Cost-Minimal Public Transport Planning

J. Pätzold, A. Schiewe, A. Schöbel

Nr. 14
Cost-Minimal Public Transport Planning*

Julius Pätzold¹, Alexander Schiewe¹, and Anita Schöbel¹

¹University of Goettingen, Lotzestr. 16-18, 37083 Göttingen, Germany
{j.paetzold, a.schiewe, schoebel}@math.uni-goettingen.de

Abstract

In this paper we discuss what a cost-optimal public transport plan looks like, i.e., we determine a line plan, a timetable and a vehicle schedule which can be operated with minimal costs while, at the same time, allowing all passengers to travel between their origins and destinations. We are hereby interested in an exact solution of the integrated problem. In contrast to a passenger-optimal transport plan, in which there is a direct connection for every origin-destination pair, the structure or model for determining a cost-optimal transport plan is not obvious and has not been researched so far.

We present three models which differ with respect to the structures we are looking for. If lines are directed and may contain circles, we prove that a cost-optimal schedule can (under weak assumptions) already be obtained by first distributing the passengers in a cost-optimal way. We are able to streamline the resulting integer program such that it can be applied to real-world instances. The model gives bounds for the general case. In the second model we look for lines operated in both directions, but allow only simplified vehicle schedules. This model then yields stronger bounds than the first one. Our most realistic model looks for lines operated in both directions, and allows all structures for the vehicle schedules. This model, however, is only computable for small instances. Finally, the results of the three models and their respective bounds are compared experimentally.

Keywords: Public Transport Planning, Integer Optimization, Line Planning, Vehicle Scheduling

1 Introduction

Public transport planning is a challenging task since it consists of several stages: network design, line planning, timetabling, vehicle- and crew scheduling. In this paper we look for a line plan in combination with a timetable and a vehicle schedule, i.e., a public transport plan. Apart from the different subproblems that need to be solved in an integrated way, there are also different objectives to be considered. A public transport plan should be passenger-friendly (mostly reflected by a short traveling time for the passengers) but also have low operating costs. For individual planning stages, such as line planning, there hence exist

*This work was partially supported by DFG under SCHO 1140/8-1 and by the Simulation Science Center Clausthal/Göttingen.
both passenger-oriented and cost-oriented models, but finding an integrated solution to this multi-stage problem is more challenging. As a result, only few papers evaluate both cost and traveling time for integrated public transport plans. A first approach in which line plans, timetables and vehicle schedules have been evaluated together under different criteria has been given in [15]. More recently, [12] proposes to measure the costs and the perceived traveling time, and evaluates public transport plans under these criteria (cf. Figure 4 in Appendix A). One possible goal could be to find the set of pareto solutions with respect to these criteria and then to choose from this set. Another goal might be to find the best possible public transport plan for the passengers while determining the minimal costs that are needed if all passengers should be transported. The former problem can be solved by a taxi-solution, providing a direct and fast connection for each origin-destination pair. Nevertheless, what a cost-optimal transportation plan would look like has not been studied so far and does not seem to be obvious. Given a line pool, [4] determine a line plan such that all origin-destination pairs can travel. The costs for the lines, however, are only approximated and not determined by the vehicle schedule. Furthermore capacities are neglected. In contrast to this work, we now take an integrated point of view and propose models for finding cost-optimal public transport plans, including lines, timetables, and vehicle schedules.

In this paper we propose models for finding cost-optimal public transport plans. More precisely, we assume that the public transport network with its stops and direct connections is given, and that the passengers’ demand is known in form of an origin-destination (OD) matrix. For a homogeneous fleet with a given capacity for each vehicle we then design a line plan, a timetable, and a vehicle schedule under the constraint that all passengers can be transported, i.e., for each passenger there exists a possible (maybe non-optimal) connection from their origin to their destination such that none of the vehicles is overloaded. We aim at solving the integrated system exactly, meaning that we do not provide iterative heuristics as in [6, 33, 36] or a sequential approach as the one in [24].

For the single planning stages line planning, timetabling, and vehicle scheduling, models and algorithms are well-researched. For line planning, cost-oriented models (e.g. [9, 17, 37]) and passenger-oriented models (e.g. [2, 7, 34]) are known, see [32] for a survey. (Periodic) timetabling focuses on the passengers and is the hardest of the three problems. Exact approaches to this problem can be found in [35, 22, 28, 18] and heuristics in [23, 16, 25] and references therein. Integrating the passengers’ routes in timetabling is an ongoing problem, see [3, 31, 13]. For vehicle scheduling we refer to the survey in [5].

2 A cost-optimal LTS-plan

In this section we formally describe what a feasible public transport plan (LTS-plan), consisting of a line plan (L), a timetable (T), and a vehicle schedule (S), is and how its quality can be evaluated. We restrict ourselves to periodic LTS-plans (including the vehicle scheduling) in this paper.

Notation 1. The following input data is needed:

- a public transport network \( PTN = (V, E) \) with a set of stops \( V \) and direct connections \( E \) between them,
• for every edge $e \in E$:
  – a length (in kilometers) $\text{length}_e$,
  – a lower bound on the travel time along the edge $L_e^{\text{drive}}$,
• a lower bound $L_{\text{wait}}$ for the time vehicles have to wait at every stop,
• a minimal turnaround time for vehicles $L_{\text{turn}}$, denoting the minimal time a vehicle has to wait at the end of a line. We assume that $L_{\text{wait}} \leq L_{\text{turn}}$.
• an OD-matrix $W$ with entries $W_{uv}$ for each pair of stops $u, v \in V$, denoting how many passengers want to travel from an origin $u$ to the destination $v$. A pair of stations $u, v \in V$ with $W_{uv} > 0$ is called an OD-pair.
• a capacity $\text{Cap}$ being the maximal number of passengers each vehicle can transport,
• cost parameters
  – $c_{\text{time}}$ costs per hour for a vehicle driving,
  – $c_{\text{length}}$ costs per kilometer for a vehicle driving.

We assume that the fixed costs (cost of a vehicle, administration, etc.) are included in the costs per hour and the costs per kilometer, as is often done in practice.

