**CASPT 2015**

# Line Planning Optimization at DSB

**Natalia J. Rezanova**

**Abstract** Solving the line planning problem is one of the fundamental steps in strategic planning of a railway operator. We present a cost-oriented line planning optimization tool developed at a Danish railway operator DSB for strategic planning and analysis purposes. Two line planning optimization models are presented, a cost-oriented integer programming model and a passenger-oriented network flow model, which are combined in order to find line planning solutions that minimize the cost of the line plan while keeping the passenger travelling time low. The line planning optimization tool was successfully used to generate and evaluate different line plans for the year 2016 DSB S-bane railway schedule.

**Keywords** Danish railway · DSB · Integer programming · Line planning · Public transportation

## 1 Introduction

Given a railway infrastructure, a passenger origin-destination travel demand, and a list of requirements to the level of service, the *line planning* problem of a railway operator is to find a set of train lines with a ceratin line frequency, which satisfy all requirements. A train line is a return trip between two terminal stations with a certain halting pattern. The *capacity* of a line is determined by its frequency and the rolling stock capacity assigned to the line. A line with a high frequency and a high rolling stock capacity can transport more passengers than the same line with a lower frequency and/or lower train capacity. The availability of the rolling stock determines the total train capacity that can be assigned to a line plan, while other oprational constraints, such as the channel availability at different segments of the railway infrastructure,

Natalia J. Rezanova
DSB, Operations, Longterm Planning, Telegade 2, 2630 Taastrup
Tel.: +45 24 68 25 65
E-mail: najr@dsb.dk

the minimum headways between trains, the lack of overtaking possibility, or the limit to the number of turning trains at certain terminal stations, set a limit to the frequency of the line.

Public service transportation operators have service agreements with the authorities about a minimum service level of operations. DSB is a railway operator responsible for the passenger traffic of all intercity trains as well as the majority of all regional trains in Denmark. DSB has a contract with the Danish Ministry of Transport, where the partners agree about a minimum level of service per station (number of stops), per track segment (number of trains per hour) and the minimum number of direct connections between certain cities. The service agreement ensures availability of train departures at all stations in Denmark, also those stations, which would otherwise not be serviced by operators because the cost of service is much higher than the ticket revenue. According to the service agreement (also called *traffic contract*), DSB recieves a reimbursement from the Danish government for providing the required level of service, even if it is not profitable otherwise. The service requirements according to the traffic contract are therefore an important part of the line plan.

Solving the line planning problem is a part of the longterm strategic planning process. Most railway operators try to keep their lines unchanged unless it is absolutely necessary. The entire Danish rail network will undergo major improvements during the next 15 years, including electrification, infrastructure expansions, and replacement of the signalling systems. The line plans would have to be adjusted accordingly. The line planning optimization tool was developed in DSB Longterm Planning department to be used for the year-to-year strategic planning whenever a change to the line plan is required, as well as for analysis connected to the service agreement negotiations between DSB and the Danish Ministry of Transport.

This paper presents optimization models implemented in DSB line planning tool, as well as an implementation example, showing how the tool was used to generate and evaluate the year 2016 S-bane line plan scenarios.

## 2 DSB line planning optimization tool

As the comprehensive review of OR models within the line planning, Schöbel (2012), points out, the line planning optimization problems fall into two categories: passenger-oriented models and cost-oriented models. Passenger-oriented models focus on either maximizing the number of direct travellers (Bussieck et al (1996)), minimizing riding time for the passengers (Pfetsch and Borndörfer (2006), Borndörfer et al (2007)), or minimizing both passenger riding time and line change time (Schöbel and Scholl (2006)) in a line plan. Cost-oriented models focus on minimizing operator's costs of a line plan: Claessens et al (1998), Goossens et al (2004), Goossens et al (2006).

We choose to adapt and expand one of the three formulation of the cost-oriented line planning model described in Goossens et al (2006). This integer programming model, called $IP_X$, makes sure that within a given period of time (typically one hour) the passenger demand is covered by a set of lines with different patterns and frequency. The passenger flow is distributed according to the known preferences, and each line is assigned a certain train capacity in order to make sure that the total passenger demand is satisfied. The objective function minimizes the operational costs calculated for the estimated train units necessary to cover the passenger demand with the given line frequency.

There are three main reasons for choosing the model suggested by Goossens et al (2006).

First, it allows considering multiple line types in the same optimization run, and the passenger demand is not split between the line types. DSB train lines have different halting patterns: fast and slow, regional and intercity. Regional networks, such as S-bane, are dense and heavily used by commuters who frequently change between fast and slow lines. Therefore, it is important to solve the line planning problem for regional and commuter networks *without* dividing the passengers into separate networks according to the so-called *system split* procedure, which allows to solve the line planning optimization problems independently for each train category, and is applicable for intercity trains.

Second, the chosen integer programming formulation of the model can easily be extended by adding a set of operational and service constraints necessary to obtain feasible line planning solutions for DSB. The definition of the train capacity used by Goossens et al (2006) is extended to incorporate different compositions of carriages, including combination of different fleet types. When extended this way, the model allows to set a limit to the estimated number of carriages used in the line plan or even to minimize the capacity in the plan.

Third, if the line planning optimization tool is to be used for what-if analysis in practice, the solutions must be obtained relatively quickly. The chosen model formulation has a better complexity than many other line planning models, and therefore allows to solve problems of medium network sizes to optimality relatively fast without using advanced integer programming techniques.

