Section 8.1. For some solutions, also tighten some operational limit constraints (Constraints (5)) to find solutions that are more conservative but might be more likely to be operated by DSB as true line plans. Figure 6 shows every solution we consider, plotted by their cost to the operator and to the passenger. The majority of the considered solutions sit close to the frontier of solutions that are optimal for some weighted sum of these two measures, while a small number appear to be poor for both measures but are optimal or close to optimal for some other measure. These unusual solutions were found with a primary objective maximizing the number of direction travellers, as in Section 5.3. A real S-tog solution, R1, is marked with an open circle, showing that it is neither optimal for operator cost nor passenger cost, and there are solutions that are better for *both* measures. In reality, there are additional concerns we do not model that may mean real solutions are more favourable, or extra requirements that may invalidate the "better" solutions.

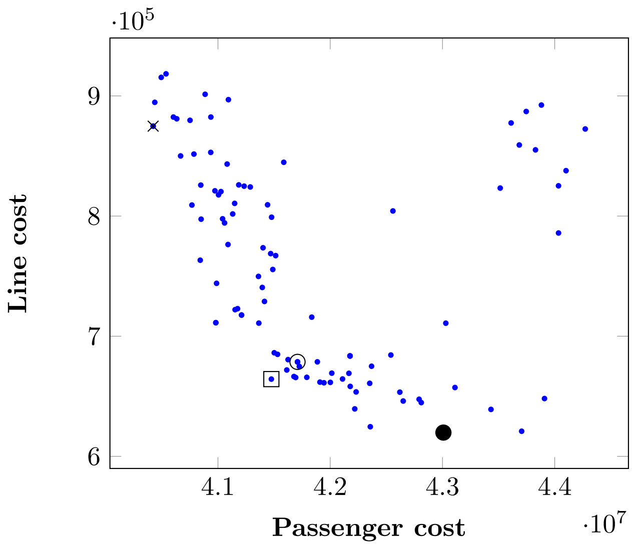

Figure 6: The passenger cost and operator cost for the considered set of noteworthy solutions. The open circle indicates real line plan R1.

On Figure 6, the optimal solutions considering only operator cost or only passenger cost can be seen. Due to the competing nature of the two measures, the solution providing the best value for one measure is poor for the other, though the best solution for the operator is not the worst solution for the passenger, and vice versa. Let us refer to these extreme solutions as SO, the operator-optimal solution; and SP, the passenger-optimal solution. Finally, we will refer to one of the solutions that dominates R1 for both **line cost** and **passenger cost** as SM (in fact four solutions strictly dominate R1 in this solution set).

Table 8: The minimum, maximum, and mean values for each metric, and the actual solution values for every metric for the four solutions: R1, SO, SP, and SM.

Indicator	Minimum	Maximum	Mean	R1	SO	SP	SM
Line cost	6.198e+05	9.184e+05	7.522e+05	6.786e+05	6.198e+05	8.748e+05	6.643e+05
Hourly turn time	3.672e+04	4.752e+04	3.869e+04	3.96e+04	3.924e+04	3.672e+04	3.816e+04
Hourly run time	2.454e+05	3.231e+05	2.71e+05	2.48e+05	2.454e+05	2.983e+05	2.563e+05
Hourly skipped stations	0	276	189.9	186	192	258	204
Unit requirement	82	108	89.82	84	82	97	85
Low-service stations	0	38	26.26	37	38	22	34
Mean station visits	33.71	53.36	41.85	37.29	36.86	44.67	41.74
Direct trips	2832	4482	3516	3132	3096	3752	3506
Passenger cost	4.042e+07	4.427e+07	4.184e+07	4.171e+07	4.301e+07	4.042e+07	4.148e+07
Average travel time	1027	1130	1050	1042	1048	1033	1038
Average switch cost	99.83	173.1	125.3	130.2	160.8	103.3	127.2
Average switching wait	45.37	107.7	68.23	66.92	99.67	50.78	73.62
Average switch penalty	42.13	72.16	57.09	63.33	61.1	52.53	53.63
Total transfers	4165	7133	5643	6260	6040	5193	5301

Table 8 shows the minimum, maximum, and mean values for the different measures for all solutions in the set we discuss here. The table also shows the metric values for four solutions: true solution R1, solution SO that is best for **line cost**, passenger-optimal solution SP, and the fourth SM that is better than R1 for both **line cost** and **passenger cost**.