With this input data we then look for an LTS-plan, whose objects are described next.

**Line plan $L$**

A line is a path through the PTN. A line plan is a set of lines $\mathcal{L}$, each of them operated once in the planning period (often an hour). A line plan is feasible if every passenger can be transported, i.e., if for every OD-pair $(u, v)$ there exist

• a set of directed paths $P_{uv}$ from $u$ to $v$, $P_{all} = \bigcup_{u, v \in V} P_{uv}$, and
• weights $w_p$ for each path $p \in P_{uv}$ such that $\sum_{p \in P_{uv}} w_p = W_{uv}$ and such that for every edge $e$ it holds that

$$\sum_{p \in P_{all} : e \in p} w_p \leq \text{Cap} \cdot |\{l \in \mathcal{L} : e \in l\}|.$$  \hfill (1)

We furthermore assume that lines are simple paths and that every line is operated in both directions. We do not forbid identical lines, i.e., there may be multiple lines with the same path. In our setting we allow any path as a possible line (as also done in [2]) in contrast to many papers which require a line pool of limited size.
Timetable T

Given a set of lines $L$, a timetable assigns a time to every departure and arrival of each line at its stops. Determining a (periodic) timetable is the hardest of the three problems line planning, timetabling, and vehicle scheduling, and even finding a feasible timetable that respects the upper and lower bounds on driving, waiting, transfer and turnaround activities is intractable. Nevertheless, in our case we assume no upper bounds on transfer and turnaround activities, hence a feasible timetable exists for every possible line plan $L$ (since the timetable for each line can then be determined separately.). Since we are interested in determining an LTS-plan with minimal costs we furthermore need not care about optimizing the travel time of the passengers, meaning that any feasible timetable is sufficient. More precisely, we can neglect the timetabling as a separate planning stage in cost-optimal planning and simply use the arrival and departure times which are determined by the vehicle schedule.

Vehicle schedule $S$

Given a line plan a vehicle schedule determines the number of vehicles and the exact routes of the vehicles for operating the lines. We construct a set of trips $L'$ which contains two directed lines for every (undirected) line $l \in L$, one in forward and the other in backward direction.

A route of a vehicle is given by the sequence of (directed) lines it passes,

$$r = (l'_1, \ldots, l'_k), \quad l'_i \in L'$$

whereby we require that the $l'_i$, $i = 1, \ldots, k$ are pairwise distinct. We assume that after having taken the last trip $l'_k$ in a route, the vehicle starts again with $l'_1$.

This sequence $r$ is interpreted as follows: A vehicle starts with operating line $l'_1$ at some point in time, $x$. At the end of line $l'_1$ it drives to the start point of line $l'_2$, operates this line, and so on. At the end of line $l'_k$ the vehicle returns to the start point of $l'_1$ and starts from the beginning. In order to ensure the required periodicity of the schedule, the vehicle needs to start after an integer multiple of the period $T$, i.e., at a time $y = x + d_r \cdot T$, whereby the integer $d_r$ is the number of periods needed for a complete operation of the route $r$.

A vehicle schedule thus consists of a set of routes $\mathcal{R}$. It is feasible if each directed line in $L'$ is contained in exactly one route, i.e., if

$$|\{r \in \mathcal{R} : l' \in r\}| = 1 \quad \forall l' \in L'.$$  \hspace{1cm} (2)

With these assumptions in place we can then define what an LTS-plan is.

**Definition 2.** An LTS-plan is a tuple $(L, \mathcal{R})$, such that

- $L$ is a feasible line plan, i.e., it satisfies (1),
- $\mathcal{R}$ is a feasible vehicle schedule for the directed lines $L'$, i.e., it satisfies (2).
Costs of an LTS-plan

The costs of an LTS-plan are given by the distance driven by all vehicles and its total duration. Since we compute a periodic schedule, we consider the costs per planning period $T$.

A vehicle route $r$ consists of (directed) lines $l' \in \mathcal{L}'$. Hence, we first determine time and duration of a line $l'$, namely,

$$\text{length}_l = \sum_{e \in l} \text{length}_e$$  \hspace{1cm} (3)

$$\text{dur}_l = (|l| - 1)L^\text{wait} + \sum_{e \in l} L^\text{drive},$$  \hspace{1cm} (4)

where $|l| := \{e \in E | e \in l\}$ and (4) uses the fact that it is always cheaper to operate a line as fast as possible. For the empty rides between a pair of lines $l_1'$ and $l_2'$ we can use the PTN to determine the parameters

$$\text{length}_{l_1',l_2'} = \text{length} \text{ when driving from the last station of } l_1' \text{ to the first station of } l_2'$$

$$\text{time}_{l_1',l_2'} = \text{time\ for\ driving\ from\ the\ last\ station\ of\ line } l_1' \text{ to the first station of } l_2'$$

The minimum turnaround time (usually accounting for a driver’s break) has to be added to the duration of an empty ride. This yields

$$\text{dur}_{l_1',l_2'} = L^\text{turn} + \text{time}_{l_1',l_2'}.$$  \hspace{1cm} (5)

The number of kilometers a given LTS-plan covers is determined by summing up the kilometers of each single route, i.e.,

$$\text{length}(\mathcal{L}, \mathcal{R}) = \sum_{l' \in \mathcal{L}'} \text{length}_{l'} + \sum_{r=(l_1',\ldots,l_k') \in \mathcal{R}} \sum_{i=1}^{k_r} \sum_{l'=l_i'}^l \text{length}_{l_i',l_{i+1}}$$

$$= \sum_{l' \in \mathcal{L}} 2 \cdot \text{length}_{l'} + \sum_{r=(l_1',\ldots,l_k') \in \mathcal{R}} \sum_{i=1}^{k_r} \sum_{l'=l_i'}^l \text{length}_{l_i',l_{i+1}}$$

with $l_{k_r+1} := l_1'$. The duration of a route $r = (l_1',\ldots,l_{k_r}') \in \mathcal{R}$ is measured by the number of time periods $\text{dur}_r$ needed. This can be formally computed by

$$\text{dur}_r = \left\lceil \sum_{i=1}^{k_r} \text{dur}_{l_i'} + \text{dur}_{l_i',l_{i+1}} \right\rceil_T$$  \hspace{1cm} (6)

with $\lceil a \rceil_T := \min\{n \in \mathbb{N} | n \cdot T \geq a\}$ for any $a \in \mathbb{R}$ and $l_{k_{r}+1} := l_1'$. The overall duration is hence given as

$$\text{dur}(\mathcal{L}, \mathcal{R}) = \sum_{r \in \mathcal{R}} \text{dur}_r.$$  \hspace{1cm} (7)