In order to compare the existing or manually generated line plans with the optimization solutions, an additional feature of the tool allows to validate a given solution built from a given line pool instead of generating line routes automatically. The line planning optimization tool is implemented in JAVA, and the integer programming model is solved using the MIP solver of MOSEK v.7 (www.mosek.com).

## 3 Modelling

The railway infrastructure network, also called a *track graph*, $G = (S, E)$ is described by a set of stations $S$ and a set of tracks $E$ between stations. Every station node has a *type* $t$ indexed from 1 to $t^{max}$. The station type represents the size of the station and determins the allowed halting patterns of the train

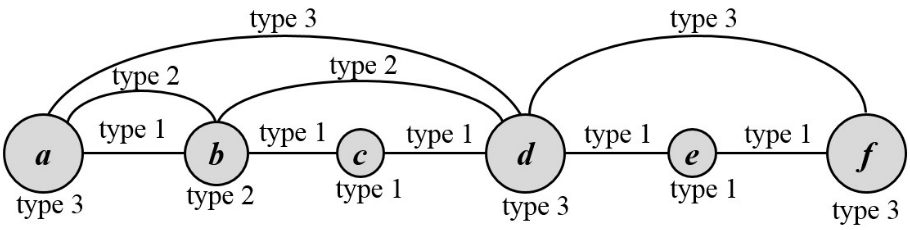


**Fig. 1** Example of a type graph $G^T = (S, E^T)$.

lines at the station. The station type $t = 1$ is the lowest type, used for the stations in small towns and villages along the railway network, while types $t > 1$ represent larger towns and cities.

Based on the railway network $G = (S, E)$ and the types of the stations in $S$, we generate a *type graph* $G^T = (S, E^T)$, as defined in Goossens et al (2006). Every edge in the type graph has at type $t$, indexed from 1 to $t^{max}$. The set of edges $E_1^T$ contains all edges in $E$ and has the type $t = 1$. Every edge of type $t > 1$ in the type graph $G^T = (S, E^T)$ connects two closest stations of the same or a higher type. A station $w_t$ is one of the *closest stations of the type* to a station $v_t$, if there exists a simple shortest path between $w_t$ and $v_t$ in the underlying railway infrastructure network $G = (S, E)$, which does not contain any other stations of type that is higer og equals to $t$.

In a type graph example shown on Figure 1, a type 3 station $d$ is connected to the closest type 1 stations $c$ and $e$ by type 1 edges, to the closest type 2 station $b$ by a type 2 edge, and to the closest type 3 stations $a$ and $f$ by type 3 edges.

Every type edge $e \in E^T$ can be expressed through a simple path $P(e)$ in the underlying network $G = (S, E)$. In the above example, $P(\{a, d\}) = \{\{a, b\}, \{b, c\}, \{c, d\}\}$.

Consider two type edges, $e_1$ and $e_2$. We say that a type edge $e_1$ *covers* a type edge $e_2$ if $P(e_2) \subseteq P(e_1)$. For example, edge $\{b, d\}$ covers edges $\{b, c\}$ and $\{c, d\}$. The covering edge always has a higer type than the covered edge.

We define $\Phi_t(e)$ as a subset of edges in $E^T$, which contains all edges of type $t$ that *cover* the edge $e$. For $e = \{b, c\}$, $\Phi_2(\{b, c\}) = \{\{b, d\}\}$, while $\Phi_3(\{b, c\}) = \{\{a, d\}\}$.

Let $L$ be the pool of possible train lines. In the present formulation of line planning model CLPP (1) - (9), a line $l \in L$ represents a *return travel* between two terminal stations. The halting pattern of a line is represented by a path in the type graph $G^T = (S, E^T)$. The path of the line does not have to follow the edges of the same type, but can be a combination of different edge types, and therefore a combination of different halting patterns in different parts of the railway network. The halting pattern is, however, unchanged in both directions of the line.

Let $F^l \subset \{1, 2, ...\}$ be a set of possible frequencies for the line $l \in L$. A frequency $f \in F^l$ is defined for a certain period of time, for instance, one hour.

A set of rolling stock compositions $C^l$ contains different combinations of the rolling stock, which are possible to assign to a line $l \in L$. A composition $c \in C^l$ contains one or more carriages of one or more rolling stock types $r \in R$. Each composition $c \in C^l$ has therefore a known number of seats $\gamma_c$ and a known cost of driving, which depends on the rolling stock type and the number of carriages in the train.

The line planning optimization problem is to find a cost-optimal combination of lines, frequencies, and rolling stock compositions assigned to the lines, such that the passenger demand is met and all defined operational and service constraints are satisfied.

For every combination $(l, f, c) \in N$ we define a binary decision variable. We use $i \in N$ to refer to a parctucular combination of $(l_i, f_i, c_i) \in N$:

$$x_i = \begin{cases} 1 & \text{if the combination of line, frequency and composition} \\ & i = (l_i, f_i, c_i) \in N \text{ is in the optimal solution.} \\ 0 & \text{otherwise.} \end{cases}$$

A decision variable $x_i$ has two cost components: an operational cost $w_i^{cost}$ and a passenger-related cost $w_i^{pax}$.