Figure 7 shows normalized boxplots for all of the metrics, with the values for four noteworthy solutions marked on each. The real line plan R1 is marked on Figure 6 with an open circle, and we can see it is below average for **line cost** and above average for **passenger cost**. Noteworthy observations are that it has a rather low unit requirement (which is a component of the total **line cost**), but has a particularly high number of total transfering passengers compared with most other solutions, and an above average passenger (moving) travel time. In contrast, the solution SO (optimal for **line cost**) also has the lowest unit requirement of all presented solutions, but has the highest number of **low-service stations**. It also has a higher total cost to the passenger compared to solution R1 although it has fewer overall transfer passengers, and surprisingly it provides a very similar number of **direct trips** as SO. Solution SP, the passenger-optimal solution, has, perhaps unexpectedly, a very large number of **hourly skipped stations**, which corresponds to many passengers saving travel time. However this is not completely at the expense of those passengers at those skipped stations being required to transfer as the solution has a below-average (but not exceptionally low) number of transferring passengers. Finally, we mark a fourth solution which is better than R1 for both **line cost** and **passenger cost** (with an open square). We observe that it is similar to solution R1 for some metrics but surprisingly different for others; for example, it provides more **direct trips** and has fewer transferring passengers than R1, but is greater for metric **average switching wait**. The interpretation of that is that while there are fewer overall passengers transferring, those that do transfer are transferring to lines with greater headway and therefore are estimated to wait longer. The **hourly skipped stations** is indicative of a lack of homogeneity in trains; with zero **hourly skipped stations** all trans stop at every station while if this value is high, many stations are bypassed. Bypassed stations must still be served, and so there must be parallel lines running on the same infrastructure stopping at the stations, therefore stopping more and running slower. This could have implications for the robustness of the subsequent timetable; some authors [Vromans et al., 2006, for example] equate high robustness with high homogeneity in running speed. We can see that the real line plan R1 has lower **hourly skipped stations** that the other marked plans, but substantially lower values are present in other plans. The measure **hourly turn time** could also be relevant for a robustness perspective, as the measure indicates the total minimum hourly utilization of turning facilities at stations, either on-platform turning or using dedicated turning track. A higher value indicates greater utilization of turning capacity and likely therefore less buffer between turning events, and subsequently less robustness for absorbing delay.

It may be expected that a solution with many **direct trips** would also be a solution with low **passenger cost**, due to the limiting of the line-switching cost in the total passenger cost. However, that does not take into account the fact that solutions with more **direct trips** tend to have fewer missed stops, resulting in longer travel times for some passengers, and that not all line switching is equally penalised as we estimate the wait time by the line frequency. Figure 8 shows

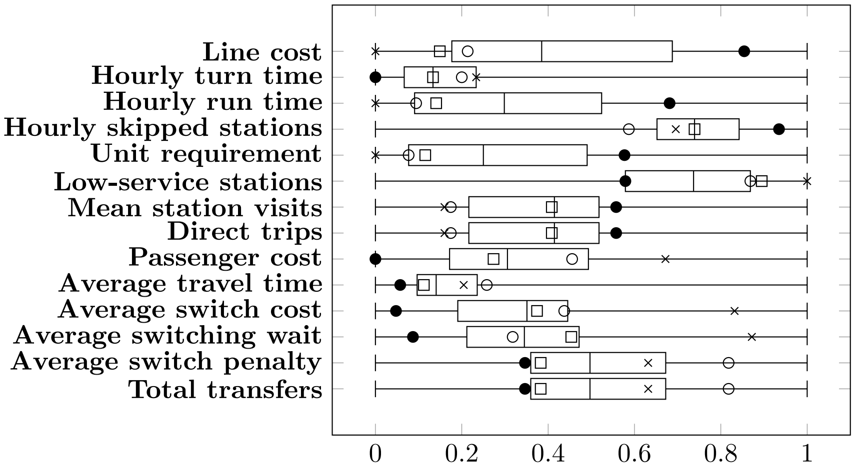

Figure 7: Normalized distributions of values for the performance indicators. The real line plan R1 is marked in each row with an open circle. The asterisk and the closed circle indicators indicate the optimal **line cost** and **passenger cost** solutions, respectively (SO and SP). The open square indicates a solution which dominates the real solution for those two measures.