Finally, the cost function is defined as

$$g(\mathcal{L}, \mathcal{R}) := c_{\text{time}} \cdot \text{dur}(\mathcal{L}, \mathcal{R}) + c_{\text{length}} \cdot \text{length}(\mathcal{L}, \mathcal{R}).$$  \hspace{1cm} (8)
This allows us to define the optimization problem we are concerned with in this paper.

**Problem (cost-opt LTS):** Given the input data from Notation 1, find a feasible LTS-plan \((\mathcal{L}, \mathcal{R})\) with minimal costs \(g(\mathcal{L}, \mathcal{R})\).

Traditionally, calculating an LTS-plan consists of solving a series of problems in a sequential order, as can be seen in [8, 10, 19]. A sequential approach, however, is flawed, since the costs are mainly determined by the vehicle schedule, which constitutes the last step of the planning process. Nevertheless, this has been tackled in [24] by a heuristic approach. The aim of our paper, however, is to find the exact cost minimum of the integrated problem. In order to address this issue we present three different models for minimizing the costs of the resulting LTS-plan (see Figure 1).

<table>
<thead>
<tr>
<th>Model 1: Load Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 2: Integrating up to Line Planning</td>
</tr>
<tr>
<td>Model 3: Integrating up to Timetabling and Vehicle Scheduling, i.e., solving it all</td>
</tr>
</tbody>
</table>

The first model aims at distributing the OD-pairs in a cost-optimal way (called load generation). Although it only concerns this very first step we can show that this determines the minimal costs of an integrated LTS-plan under certain conditions. The second model integrates load generation and line planning, minimizing a cost function that approximates (now in greater detail) the costs of a resulting LTS-plan. Finally, the third model presents an IP formulation for integrating load generation, line planning, timetabling, and vehicle scheduling; it hence provides an exact model for (cost-opt LTS).

### 3 Model 1: Creating a Cost-efficient Load

Line planning is often decomposed into two steps. In the first step, all OD-pairs \((u, v)\) are routed through the PTN resulting in paths \(P_{uv}, \bigcup_{u,v \in V} P_{uv}, \) and weights \(w_p\) for every path \(p \in P_{uv}\) (with \(\sum_{p \in P_{uv}} w_p = W_{uv}\)). This data is then used to define the loads

\[
    f_{e}^{\min} = \left[ \sum_{p \in P_{uv}, e \in p} w_p \cdot \frac{1}{\text{Cap}} \right]
\]
specifying how often an edge $e \in E$ in the PTN has at least to be served by some vehicle. In the second step, the line planning problem is solved using these minimal frequencies. Normally the $f_{e}^{\text{min}}$ are calculated assuming that all passengers travel on their shortest path in the PTN to their destination. Since we are interested in finding a cost-minimal LTS-plan, we do not want to work with that assumption. In our system we require just enough capacities so that every passenger has some possibility to travel to their destination. We use this insight to find a load that eventually even leads to a cost-minimal LTS-plan.

Of course, in this early planning stage we do not yet have all information to exactly determine the costs of the resulting LTS-plan, since they depend on the line plan and the vehicle schedule. Nevertheless, we can already approximate the costs with the following model.

**Model 1.** Given the input data from Notation 1, calculate a load (i.e., $f_{e}^{\text{min}}$ for all $e \in E$) that aims at minimizing the cost of an LTS-plan.

$$\min \ c_{\text{time}} \cdot \text{dur} \cdot T + c_{\text{length}} \sum_{e \in E} 2 \cdot \text{length}_e \cdot f_{e}^{\text{min}} \quad (9)$$

s.t. $\sum_{e \in E} 2f_{e}^{\text{min}}(L_{e}^{\text{drive}} + L_{e}^{\text{wait}}) \leq T \cdot \text{dur}$ \quad (10)

$$\sum_{u \in V} f_{(i,j),u} \leq f_{e}^{\text{min}} \cdot \text{Cap} \quad \forall i, j \in V \text{ with } \{i, j\} \in E$$ \quad (11)

$$\sum_{i \in V \setminus \{i, v\} \in E} f_{(i,v),u} = W_{uv} + \sum_{i \in V \setminus \{v, i\} \in E} f_{(v,i),u} \quad \forall u \in V \quad \forall v \in V \setminus \{u\}$$ \quad (12)

$$\sum_{i \in V \setminus \{u, i\} \in E} f_{(u,i),u} = \sum_{v \in V} W_{uv} \quad \forall u \in V$$ \quad (13)

**Variables:**

- $f_{(i,j),u}$ – number of passengers starting from stop $u \in V$ traveling on arc $(i, j)$ for some $i, j \in V$ with $\{i, j\} \in E$ (non-negative, continuous)
- $f_{e}^{\text{min}}$ – how often edge $e$ has to be covered (integer)
- $\text{dur}$ – total duration (counted in periods) (integer)

In this model we define from every stop $u \in V$ in the PTN some passenger flow going to all destinations $v \in V$. In order not to mix up passengers starting from different stations we accordingly have to define $|V|$ different flows. The constraints (12) and (13) describe the flow conservation constraints. In order to restrict the number of passengers traveling on a certain edge in the network we defined the capacity constraints (11). Note that the flow variables $f_{(i,j),u}$ for $u \in V$ are defined on directed edges $(i, j)$ whereas the minimal frequencies $f_{e}^{\text{min}}$ are defined on undirected edges $\{i, j\} = e \in E$. Finally constraint (10) rounds the minimal duration up to the next multiple of a time period $T$ and the objective function gives the costs which are needed in the best case, namely for a vehicle schedule without any empty ride and as few time loss (through the periodicity) as possible.