For a given line circulation time and frequency, we calculate the minimum number of trains necessary to cover the line. The circulation time of a line is the total time in minutes that a train of line $l \in L$ needs to complete its return trip between two terminal stations, including the minimum turn time at the terminal stations and a halting time at all intermediate stations. The minimum number of trains necessary to cover the line within a given time period (let us say, 60 minutes) is calculated as $\lceil \frac{\text{circulation time} \times f}{60} \rceil$, where $f \in F^l$ is the frequency of the line. This calculation is beautifully illustrated in Goossens et al (2004).

Knowing the driving time of the line, the minimum number of trains necessary to cover the line, and a number of carriages of each rolling stock type in every train, we can use three following operational cost components when calculating $w_i^{cost}$ for every $x_i$:

1. The cost of having a specific line $l_i$ in the solution.
2. The cost of using one carriage of a rolling stock type $r \in R$ in the plan.
3. The cost of one *train minute* of driving the train, independently of the train composition.
4. The cost of one *rolling stock kilometer* of driving each carriage in the composition.

A passenger-related cost $w_i^{pax}$ measures the *unattractivenes* of the combination $(l_i, f_i, c_i) \in N$ from the passenger point of view. As a rule, train lines with a higher frequency and a larger train capacity are more attractive to the passengers. From an individual passenger's point of view, the train line is attractive when it it is fast and has as few stops along the way between the passenger's origin station O and the destination station D. We can therefore assume that for every passenger OD-relation, the shortest path between O and D in the type network represents the passenger's most attractive line route.

In order to measure the attractiveness of the line from the point of view of *all* passengers who would potentially use the train line, we introduce a directed travel network $G^{TA} = (S, A^T)$ based on the type graph $G^T = (S, E^T)$, where each edge $e \in E^T$ is replaced by two directed arcs, $\overrightarrow{a} \in A^T$ and $\overleftarrow{a} \in A^T$, running in opposite directions. Given an asymmetric origin-destination (OD) matrix of passenger relations, and assuming that every passenger's most attractive route in the network is the shortest path in $G^{TA} = (S, A^T)$, we can calculate the total passenger demand $H(a)$ on every arc $a \in A^T$ in the directed travel network $G^{TA} = (S, A^T)$.

Decision variables $x_i$ do not allow to change the assigned train capacity along different parts of the line: since the line represents a return trip, the same train capacity is assigned to the line in both directions. Due to this assumption, we calculate the passenger demand $H(e)$ of the type edge $e \in E^T$ as the largest demand of the two corresponding arcs $\overrightarrow{a} \in A^T$ and $\overleftarrow{a} \in A^T$: $H(e) = \max\{H(\overrightarrow{a}), H(\overleftarrow{a})\}$. This assumption makes it possible to reduce the number of decision variables in the model, while unfortunately introducing a risk of assigning too much train capacity to the line plan in cases where the difference between $H(\overrightarrow{a})$ and $H(\overleftarrow{a})$ is big for all edges in a parcitular line route, and where it would make sense to use a different capacity in one direction compared to the other direction of the line or change the capacity at different line segments. Examples of such differences in demand are the morning rush hours, where the passenger demand towards larger cities is larger than the demand in the opposite direction. In practice, a rolling stock capacity is driven by the passenger demand, and coupling and decoupling of carriages happens either at the terminal stations of the line or at larger depot stations along the line. The flexible train capacity assignment to the different segments of the line is currently in the pipeline for implementation in the line planning optimization tool.

The passenger-related cost $w_i^{pax}$ uses three user-defined parameters, which are able to punish the unattractiveness of the decision variable $x_i$ independently of the presence of other decision variables in the solution:

1. The frequency parameter, which is a negative number multiplied to the frequency $f_i \in F$ of the combination $i \in N$.
2. The train capacity parameter, which is a negative number multiplied to the number of seats $\gamma_{c_i}$ in the train composition $c_i \in C$.
3. The passenger demand parameter, which is a negative number multiplied to the largest edge demand $H(e)$ found on the path along the line $l_i \in L$.

Several alternative objectives can be achieved by altering the cost components $w_i^{cost}$ and $w_i^{pax}$. By setting the cost of using one carriage to the same positive number, while setting all other costs to zero, the objective function of the model is transformed into finding a line plan that minimizes the number of carriages assigned to the lines. A special case of this objective is to set the cost of using a carriage to the number of seats $\gamma_c$ in the train composition $c_i \in C^l$, while all other costs are set to zero. Then the objective function of the model is transformed into minimizing the total line plan capacity. It is