passenger cost and the number of **direct trips** for the range of solutions. Surprisingly, there is a set of solutions that have many direct trips but are also high in passenger cost. Excluding those, there may be some negative correlation between **passenger cost** and the number of **direct trips**, but this is not strong. The real solution R1 is marked, and is not particularly extraordinary for either measure. We see from the figure that simply maximizing the number of direct travellers, a measure used by some other work in the field (for example Bussieck et al. [1997]), is not appropriate for this particular problem, as those solutions with the highest number of direct travellers are particularly bad for the passenger.

We may consider the proportion of the total cost to the passenger that is attributable to travel time, and to switching lines. Further, we can consider the components of the cost of switching lines. That is, we consider as a fraction of the per-person passenger cost, the cost attributed to moving (**average travel time**) and the cost attributed to waiting to transfer (**average switch cost**). Of the waiting time cost, there is the estimate of the waiting time itself (**average switching wait**) and fixed penalty (**average switch penalty**). Table 9 shows a summary of the percentage of passenger cost that can be attributed to the different constitutive components. For these solutions the vast majority of the total passenger cost is attributable to actual moving time in trains. All passengers spend time travelling while only some must switch trains, and for many journeys with a switch, the actual travel time is greater than the switching time estimate and penalty combined, so it is unsurprising that travelling time is greater overall. We see that, for all considered solutions here, at least 85.81% of passenger cost is attributable to actual travelling in a train, while up to 14.09% is due to the cost of switching trains.

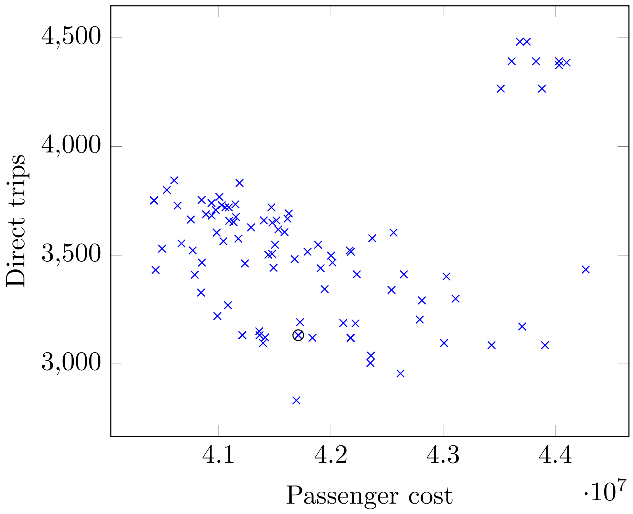

Figure 8: The total passenger cost and the number of direct trips for the set of solutions. The real line plan R1 is marked with an open circle.

Table 9: Summary of the percentage of total passenger cost attributable to passenger travel time, switching cost, and the components of switching cost.

Component	Minimum	Mean	Maximum
Average travel time	85.91	89.33	91.88
Average switch cost	8.12	10.65	14.09
Average switching wait	3.92	5.79	8.77
Average switch penalty	3.40	4.86	6.16