The following theorem shows that Model 1 is indeed an approximation of (cost-opt LTS), as its optimal solution yields a lower bound.
Theorem 3. The optimal objective value of Model 1 is a lower bound on the optimal objective value of (cost-opt LTS).

Proof. See Appendix C.

For large problem instances a speed-up of the solution process is possible by adding the following valid inequalities to Model 1.

Lemma 4. Let \((X, Y)\) be some cut, i.e., some disjoint partition of all nodes in the PTN with \(E_{\text{cut}} = \{(i, j) = e \in E | i \in X \text{ and } j \in Y\}\) being all cut edges. Then it holds that

\[
\sum_{u \in X} \sum_{v \in Y} W_{uv} \leq \text{Cap} \cdot \sum_{e \in E_{\text{cut}}} f_{e}^{\text{min}}.
\]

Proof. See Appendix C.

In the computational experiments (Section 6) we investigated adding these valid inequalities, which resulted in an improvement of the runtime of up to 50%.

Model 1 does not only yield some lower bound, but we can even construct an optimal solution to (cost-opt LTS) if a particular assumption is met.

Theorem 5. Let \(L^{\text{wait}} = L^{\text{turn}}\) and let the graph \(G = (V, \bar{E})\) with \(\bar{E} = \{e \in E : f_{e}^{\text{min}} > 0\}\) for an optimal solution \(f_{e}^{\text{min}}\) of Model 1 be connected. Then the optimal objective of Model 1 is equal to the optimal objective of (cost-opt LTS).

Proof. See Appendix C.

In case the assumption \(L^{\text{wait}} = L^{\text{turn}}\) does not hold, we still get a feasible solution and therefore an upper bound for (cost-opt LTS), when we slightly modify Model 1.

Definition 6. We define an adjusted version of Model 1, where \(L^{\text{wait}}\) is replaced by \(L^{\text{turn}}\) in constraint (10), to be Model 1*.

Corollary 7. The solution \((L, R)\) constructed in the proof of Theorem 5 is an upper bound for (cost-opt LTS) and can be found by solving Model 1*.

If we allow that lines do not have to be bidirectional and simple paths in the PTN, we can always obtain an optimal solution to (cost-opt LTS) by just solving Model 1. This can be done by converting the Eulerian Cycle constructed the proof of Theorem 5 into one big line.

Corollary 8. Let \(L^{\text{wait}} \leq L^{\text{turn}}\). Then the optimal objective value of Model 1 is equal to the optimal objective of (cost-opt LTS) if we allow directed and non-simple lines.

This, of course, may lead to non-practical lines, as can be seen in the following example.

Example 9. We examine the solution provided by Corollary 8 on a small example. Consider the PTN given in Figure 2, with \(\text{Cap}\) passenger traveling from \(v_1\) to \(v_5\) and 1 passenger traveling from \(v_2\) to \(v_3\). Then the solution provided by Model 1 is given by lower bounds of \([1, 2, 1, 1]\) and the vehicle schedule of Corollary 8 is depicted in Figure 2, where the edges are numbered in the order of their usage. As can be seen here, the resulting line structure is not suitable for a practical public transport system, since it contains a cycle.
4 Model 2: Integrating Load Generation and Line Planning

Although we can already find a cost-optimal solution using Model 1, this only works in the special case of $L^{\text{wait}} = L^{\text{turn}}$. We have seen that for $L^{\text{wait}} < L^{\text{turn}}$ the resulting line plan consists of directed lines (without their symmetric counterparts) and the lines may contain circles. We therefore further explore the next steps for obtaining an LTS-plan in which the lines satisfy the usual requirements. To this end, we combine the load generation of Model 1 with line planning to improve the approximation of the cost objective of the overall LTS-plan. This idea is approached by the following model.

Model 2. Given the input data from Notation 1, calculate a load $f_e^{\text{min}}$ and a line plan $\mathcal{L}$ that aim at minimizing the costs of an LTS-plan.

$$\begin{align*}
\min & \quad c_{\text{time}} \cdot \text{dur} \cdot T + c_{\text{length}} \sum_{l=1}^{L} \sum_{e \in E} 2x_{e,l} \cdot \text{length}_e \\
\text{s.t.} & \quad (11) - (13) \\
& \quad \sum_{l=1}^{L} \left( 2z_l(L^{\text{turn}} - L^{\text{wait}}) + \sum_{e \in E} 2(L^{\text{drive}} + L^{\text{wait}}) \cdot x_{e,l} \right) \leq \text{dur} \cdot T \\
& \quad \sum_{l=1}^{L} x_{e,l} \geq f_e^{\text{min}} \quad \forall e \in E \\
& \quad x_{e,l} \leq z_l \quad \forall e \in E \forall l \in [L] \\
& \quad \sum_{e \in E} x_{e,l} \geq z_l \quad \forall l \in [L] \\
& \quad \sum_{e \in E: s \in e} x_{e,l} \leq 2 \quad \forall s \in V \forall l \in [L] \\
& \quad 2x_{e,l} \leq y_{i,l} + y_{j,l} \quad \forall l \in [L] \forall (i, j) = e \in E \\
& \quad \sum_{s \in V} y_{s,l} = \sum_{e \in E} x_{e,l} + z_l \quad \forall l \in [L] \\
& \quad \sum_{(i,j) \in E: i \in C \text{ and } j \in C} x_{e,l} \leq |C| - 1 \quad \forall \text{ circles } C \subseteq E \forall l \in [L]
\end{align*}$$

Coefficients:
- $L$ – maximal possible number of lines (integer) and $[L] := \{1, \ldots, L\}$. 

Figure 2: Solution of Model 1 for Example 9
Variables:

- $z_l$ – is 1 iff line $l$ is non-empty. (binary)
- $y_{s,l}$ – is 1 iff stop $s$ is contained in line $l$. (binary)
- $x_{e,l}$ – is 1 iff edge $e$ is contained in line $l$. (binary)
- $dur$ – total duration of all lines (counted in periods) (integer)
- $f_{e}^{\text{min}}$ – as in Model 1, including the variables $f_{e,u}$ and constraints (11) - (13) from Model 1.