**Table 1** Notation and terminology used in CLPP

| | |
|---|---|
| $L$ | set of lines, indexed $l \in L$. |
| $F^l$ | set of line frequencies for a line $l \in L$, indexed $f \in F^l$. |
| $C^l$ | set of train compositions for a line $l \in L$, indexed $c \in C^l$. |
| $N$ | set of combinations of lines, frequencies and compositions, referred to by $i = (l_i, f_i, c_i) \in N$. |
| $l_i$ | line in the combination $i \in N$. |
| $f_i$ | frequency in the combination $i \in N$. |
| $c_i$ | train composition in the combination $i \in N$. |
| $\gamma_{c_i}$ | capacity of $c_i$ measured in the number of seats. |
| $S$ | set of stations, indexed $s \in S$. |
| $s_i$ | a station where the line $l_i$ in $i \in N$ halts or turns. |
| $E^T$ | set of edges in the type graph $G^T = (S, E^T)$, indexed $(e) \in E^T$. |
| $t$ | arc type, $t = 1, ..., t^{max}$. |
| $E_1^T$ | set of edges in $e \in E^T$ belonging to a type $t = 1$, also represents the edges in the track graph $G = (S, E)$. $E_1^T \subseteq E^T$. |
| $e_i$ | an edge in $E^T$, which is part of the line $l_i$ belonging to $i \in N$. |
| $\Phi_t(e)$ | set of edges of type $t$ which cover the edge $e$. |
| $H(e)$ | total passenger demand on edge $e \in E^T$. |
| $R$ | set of rolling stock types, indexed $r \in R$. |
| $r_i$ | rolling stock type that is part of the composition $c_i$ in $i \in N$. |
| $\alpha_{r_i}$ | number of carriages of the rolling stock type $r \in R$ in capacity $c_i$ in $i \in N$. |
| $B^r$ | number of carriages of type $r \in R$ in the fleet. |
| $V_e^{min}$ | minimum number of trains on a track $e \in E_1^T \subseteq E^T$. |
| $V_e^{max}$ | maximum number of trains on a track $e \in E_1^T \subseteq E^T$. |
| $V_s^{min}$ | minimum number of trains stopping at a station $s \in S$. |
| $V_s^{max}$ | maximum number of trains stopping at a station $s \in S$. |
| $\Psi(l)$ | set of lines that cannot coexist with the line $l \in L$, $\Psi(l) \subset L$. |
| $L^{\Psi}$ | set of lines that contains in at least one of the sets $\Psi(l)$. $L^{\Psi} \subset L$. |

important to mention that the line planning model is by no means a rolling stock optimization model, but these alternative objective functions are very useful when different line plan scenarios are considered.

The cost component of having a specific line $l_i$ in the solution is used in scenarios where we want to minimize the number of lines with specific characteristics. As an example, we can punish the lines starting and terminating at certain stations or running in specific network regions. When the cost of all lines is set to the same positive number, while all other costs are set to zero, we can minimize the number of lines in the solution. By minimizing the number of lines, we achieve a streamlined line plan, containing a small number of lines with the same pattern and frequency, which is easy to remember for the passengers.

By setting the operational cost component $w_i^{cost}$ to zero, while keeping a non-zero cost component $w_i^{pax}$, the objective function is transformed into minimizing the passenger unsatisfaction. Unfortunately, the model formulation does not allow to optimize the passenger flow directly, and it is not possible to minimize the number of line changes or the line change duration time for each passenger relation. Some passenger-oriented line planning models allow to do so, e.g. Schöbel and Scholl (2006), but the size of these models grows quite heavily with the network size and the number of passenger OD-relations.

The integer programming formulation of the implemented cost-optimization line planning problem (CLPP) is described below. Table 1 gives an overview of notation and terminology used in the model.

$$\text{(CLPP)} \qquad \min \sum_{i \in N} (w_i^{cost} + w_i^{pax}) x_i \tag{1}$$

$$\text{s.t.} \qquad \sum_{i \in N | l_i = l} x_i \leq 1 \quad \forall l \in L \tag{2}$$

$$\sum_{i \in N | e_i = e} \gamma_{c_i} f_i x_i \geq H(e) \quad \forall e \in E_1^T \subseteq E^T \tag{3}$$

$$\sum_{i \in N | e_i = e} \gamma_{c_i} f_i x_i + \sum_{t>1} \sum_{i \in N | e_i = e' \in \Phi_t(e)} \gamma_{c_i} f_i x_i \geq H(e) + \sum_{t>1} \sum_{e' \in \Phi_t(e)} H(e') \quad \forall e \in E_1^T \subseteq E^T \quad \forall \Phi_2(e), ..., \forall \Phi_{t^{max}}(e) \tag{4}$$

$$\sum_{i \in N | r_i = r} \alpha_{r_i} x_i \leq B^r \quad \forall r \in R \tag{5}$$

$$V_e^{min} \leq \sum_{i \in N | e_i = e} f_i x_i \leq V_e^{max} \quad \forall e \in E_1^T \subseteq E^T \tag{6}$$

$$V_s^{min} \leq \sum_{i \in N | s_i = s} f_i x_i \leq V_s^{max} \quad \forall s \in S \tag{7}$$

$$\sum_{i \in N | l_i = l} x_i + \sum_{i \in N | l_i = l' \in \Psi(l)} x_i \leq 1 \; \forall l \in L^{\Psi} \subset L, \forall \Psi(l) \subset L \tag{8}$$

$$x_i = \{1, 0\} \; \forall i = (l_i, f_i, c_i) \in N \tag{9}$$

The objective function (1) minimizes the operational and passenger-related costs as described above. Constraints (2) ensure that only one combination $i \in N$ with a specific line $l \in L$ is chosen in the plan. Constraints (3) - (4) are the so-called *capacity-subset* constraints. These constraints ensure that the passenger demand in the line plan is satisfied. Constraints (3) are necessary in order to ensure sufficient train capacity to those passengers who's best route contains one of the track edges (type 1 edges $E_1^T$). Constraints (4) are generated for *all* combinations of $\Phi_t(e), t > 1$. Each constraint then gives a possibility for a passenger who's best route contains one of the "fast" edges of type $t > 1$ to either use a capacity on one of the fast lines or on one of the slow lines running along the type 1 edges. A detailed description and a proof of the validity of the capacity-subset constraints (4) can be found in Goossens et al (2006). Many network segments in the Danish railway network are represented by a track graph, which is a simple path, just as an example shown on Figure 1. Therefore, the number of possible cover-subsets $\Phi_t(e)$ for each edge $e$ is at