10 Solving only lines model

Solutions are a valid line plan if they satisfy constraints (2)–(6), operationally and contractually, though these do not guarantee sufficient capacity for all passengers. However, the minimum-visit requirements at many stations do mean that most line plans satisfying (2)–(6) have close to sufficient capacity, or sufficient capacity, for passengers. To determine the passenger-feasibility of a solution, and the cost to the passenger, we can construct the graph described in Section 4.1 for only those lines in the given solution, and solve just the passenger flow problem. Alternatively, as in Section 8.1, we may take the lines but not their frequencies from a given solution, or construct a neighbourhood of lines, and then solve (1)–(6), and (7)–(9) to, with greater likelihood, find a feasible solution for passengers similar to the line plan given. It is possible to forbid specific solutions, either by specific line-frequency (constraint (15)), or by just line at any frequency (constraint (16)), as described in Section 6.2. A potential method for finding solutions is to iteratively solve the MIP (2)–(6) to optimality with some objective, apply a constraint ((15) or (16)), and repeat the process. The found solutions can be assessed in terms of passenger quality by solving the passenger flow problem, if frequencies are known, or a limited line model if the frequencies are still to be determined. It is possible that with a well-chosen objective for the reduced problem, good solutions can be found. As the limited problem of finding solutions is small (even with hundreds or thousands of additional solution-forbidding constraints), many solutions can be discovered quickly. The passenger flow problem is also relatively fast to solve for a fixed solution and so such solutions can be quickly assessed for quality and feasibility for passengers.

Of the metrics introduced in Section 9, the following do not depend on passenger flows:

- **Line cost**
- **Hourly turn time**
- **Hourly run time**
- **Hourly skipped stations**
- **Unit requirement**
- **Low-service stations**
- **Mean station visits**
- **Direct trips**

As we describe in Section 5.2 and Section 5.3, we can use other objective functions (**line cost** or **direct trips**, respectively). In fact we could use all of the metrics here, some more readily than others, as an objective function. However we observe that none is a good substitute for the total passenger cost objective we primarily use, and none gives solutions that are either particularly good for the passenger or that are similar to real DSB solutions. As already seen, the real DSB solutions are not extraordinary for any of the metrics we define, and so using any one metric as an objective function does not give similar solutions. Furthermore, although we can quickly find, qualify and forbid solutions and therefore assess a large number in reasonable time, we also see that there is a very large number of possible solutions to the problem. Such an approach, with the wrong objective, tends to find many solutions that are not of particular interest. An exception is when using the **line cost** objective; instead we see many infeasible solutions (for the ignored passengers), but the few that are feasible are interesting for having minimal **line cost**. However we note that when ignoring passengers, the first feasible solution is number 1305 discovered (i.e. 1304 other solutions were discovered with better **line cost** that were not feasible for the passenger). Subsequent feasible solutions occurred as solution numbers 2018, 2251, 2155 and 2274 in the first 2500 solutions discovered.

11 Conclusions

Here we use an arc-flow formulation to attempt to solve a line planning problem for the S-tog network, focusing on passengers. The integer programme we formulate can generally not be solved to optimality for our instances, but we can find relatively good solutions in reasonable time with an LP based heuristic method. Given more time the full formulation itself can also find good quality solutions but not generally prove optimality (without excessive additional time). We also show we can find good solutions quickly when we restrict ourselves to lines that are similar to currently-operated lines, and this is perhaps a natural restriction as it is unlikely that the operator would change all lines at once.

The passenger focus means that the lines are of good quality for the passenger and tend to be at the upper limit of whatever cost limit we enforce. By reducing the cost limit to be lower than real operated plans, we can show that there are plans which are both better for the average passenger and cheaper in line cost (though our line cost does not necessarily reflect all components of the true operating cost or other important measures). Considering some other measures, but not including them as constraints or objectives for the optimization, we can compare solutions by more than just one or two objectives, and see that there are significant differences between apparently similar solutions

We show that for this problem, the arc-flow model, though large, can be directly applied and solved to find solutions of reasonable quality, and we show a simple LP based heuristic approach to find good solutions more quickly. In our experiments we have found that the model is also applicable to the same problem but with many more frequencies per line, without becoming unsolvable. However, given that the operational requirements are created considering the given frequency options, there are fewer interesting solutions with additional frequencies.

One limitation of the model is that, while we try to minimize switching time, we can only estimate this time as we have no timetable. In fact, the lines constraints (constraints (1)–(6)) do not capture everything necessary to ensure that a timetable can be created for the line plan at all; it may be that our proposed line plans are infeasible. However, assuming that a valid timetable does exist, we can still only estimate the wait time.