This model finds some feasible line plan. First the $z_l$-variables determine if line number $l$ is a line or empty. Constraint (17) and (18) ensure this. Now we need for every index $l$ that for every stop of some line there are at most two incident edges (constraint (19)). This ensures that the $x_{e,l}$ variables form circles or paths. To ensure that they form only one connected path we could consider them as flow variables. Here, we decided to add $y$-variables for every visited stop and count the number of stops that a line visits. The $y$-variables are set to one for the incident nodes of all edges the line visits in (20). We then can ensure that there is some connected path by requiring that there exists exactly one more stop than edges in a line in constraint (21). Finally we need to rule out subtours which is done by constraint (22) (As usual they are added by constraint generation procedures). The variables $f_{e}^{\text{min}}$ taken from Model 1 help us to determine feasibility of the line plan, which is done by constraint (16). Finally we round the duration up to the next multiple of a time period, which is done by (15).

The objective function is again a lower bound on the exact costs of an LTS-plan. This is shown in the next theorem.

**Theorem 10.** The optimal objective value of Model 2 is a lower bound on the optimal objective value of (cost-opt LTS) and an upper bound to the optimal objective value of Model 1.

**Proof.** See Appendix C. □

We can again construct a feasible solution for (cost-opt LTS) from the solution of Model 2 in the case that we are only interested in line-pure vehicle schedules. In such schedules, every vehicle serves the same line, alternating between its forward and its backward direction. More formally:

**Definition 11.** A solution to (cost-opt LTS) is called line-pure if $\mathcal{R} = \{ r_l : l \in \mathcal{L} \}$, with $r_l = (l^+, l^-)$ being the route that contains only the forward and backward direction of line $l \in \mathcal{L}$.

We now show that the following slight modification of Model 2 can find a cost-optimal LTS-plan under the restriction that only line-pure vehicle schedules are allowed.

**Definition 12.** Consider Model 2 and replace constraint (15) by

$$2z_l(L_e^{\text{turn}} - L_e^{\text{wait}}) + \sum_{e \in E} 2(L_e^{\text{drive}} + L_e^{\text{wait}}) \cdot x_{e,l} \leq d_l \cdot T \quad \forall l \in [L]$$

(23)

$$\sum_{l=1}^{L} d_l = dur$$

(24)
with integer variables $d_i \in \mathbb{N}$. We call this modified version Model 2*.

Restricting ourselves to a special structure of the vehicle schedules, we are still able to obtain the optimal solution to \((\text{cost-opt LTS})\) (under some assumptions) by simply considering loads and the lines. This is the main result of this section.

**Theorem 13.** An optimal solution to Model 2* solves \((\text{cost-opt LTS})\) under the restriction that only line-pure vehicle schedules are allowed.

**Proof.** See Appendix C.

For the general case of \((\text{cost-opt LTS})\), Model 2* still finds a feasible solution and therefore provides an upper bound to \((\text{cost-opt LTS})\).

**Corollary 14.** The optimal objective value to Model 2* imposes an upper bound on the optimal objective value of \((\text{cost-opt LTS})\).

**Example 15.**

![Figure 3: Solution of Model 2](image)

We continue Example 9 and now consider the solution constructed in Theorem 10. These now provide simple lines, resulting in the line-pure vehicle schedule depicted in Figure 3, improving on the line structure of Example 9. The first line is depicted in red, the second is dashed in green. The lines here look much more reasonable for practical implementation than the solution which was obtained by Model 1*.

## 5 Model 3: Integrating Timetabling and Vehicle Scheduling

In Model 1 and Model 2 we did not consider all subproblems of \((\text{cost-opt LTS})\), especially we did not include a proper vehicle scheduling. With the following model we want to overcome this issue and formulate the whole problem in an integrated way.

**Model 3.** Given the input data from Notation 1, find a feasible LTS-plan \((\mathcal{L}, \mathcal{R})\) with minimal costs, i.e., minimizing $g(\mathcal{L}, \mathcal{R})$.

This problem can be formulated as an integer linear program. It is described in Appendix B.

As can be seen in Section 6, the integrated problem cannot be solved even for instances of small size. This is due to its enormous number of variables including a trip for every possible line in the network. Nevertheless, Model 3 can be used
if enough variables are fixed. We hence can combine it with Model 2 by fixing the lines in Model 3 to the optimal lines computed by Model 2. This means that we only need to consider the constraints (25)-(28) and (34), additionally guaranteeing that every trip in \( L' \) is covered exactly once. The result is a tractable model for medium-sized instances.

6 Experiments

In the computational experiments we implemented the three proposed models with LinTim (see [15, 30, 1]) and tested them on four different datasets. These datasets are described in in Table 1 and depicted in Figure 5, Appendix A.

<table>
<thead>
<tr>
<th>Instance</th>
<th>Nodes</th>
<th>Edges</th>
<th>Passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>5</td>
<td>4</td>
<td>141</td>
</tr>
<tr>
<td>Toy</td>
<td>8</td>
<td>8</td>
<td>2622</td>
</tr>
<tr>
<td>Grid</td>
<td>25</td>
<td>40</td>
<td>2546</td>
</tr>
<tr>
<td>Germany</td>
<td>250</td>
<td>326</td>
<td>385868</td>
</tr>
</tbody>
</table>

Table 1: Properties of the examined datasets

We implemented Model 1, Model 1*, Model 2, Model 2* and Model 3 using Gurobi 8.0 as MIP solver with default settings. We tested all implementations on a compute server (6 cores of Intel(R) Xeon(R) CPU X5650 @ 2.67GHz, 78 GB RAM) with a time limit of 3 hours per test case. For each model and each instance we considered two different cases: Either \( L_{\text{turn}} = L_{\text{wait}} \) or \( L_{\text{turn}} > L_{\text{wait}} \) to distinguish the cases where Model 1* is able to find an optimal solution and where it is not. We obtained the results depicted in Tables 2 and 3. A symbol ◦ denotes that the problem has not been solved to optimality and hence only the best found upper or lower bound is presented.