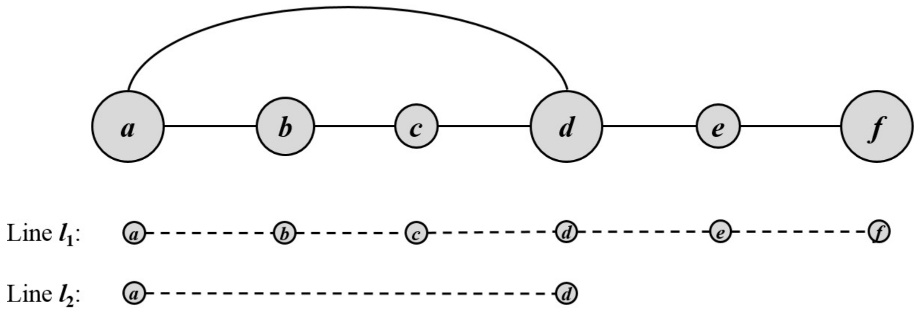


**Fig. 2** Example of a solution network $G^* = (S, A^*, \beta)$, based on a line planning solution containing two lines, where each arc $a \in A^*$ is illustrated with an edge.

most $t^{max}$ for a substantial number of edges in the network, so the complexity of the model is attractive compared to the traditional flow models.

Constraints (5) - (8) are added to ensure operational feasibility and service requirements by the Ministry of Transport: Constraints (5) ensure that the estimated number of carriages of each fleet type $r \in R$ does not exceed a given upper bound $B^r$. Constraints (6) set upper and lower bounds to the number of trains per track segment. Constraints (7) set upper and lower bounds to the number of trains halting at a station. Constraints (8) make sure that only one line is chosen from the set of lines, which cannot coexist in the same solution. Finally, constraints (9) define the feasible region for the decision variables.

## 4 Measuring the passenger quality of the solution

### 4.1 Passenger travelling time

When the optimal line planning solution to CLPP is found, we can measure the quality of the solution from the passenger point of view. With a given line plan and a set of passenger OD-relations, we can calculate the passenger flow along the type arcs in $G^{TA} = (S, A^T)$, given a fixed train capacity on all arcs. For each edge $e \in E^T$ in the type graph $G^T = (S, E^T)$, we define an *edge capacity* $\beta(e) = \sum_{i \in N^* | e_i = e}(f_i \cdot \gamma_i)$, where $N^* \subseteq N$ is the optimal combination of the lines, frequencies and train compositions. Since the line represents a return trip, the arc capacity of the corresponding arcs $\overrightarrow{a} \in A^T$ and $\overleftarrow{a} \in A^T$ is the same as the edge capacity $\beta(e)$.

Let $K$ be the set of all passenger OD-relations, and $M^k$ be the demand of the OD-relation $k \in K$. A capacitated network $G^* = (S, A^*, \beta)$ is a subgraph of $G^{TA} = (S, A^T)$, containing the subset of arcs $A^* \subseteq A^T$ with the positive capacity $\beta(a) > 0$. All arcs in $A^*$ belong to at least on line in the optimal line planning solution $N^*$.

Let us assume that the optimal solution to CLPP solved on the type graph shown on Figure 1 contains two lines: $l_1 = \{\{a, b\}, \{b, c\}, \{c, d\}, \{d, e\}, \{e, f\}\}$

and $l_2 = \{\{a, d\}\}$. The first line is a slow line, stopping at all stations between stations $a$ and $f$. The second line is a fast non-stopping line between stations $a$ and $d$. Given this optimal line plan solution, the graph $G^* = (S, A^*, \beta)$ is shown on Figure 2. Notice that the graph $G^*$ only contains edges used in the line planning solution.

Given the capacitated network $G^* = (S, A^*, \beta)$, we define a linear flow variable $y^k(v, w)$ for each $k \in K$ and $(v, w) = a \in A^*$. Let $d(v, w)$ be the traveling time on the arc $(v, w) = a$ including the halting time at the departing station. In order to find the distribution of the passenger flow in the line planning solution $N^*$, we solve the following multi-commodity network flow problem with capacity constraints (MCNFC):

$$\text{(MCNFC)} \qquad \min \sum_{k \in K} \sum_{(v,w) \in A^*} d(v, w) \cdot y^k(v, w) \tag{10}$$

$$\sum_{w \in S} y^k(v, w) - \sum_{w \in S} y^k(w, v) = \begin{cases} M^k & \text{if } v = O^k \in S, \ \forall k \in K \\ -M^k & \text{if } v = D^k \in S, \ \forall k \in K \\ 0 & \text{if } v \neq O^k, v \neq D^k, \ \forall k \in K \end{cases} \tag{11}$$

$$\sum_{k \in K} y^k(w, v) \leq \beta(v, w) \quad \forall (v, w) \in A^* \tag{12}$$

$$y^k(v, w) > 0, \quad \forall y^k(v, w) \in \mathbb{R}, \forall k \in K \tag{13}$$

Objective function (10) minimizes the travel time for the passengers, while the flow conservation constraints (11) distribute the passenger flow in the network. The capacity in the network is ensured through constraints (12).