Another limitation is that we penalise the cost of switching to one specific line. However, for many trips, when boarding a subsequent line on a trip, a passenger can have several similar options in some line plans. For example, a passenger may begin on line l_1 and exit at some station to wait for a train to their destination, and there are two lines l_2 and l_3 stopping at both their intermediate and destination stations, and a real passenger would likely board the first of those to arrive (even if the first train provided longer travel time). Therefore if the two lines operate at a low frequency, our long estimated wait time is pessimistic because the combined frequency of the lines is not low.

Finally, we restrict ourselves to a predetermined line pool which is not an exhaustive set of every feasible line. However the more limited number of lines ensures a reasonably sized problem, and also provides more certainty that every line in the pool is feasible (alone) because they may all be explicitly checked by an expert.

12 Future work and extensions

As we have said, we restrict ourselves to a predetermined line pool of 174 lines for the S-tog problem. For some different problems not expanded upon here, but, for example, finding a night-time line plan, we use a different pool of lines due to the different rules defining a feasible line and line plan. All such limited pools can be viewed as being subsets of the set of any possible line, which, due to the possibility of arbitrary stopping patterns, is very large. However, we have experimented with expanding our limited pool by considering new lines not in the original pool but with some similarity to pool lines. By fixing some of the lines in a solution, but for the non-fixed part introducing a large variety of new, out-of-pool lines, we can find line plans with some different lines relatively quickly. By performing such moves within a heuristic framework, there seems to be potential to explore all possible lines in a reasonable way that is not possible with the MIP formulation we present.

The line plan solutions we find may possibly, but not necessarily, be operated in practice by DSB. Despite meeting all operational and contractual requirements, and not containing any explicitly forbidden line pairs which cause problems for creating a timetable, it can still be the case that it is not possible to create a feasible timetable for a line plan. Even if there are feasible timetables for a line plan, there may be no *good* timetables for the line plan; the measures used to compare timetables are not necessarily the same as those we estimate for the line plans (operator cost and passenger cost) and may include other measures such as minimal headways. On the other hand, our estimated operator cost and estimated passenger waiting times can be more precisely assessed when the timetable is determined. One direction for future work is more closely integrating the creation of a line plan with the subsequent timetable problem that takes a line plan as input. This would avoid the risk of finding infeasible line plans (from the point of view of timetable creation), and could allow more precise operator cost and waiting time estimates. Looking further in the planning process, the timetable itself may not facilitate the creation of good or even feasible rolling stock plans. Closer integration of all planning stages of rail is obviously desirable to achieve overall optimality of plans, but the complexities of each problem alone prove to be a challenge, and so complete integration of every problem stage is unlikely.

References

Ralf Borndörfer, Martin Grötschel, and Marc E Pfetsch. A column-generation approach to line planning in public transport. *Transportation Science*, 41(1): 123–132, 2007.

Michael R Bussieck, Peter Kreuzer, and Uwe T Zimmermann. Optimal lines for railway systems. *European Journal of Operational Research*, 96(1):54–63, 1997.

Marc Goerigk, Michael Schachtebeck, and Anita Schöbel. Evaluating line concepts using travel times and robustness. *Public Transport*, 5(3):267–284, 2013.

Jan-Willem Goossens, Stan Van Hoesel, and Leo Kroon. A branch-and-cut approach for solving railway line-planning problems. *Transportation Science*, 38(3):379–393, 2004.

Jan-Willem Goossens, Stan van Hoesel, and Leo Kroon. On solving multi-type railway line planning problems. *European Journal of Operational Research*, 168(2):403–424, 2006.

Karl Nachtigall and Karl Jerosch. Simultaneous network line planning and traffic assignment. In *ATMOS*, page 105, 2008.

Natalia J Rezanova. Line planning optimization at DSB. In *13th Conference on Advanced Systems in Public Transport*, 2015.

Anita Schöbel. Line planning in public transportation: models and methods. *OR spectrum*, 34(3):491–510, 2012.

Anita Schöbel and Susanne Scholl. Line planning with minimal traveling time. In *ATMOS 2005-5th Workshop on Algorithmic Methods and Models for Optimization of Railways*. Internationales Begegnungs-und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, 2006.

Michiel JCM Vromans, Rommert Dekker, and Leo G Kroon. Reliability and heterogeneity of railway services. *European Journal of Operational Research*, 172(2):647–665, 2006.