<table>
<thead>
<tr>
<th>Instance</th>
<th>Model 1</th>
<th>Model 1*</th>
<th>Model 2</th>
<th>Model 2*</th>
<th>lb</th>
<th>ub</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>130</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Toy</td>
<td>1424</td>
<td>1424</td>
<td>1424</td>
<td>1696</td>
<td>1270(^o)</td>
<td>1460(^o)</td>
</tr>
<tr>
<td>Grid</td>
<td>1034</td>
<td>1034</td>
<td>1034</td>
<td>1034</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Germany</td>
<td>61354(^o)</td>
<td>84694(^o)</td>
<td>54148(^o)</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Objective values for the case of \( L_{\text{turn}} = L_{\text{wait}} \)

For each of the three models there exist two columns. The left column contains a lower bound to (cost-opt LTS), whereas the right column contains an upper bound, i.e., the objective value of the best found feasible solution.

We observe for Model 1 that in the case \( L_{\text{turn}} = L_{\text{wait}} \) it almost always finds the optimal objective value within the specified time limit of 3 hours. Only in our biggest instance we cannot get an optimal solution within the time limit (we still have a gap of 27.5% here). For the case \( L_{\text{turn}} > L_{\text{wait}} \) there exists a
Model 2 can solve the two smallest instances easily, but starts having trouble with the time limit for Grid. For Germany it is not able to find a feasible solution within the specified time limit. Regarding the solution quality, we see that the lower bound given by Model 2 is only in a single case sharper than the lower bound given by Model 1. On the other hand, the upper bounds found by Model 2* never have smaller objective values than Model 1*.

Model 3 is already on the toy instance not able to find an optimal solution within 3 hours. The obtained objective values for Linear and the bounds for Toy are consistent with the values given in Models 1 and 2. For the bigger instance, even the precomputation of the complete line pool for Model 3 was not possible anymore.

We included the optimal solution provided by Model 1 in the ongoing competition on Grid started in [12]. The current progress can be found at [29]. The solution with the smallest costs so far was P5, provided in [20]. As can be seen in Figure 4 in Appendix A, our cost-minimal solution decreases the costs by about 23%, compared to P5.

We finally investigate the influence of valid inequalities introduced in Lemma 4 on the runtime of Model 1. We restricted this investigation to Grid, since the runtime for the smallest two instances is already less than a second, and for Germany it is already non-trivial to determine “good” cuts of the network. For Grid, however, we took all horizontal and all vertical cuts of the network, whose PTN is depicted in Figure 5, into the model. With this improvement we were able to speed up the solution process significantly with respect to runtime and number of explored MIP nodes, as can be seen in Table 4.

<table>
<thead>
<tr>
<th>Instance</th>
<th>Model 1</th>
<th>Model 1*</th>
<th>Model 2</th>
<th>Model 2*</th>
<th>lb</th>
<th>ub</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>80</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
<td>130</td>
</tr>
<tr>
<td>Toy</td>
<td>1424</td>
<td>1474</td>
<td>1424</td>
<td>1696</td>
<td>1288°</td>
<td>1539°</td>
</tr>
<tr>
<td>Grid</td>
<td>1034</td>
<td>1134</td>
<td>1030°</td>
<td>1140</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Germany</td>
<td>61354°</td>
<td>86086°</td>
<td>54148°</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Objective values for the case of $L_{\text{turn}} > L_{\text{wait}}$

<table>
<thead>
<tr>
<th>parameters</th>
<th>no cuts</th>
<th>cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 1*</td>
</tr>
<tr>
<td>Nodes explored</td>
<td>46557</td>
<td>26391</td>
</tr>
<tr>
<td>Runtime in sec</td>
<td>23.18</td>
<td>12.6</td>
</tr>
</tbody>
</table>

Table 4: Runtime improvements with Lemma 4 on Grid for $L_{\text{turn}} > L_{\text{wait}}$
7 Outlook

We propose three models to compute cost-optimal public transport plans. For the first two models we derive optimality conditions and with the third model we present an IP formulation for the integrated exact model. The computational experiments show that the implementation of the models is computationally tractable.

Model 1 is able to compute cost-optimal solutions up to Grid outperforming previous approaches to tackle this problem. For large networks the model provides bounds of good quality in a reasonable amount of time. Model 2 finds optimal line-pure LTS-plans. Finally, Model 3 yields a cost-optimal LTS-plan without requiring any further assumptions.

For future work we plan to sharpen the formulation of Model 1 by identifying good cuts. It would hopefully be the case that better cuts lead to a further decrease of the computation time, especially for the large instances.

Furthermore it would be interesting to not only find a solution with minimal costs, but to find a lexicographic solution, i.e., the cost-optimal solution with the best travel time for the passengers. To this end, we can include the passengers’ travel time in Model 3 which will most likely further increase the computation time of the model. To use this model effectively, more work in speed-up techniques is necessary. Promising ideas include column generation and decomposition techniques, similar to the methods presented in [21].
A Figures

Figure 4: Multiple solutions for Grid (see [29]), evaluated by their cost per hour and perceived travel time. With our models we were able to find a cost-minimal solution. Its objective value is depicted as a red line.

(a) The Linear network  (b) The Toy network  (c) The Grid network

Figure 5: The instances used in the experiments

B The integrated model

To formulate the integrated model, we need a notation for the event-activity network $\mathcal{N} = (\mathcal{E}, \mathcal{A})$ (see, e.g., [18, 19, 22, 26, 27]). The set of events $\mathcal{E}$ consists of all departures and all arrivals of all lines at all stops and two additional OD-events $((u, \text{dep}), (u, \text{arr}))$ per stop $u$ for passengers to enter and leave the network, denoted as $\mathcal{E}_{OD}$. The set $\mathcal{A}$ connects the events by driving, waiting and transfer activities. The OD-events are connected to each departure event of the corresponding stop using OD-activities ($\mathcal{A}_{OD}$). Using this, we can now
formulate the integrated model. Let further denote with $A\setminus A_{OD}$ all activities in $A \setminus A_{OD}$ that are included in a directed line $l' \in L'$.