Let $\Omega(A^*, \beta)$ be the optimal traveling time in $N^*$, calculated by solving the MCNFC. When constraints (12) are omitted from the model, it corresponds to finding the shortest paths in the network for all commodities. The shortest paths in the network $G^* = (S, A^*)$ correspond to the passengers' best routes in the given line plan found by CLPP. The optimal solution to this problem gives the optimal traveling time for all passengers in a line planning solution, given the unlimited capacity of the trains, and is denoted $\Omega(A^*)$.

By solving MCNFC without capacity constraints (12) on the graph $G^{TA} = (S, A^T)$, we get an optimal flow solution $\Omega(A^T)$, representing the best routes for all passengers with respect to the travel time, given the existence of all possible line combinations with an unlimited capacity.

If $\Omega(A^*, \beta) = \Omega(A^*) = \Omega(A^T)$, then we found a cost-optimal line planning solution CLPP, which is accidentally also optimal from passengers point of view with respect to the passenger traveling time. If $\Omega(A^*, \beta) = \Omega(A^*) > \Omega(A^T)$, then the cost-optimal line planning solution to CLPP contains enough train capacity for all passengers to take their favourite route given the set of lines in CLPP, but the choice of the lines is not optimal from the passenger traveling time point of view. If $\Omega(A^*, \beta) > \Omega(A^*)$, then the optimal solution to CLPP is suboptimal from the passenger point of view with respect to the travelling time.

Looking at the solution example on Figure 2 compared to the type graph on Figure 1, it is clear that passengers travelling between stations $d$ and $f$ did not get their favourite line routes in the solution, but would have to use the line that stops at the intermediate station $e$. In this case one would expect that at least $\Omega(A^*, \beta) > \Omega(A^T)$. Since our example does not contain any train compositions, line frequencies or the size of the passenger OD-relations, we cannot say for sure if $\Omega(A^*, \beta) > \Omega(A^*)$ or if $\Omega(A^*, \beta) = \Omega(A^*)$.

### 4.2 Passenger train line changes

Passenger traveling time $\Omega(A^*, \beta)$ does not include the time necessary to change platforms to change the train line or the waiting time between departures. From a line plan, it is not possible to calculate the passenger waiting time between the arrival of one line and the departure of the other, since the timetable is not known. However, one can assume a minimum time necessary for the passenger to change lines. Alternatively, the quality of the line plan can be measured by the number of times a passenger needs to change train lines along the route from O to D. In order to measure the quality of the line plan with respect to the number of line changes we introduce a *change-and-go* graph inspired by Schöbel and Scholl (2006).

Given an optimal line planning solution to CLPP, $N^*$, a set of stations $S$ and a set of passenger OD-relations $K$, we generate a change-and-go network $G^{CG} = (V, A^{CG})$. The set of nodes $V$ in the network contains two subsets: a set of nodes $V^{CG}$ representing station-line pairs and a set of station nodes $S$, representing passenger origin and destination stations. The set of arcs $A^{CG}$ contains three types of arcs: *travelling arcs* between the nodes in $V^{CG}$ belonging to the same line, which represent the driving activities, *change arcs* between the nodes in $V^{CG}$ belonging to the same station, which represent the change of the line at the station $s \in S$, and *origin-destination arcs* between station nodes $S$ and change-and-go nodes $V^{CG}$, representing passengers arriving or leaving the station.

A change-and-go graph example shown on Figure 3 is based on the line planning solution shown on Figure 2. The thin dotting lines represent the origin-destination arcs (for simplicity illustrated as edges), the thick dotting lines are the change arcs, while the solid lines represent the travelling arcs.

The cost $d(v, w)$ on travelling arcs is given by the travel time, while the cost of change arcs is given by the time required for changing the lines. If the travelling arcs have a capacity corresponding to $\beta(v, w) = f_i \cdot \gamma_i$, where $i \in N^*$ is part of the optimal solution to CLPP, then solving the MCNFC on the change-and-go graph would give the optimal passenger solution with respect to the travel time *and* change time.

By setting a positive fixed cost only on change arcs, the objective function of MCNFC transforms into finding the optimal number of line changes in the solution $N^*$.

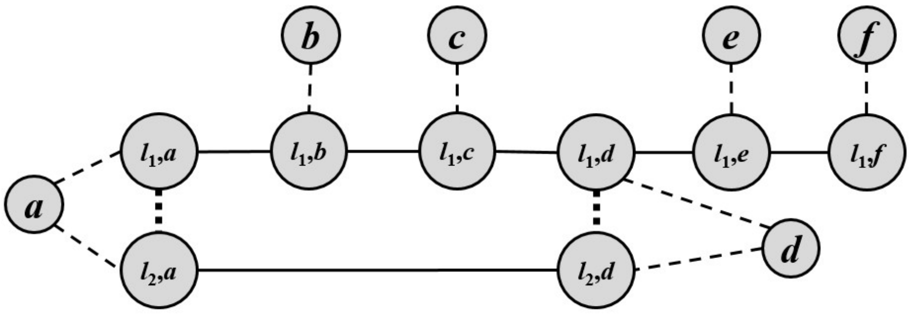


**Fig. 3** Example of a change-and-go solution network $G^{CG} = (V, A^{CG})$, based on a line planning solution shown on Figure 2, where each arc $a \in A^{CG}$ is illustrated with an edge.