**Model** (Model 3 continued). Given the input data from Notation 1, find a feasible LTS-plan $(L, R)$ with minimal costs, i.e., minimizing $g(L, R)$.

$$\min \sum_{v \in V} \text{cost}_v$$

s.t. 

$$\text{dur}_r \geq \frac{1}{T} \cdot \sum_{l' \in L'} x_{l',r} \cdot \text{dur}_{l'} + \sum_{l',l'' \in L'} x_{(l',l''),r} \cdot \text{dur}_{l',l''} \quad \forall r \in [R]$$  \hspace{1cm} (25)

$$\text{length}_r \geq \sum_{l' \in L'} x_{l',r} \cdot \text{length}_{l'} + \sum_{l',l'' \in L'} x_{(l',l''),r} \cdot \text{length}_{l',l''} \quad \forall r \in [R]$$  \hspace{1cm} (26)

$$\text{cost}_r \geq c_{\text{length}} \cdot \text{length}_r + c_{\text{time}} \cdot \text{dur}_r \quad \forall r \in [R]$$  \hspace{1cm} (27)

$$\sum_{l' \in L'} x_{l',r} = x_{b(l'),r} \quad \forall l' \in L', \forall r \in [R]$$  \hspace{1cm} (28)

$$\sum_{r \in R} x_{l',r} = \sum_{v \in V} x_{b(v'),r} \quad \forall l' \in L'$$  \hspace{1cm} (29)

$$\sum_{r \in R} \sum_{u,v \in V} f_{a,(u,v)} \quad \forall l' \in L', \forall a \in A_{pl}$$  \hspace{1cm} (30)

$$\sum_{(i,j) = a \in A} f_{a,(u,v)} = \sum_{(j,i) = a \in A_t} f_{a,(u,v)} \quad \forall p \in P, \forall j \in E \setminus E_{OD}$$  \hspace{1cm} (31)

$$\sum_{(i,j) = a \in A_{OD}} f_{a,(u,v)} = W_{uv} \quad \forall u, v \in V, \forall j = (v, \text{arr}) \in E_{OD}$$  \hspace{1cm} (32)

$$\sum_{(j,i) = a \in A_{OD}} f_{a,(u,v)} = W_{uv} \quad \forall u, v \in V, \forall j = (u, \text{dep}) \in E_{OD}$$  \hspace{1cm} (33)

$$\sum_{(l'_1,l'_2) \in U'} x_{(l'_1,l'_2),r} \leq |U'| - 1 \quad \forall U' \subseteq L' \times L', \forall r \in [R]$$  \hspace{1cm} (34)

$$\text{dur}_r \in \mathbb{N} \quad \forall r \in [R]$$  \hspace{1cm} (35)

**Coefficients:**

- $R$: number of possible vehicle routes, we assume it to be sufficiently large
- $L'$: the set of all possible directed lines in the network, $b(l')$ denotes the backwards direction for a directed line $l'$, $l$ is the corresponding undirected line.

**Variables:**

- $x_{l',r}$: is 1 iff the directed line $l'$ is part of route $r$
- $x_{(l'_1,l'_2),r}$: is 1 iff lines $l'_1$ and $l'_2$ are served directly after each other in route $r$
This model finds a cost-optimal LTS-plan (i.e., line plan, timetable and vehicle schedules). The \( f \) variables determine the passenger flow, satisfying the classical flow conservation constraints (31)-(33) and creating coupling constraints for the vehicle routes \( r \) in (30), determined by the \( x \)-variables. The duration and length of the routes are determined in (25) and (26) and then combined in (27) to determine the costs. Of course, the vehicle routes need to satisfy flow conservation as well (see (28)). (34) are the subtour elimination constraints. Constraint (29) ensures that every line is served in both directions.

The model is too large to be solved for realistic instances. One possibility to reduce its size is to start with a line pool of limited size (e.g. as generated in [14] or from Model 2) or to use column generation approaches as in [2].

### C Proofs

**Proof of Theorem 3.** Let \((\mathcal{L}, \mathcal{R})\) be some feasible solution to (cost-opt LTS). Since the line plan is feasible we can construct some feasible flow from it by setting \( f_{e}^{\text{min}} = |\{ l \in \mathcal{L} | e \in l \}| \) and \( f_{e,u} = \sum_{p \in P : e \in p} w_{p} \). Now we get for all \( i,j \in V \) with \( \{ i,j \} \in E \)

\[
\sum_{u \in V} f_{(i,j),u} = \sum_{p \in P : (i,j) \in p} w_{p} \leq f_{e}^{\text{min}} \cdot \text{Cap}
\]

by definition of feasibility of a line plan, i.e., constraint (11) is satisfied. Since the \( w_{p} \) correspond to paths in the PTN the flow conservation constraints (12) and (13) are also satisfied. By setting

\[
\text{dur} = \left\lceil \sum_{e \in E} 2f_{e}^{\text{min}} (L^{\text{drive}} + L^{\text{wait}}) \right\rceil
\]

we finally have constructed a feasible solution to Model 1.

We now show that the objective function value of the constructed solution is better than \( g(\mathcal{L}, \mathcal{R}) = c_{\text{time}} \cdot \text{dur}(\mathcal{L}, \mathcal{R}) + c_{\text{length}} \cdot \text{length}(\mathcal{L}, \mathcal{R}) \).