## 5 Multi-criteria optimization heuristic

The two cost components $w_i^{cost}$ and $w_i^{pax}$ in the objective function of CLPP are contradictory: the passengers want more capacity, which is costly for the operator. The goal of optimization is to find a line planning solution, where passenger satisfaction is maximized while the costs are kept low.

We introduce a simple greedy heuristic, which attempts to find Pareto-optimal solutions to CLPP by adjusting the weights in the objective function iteratively. The challenge with this optimization is that the two optimization objectives are not expressed within the same mathematical model. We can only assume that the passenger-oriented solution to MCLFC would be improved by altering the passenger-related cost component $w_i^{pax}$ of CLPP.

The heuristic is based on solving the optimization problems CLPP (1) - (9) and MCNFC (10) - (13) iteratively with different values of the passenger-related cost component $w_i^{pax}$. By adjusting $w_i^{pax}$, we achieve different solutions, among which we only choose Pareto efficient ones.

Let $W^{cost}(N^*)$ be the operational cost value of the optimal CLPP solution $N^*$. Let $\Omega(N^*)$ denote the optimal passenger travelling time $\Omega(A^*, \beta)$, which is obtained by solving MCNFC on the capacitated type graph $G^* = (S, A^*, \beta)$ built from the optimal solution $N^*$.

The purpose of the heuristic is to find a set of line planning solutions $\Gamma(N^*)$, where all solutions are Pareto efficient with respect to $W^{cost}(N^*)$ and $\Omega(N^*)$.

The heuristic is presented in Algorithm 1. The algorithm begins by finding $\Omega(A^T)$, which is the shortest travelling time for all passengers in the type graph without capacity constraints $G^{TA} = (S, A^T)$. $\Omega(A^T)$ is found by finding shortest paths for all passenger OD-relations in the network.

At every iteration $j = 0, 1, ...J^{max}$ of the algorithm, the CLPP is solved with some values $w_i^{pax}(j)$, obtaining a solution $N_j^*$ with the operational cost $W^{cost}(N_j^*)$. Then a MCNFC is solved, obtaining $\Omega(N_j^*)$. The initial values of $w_i^{pax}(j)$ are set to zero, while the cost component $w_i^{cost}(j)$ only includes the *real* operational costs, i.e. no artificial weights and costs are added to it. The

**Algorithm 1** Greedy multi-criteria optimization heuristic.

```
Input: type graph G^T = (S, A^T), the shortest travelling time in the network Ω(A^T)
Initialize j := 0, w_i^pax(0) := 0
while j ≤ J^max and Ω(N_j^*) > Ω(A^T) do
   w_i^pax(j) ← w_i^pax(j − 1) + Δ_i
   Find W^cost(N_j^*) by solving CLPP
   Find Ω(N_j^*) by solving MCNFC
   if N_j^* is not dominated by any other solutions in Γ(N^*) then
      Add N_j^* to Γ(N^*)
      if N_j^* dominates any other solution N_k^* in Γ(N^*) then
         Remove N_k^* from Γ(N^*)
      end if
   end if
   Adjust Δ_i
   j ← j + 1
end while
```

values of $w_i^{pax}(j)$ are increased by a certain $\Delta_i$, which is adjusted along the run of the algorithm. Adjustments and initial values of $\Delta_i$ depend on the size of the cost component $w_i^{cost}(j)$ and on the solution values of $N_j^*$. The way of adjusting $\Delta_i$ resembles adjusting the *temperature* in Simulated Annealing metaheuristics.

If $W^{cost}(N_j^*)$ and $\Omega(N_j^*)$ are not dominated by any other solutions in $\Gamma(N^*)$, the solution $N_j^*$ is added to $\Gamma(N^*)$. If solution $N_j^*$ dominates any other solutions in $\Gamma(N^*)$, those are removed from the set.

We use $\Omega(A^T)$ as one of the termination chriteria to the algorithm. The algorithm terminates either when $\Omega(N_j^*) = \Omega(A^T)$ or the maximum number of iterations $J^{max}$ is reached.

## 6 Line planning 2016 for S-bane

Copenhagen S-bane is an isolated commuter network in the Greater Copenhagen area. Figure 4 shows the 2014 line plan on S-bane. The S-bane has 6 branches, a central section, and half circular segment intersecting with the branches. The S-bane will be equipped with a Communication Based Train Control (CBTC) signalling system by the end of 2018. Figure 5 outlines the segments of the S-bane estimated to be equipped with CBTC before the beginning of the 2016 schedule start.

Major network improvement projects are always a challenge for the railway operators. The installation of the new signalling system on S-bane is not an exception. Apart from the closures of network segments, which is mostly done during the nighttime, the installation of the new signalling system requires an upgrade of the rolling stock fleet and a train driver education. Not all drivers were planned to be licensed for CBTC at the beginning of 2016. In order to avoid inefficient train driver duties with short trips caused by the necessity of driver changes between licensed and non-licensed drivers, the line plan can


**Fig. 4** S-bane line plan 2014.


**Fig. 5** Network segments of S-bane estimated to be equipped with CBTC until 2016.

be adjusted so there is as little overlap as possible between operations on two signalling systems.

If the 2014 line plan was kept unchanged, there would be four lines running under both signalling systems in 2016: the purple line E, the green line B, the orange line C, and yellow line F. Efficient train driver duties contain as many long trips as possible in order to avoid driver changes. A long trip is a round-trip on one line. Based on the line plan 2014, many long trips in the driver duties in 2016 would contain some parts of the network with CBTC, and some parts of the network with the old ATC system. Since not all drivers would be licensed to drive on CBTC, the duties would contain many short trips between line terminal stations and Copenhagen Central station, where the main driver depot is situated. It has been observed that many driver changes at Copenhagen Central station contribute to not only more expensive duties, but also are more vulnerable to disruptions.

Therefore, at the initial stage of the S-bane 2016 line planning optimization project, the focus was entirely on obtaining a line plan with as few lines that cover both signalling systems as possible. After the first runs, the objective function was adjusted to minimize the total train driving time of the lines that cover both signalling systems in order to be able to generate as few duty time across the signalling systems as possible.

The optimization result of such a line plan is shown on Figure 6. For comparison, we evaluated five different manually generated line plan suggestions for 2016, as well as the line plans for 2014 and 2015, with respect to the same


**Fig. 6** S-bane line plan 2016 optimized with respect to minimizing the train driving minutes on lines running across two signalling systems.

objectives. The evaluation showed that 3 of the 5 manually generated plans did not satisfy all operational constraints, while the optimal line planning solution contained over 30% less train driving minutes on lines that covered both signalling systems compared with the line plans for 2014 and 2015, respectively, and 20% less than the two valid manual line plans. The generated plan measured a smaller passenger travel time than the other suggested solutions, but had a larger number of passengers who would need a line change.

The new line plan was quite different from the 2014 line plan and the 2015 line plan shown on Figure 7, which was ready for production from December 2014. Drastic changes to the line patterns are not reasonable from the passenger service perspective.

Furthermore, we received new information about the educational plans for the drivers. We learned that all drivers belonging to the depots at Hillerød and Køge stations (terminal stations of the current E line) were planned to be licensed for the new signalling system prior the first day of operation of the 2016 schedule. Therefore, it would be wise to add lines connecting these two depots in order to be able to generate long trips for licensed drivers.

Based on this information, the objective function of the line planning optimization was changed to penalize lines with following properties:

P1. Lines that connect Hillerød and Køge with other depots than themselves.
P2. Lines that are different from the lines in the 2015 schedule.


**Fig. 7** S-bane line plan 2015.


**Fig. 8** S-bane line plan 2016 optimized with respect to connecting Hillerød and Køge depots and minimizing difference from 2015.

We run a range of experiments with different parameters and weight combinations in order to determine a Pareto optimal set of solutions. The chosen line planning solution is shown on Figure 8. Only 5% of train minutes in the new plan belong to the lines penalized by P1, while less than 20% train minutes belong to the lines penalized by P2. The suggested 2016 plan is not as good as 2015 plan from the passengers point of view, but the difference is much smaller than compared to the five manually generated scenarios, which were first suggested.

The described analysis were performed at the end of year 2014, with the available information and desired objectives at that time. During the spring of 2015 the situation changed. First of all, the delay of CBTC installation plan means that the current 2015 line plan is still feasible until August 2016. Second, hence the installation is delayed, there is enough time for all train drivers to be trained, and the objectives which were important just half a year ago, are no longer relevant.

The line planning optimization tool will be useful for S-bane scheduling again in the near future, when the next change of the line plan would be necessary due to future infrastructural changes on the network.

## 7 Future implementation

Not all line plans generated by the line planning tool can be transferred into a feasible timetable. We used a simple timetable generator for evaluating the feasibility of the timetable during the work on the year 2016 line plan for S-bane, and used RailSys to simulate the timetable of the most promising solutions. In case of infeasibility, we manually adjusted line planning solutions and evaluated them in the line planning tool to make sure the changes did not have a large impact on the solution quality. The line planning tool would be advanced if some of the timetabling logic was incorporated into the constraint set of the integer programming model. A fully integrated line planning and timetabling optimization is another possibility.

Using the line planning tool to generate line plans for DSB intercity and regional trains is currently in the pipeline. Some extra operational and service requirements would need to be added to the model, such as the minimum level of direct line connections between certain cities.

We are currently supervising a Ph.D. project at the Technical University of Denmark, where a passenger-oriented line planning model is under development. A combination of a cost-oriented model and a passenger-oriented model would give DSB a powerful tool for strategic analysis.

## References


Borndörfer R, Grötschel M, Pfetsch M (2007) A column generation approach to line planning in public transport. Transportation Science 41(1):123–132

Bussieck M, Kreuzer P, Zimmermann U (1996) Optimal lines for railway systems. European Journal of Operational Research 96(1):54–63

Claessens M, van Dijk N, Zwaneveld P (1998) Cost optimal allocation of rail passenger lines. European Journal of Operational Research 110(3):474–489

Goossens JW, van Hoesel S, Kroon L (2004) A branch and cut approach for solving line planning problems. Transportation Science 38(3):379–393

Goossens JW, van Hoesel S, Kroon L (2006) On solving multi-type railway line planning problems. European Journal of Operational Research 168(2):403–424

Pfetsch M, Borndörfer R (2006) Routing in line planning for public transport. In: Operations research proceedings 2005, Springer, Berlin, pp 405–410

Schöbel A (2012) Line planning in public transportation: models and methods. OR Spectrum 34:491–510

Schöbel A, Scholl S (2006) Line planning with minimal travel time. In: 5th workshop on algorithmic methods and models for optimization of railways (ATMOS 2005), Dagstuhl seminar proceedings