We first consider \( \text{length}(\mathcal{L}, \mathcal{R}) \): We know that for the constructed solution it holds that \( f_{e}^{\text{min}} = |\{ l \in \mathcal{L} | e \in l \}| \), hence

\[
\text{length}(\mathcal{L}, \mathcal{R}) \geq \sum_{l', e \in l'} \text{length}_{l'} = \sum_{l \in \mathcal{L}} \sum_{e \in l} 2\text{length}_{e} \geq \sum_{e \in E} 2\text{length}_{e} f_{e}^{\text{min}}.
\]
For \( \text{dur}(L, R) \) we calculate

\[
\text{dur}(L, R) = \sum_{r \in R} \text{dur}_r = \sum_{r \in R} \left[ \sum_{l' \in r} (\text{dur}_{l'} + L_{\text{turn}}) \right]_T \geq \sum_{r \in R} \sum_{l' \in r} (\text{dur}_{l'} + L_{\text{turn}})_T
\]

\[
= \left( \sum_{r \in R} \sum_{l' \in r} \left( (|l| - 1)L_{\text{wait}} + L_{\text{turn}} + \sum_{e \in l'} L_{\text{drive}} \right) \right)_T
\]

\[
\geq \left[ \sum_{l \in L} 2 \left( \frac{L_{\text{turn}} - L_{\text{wait}}}{\geq 0} + \sum_{e \in l} (L_{\text{drive}} + L_{\text{wait}}) \right) \right]_T
\]

\[
\geq \sum_{e \in E} 2f_{e}^{\text{min}} (L_{e} \text{drive} + L_{e} \text{wait}) = \text{dur} \cdot T.
\]

Overall it holds that

\[
g(L, R) = c_{\text{time}} \text{dur}(L, R) + c_{\text{length}} \text{length}(L, R) \geq c_{\text{time}} \text{dur} \cdot T + c_{\text{length}} \sum_{e \in E} 2\text{length}_e \cdot f_{e}^{\text{min}}.
\]

Thus every feasible solution to (cost-opt LTS) can be transformed to a solution for Model 1 whose objective is smaller than \( g(L, R) \). Hence, the optimal objective function value of Model 1 yields a lower bound to (cost-opt LTS).

\( \square \)

**Proof of Lemma 4.** We start with constraint (12), i.e.,

\[
\sum_{i \in V: \{i, v\} \in E} f_{i, u} = W_{uv} + \sum_{i \in V: \{v, i\} \in E} f_{(v, i), u} \quad \forall u \in V \forall v \in V \setminus \{u\}
\]

and argue that for any \( u \in X \) it holds that

\[
\sum_{v \in Y} \sum_{i \in V: \{i, v\} \in E} f_{i, u} = \sum_{v \in Y} \left( W_{uv} + \sum_{i \in V: \{v, i\} \in E} f_{(v, i), u} \right)
\]

\[
\sum_{v \in Y} \left( \sum_{i \in X: \{i, v\} \in E} f_{i, u} + \sum_{i \in Y: \{v, i\} \in E} f_{(v, i), u} \right)
\]

\[
= \sum_{v \in Y} \left( W_{uv} + \sum_{i \in X: \{v, i\} \in E} f_{(v, i), u} + \sum_{i \in Y: \{v, i\} \in E} f_{(v, i), u} \right)
\]

\[
\sum_{v \in Y} \left( \sum_{i \in X: \{i, v\} \in E} f_{i, u} = \sum_{v \in Y} \left( W_{uv} + \sum_{i \in X: \{v, i\} \in E} f_{(v, i), u} \right)
\]

\[
\sum_{v \in Y: \{v, i\} \in E_{\text{out}}} f_{(v, i), u} = \sum_{v \in Y} W_{uv} + \sum_{v \in Y: \{v, i\} \in E_{\text{out}}} f_{(v, i), u}
\]

18
Hence we can conclude
\[ \sum_{i \in X, v \in Y} f_{(i,v)} \geq \sum_{u \in X} W_{uv} \quad \forall u \in X. \tag{36} \]

Thus we get that
\[ \text{Cap} \cdot \sum_{e \in E_{\text{cut}}} f^\text{min}_e \geq \sum_{i \in X, v \in Y} f_{(i,v),u} \geq \sum_{u \in X} \sum_{v \in Y} W_{uv}. \]

\[ \sum_{u \in X} \sum_{v \in Y} W_{uv}. \]

\[ (L, \mathcal{R}) \] to (cost-opt LTS) as follows: We define the line plan \( \mathcal{L} \) that contains for each edge \( e \in E \) exactly \( f^\text{min}_e \) lines containing exactly this one edge \( e \), i.e., \( \mathcal{L} := \{e^1, \ldots, e^{f^\text{min}_e} : e \in E\} \).

Since \( f^\text{min}_e = \{|l \in \mathcal{L}[e] \mid e \} \) and \( f^\text{min}_e \) admits a feasible load, the line plan \( \mathcal{L} \) is feasible.

For this line plan we now generate a vehicle schedule \( \mathcal{R} \) that consists of only one large route. To this end, we consider the resulting set of directed lines \( \mathcal{L}' \)

\[ \mathcal{L}' = \{(i,j)^1, \ldots, (i,j)^{f^\text{min}_e}, (j,i)^1, \ldots, (j,i)^{f^\text{min}_e} : e = (i,j) \in E\} \]

which contains \( f^\text{min}_e \) copies of both directions of every edge \( e \in E \). This is a set of directed edges which creates a directed multigraph \((V, \mathcal{L}')\). Due to the assumption in the theorem, this graph is strongly connected and every node in \((V, \mathcal{L}')\) has the same indegree as outdegree. Hence we can find an Eulerian Cycle on it (see e.g. [11]). This means that we can form a route containing all directed lines \( r = (l'_1, \ldots, l'_k) \) (with \(|r| = |\mathcal{L}'|\) such that \( \text{length}_{l_i, l_{i+1}} = 0 \) and \( \text{time}_{l_i, l_{i+1}} = 0 \). So we set the vehicle schedule \( \mathcal{R} = \{r\} \) to contain exactly this route \( r \).

We hence have constructed some solution \((L, \mathcal{R})\) to (cost-opt LTS) with

\[ \text{length}(L, \mathcal{R}) = \sum_{l \in \mathcal{L}'} \text{length}_l + \sum_{r=(l'_1, \ldots, l'_k) \in \mathcal{R}} \sum_{i=1}^{k_r} \text{length}_{l_i, l_{i+1}} \]

\[ = \sum_{l \in \mathcal{L}} 2 \cdot \text{length}_l \sum_{f_{\text{min}} = \{|e \in \mathcal{L}[e] \mid e \}} \sum_{e \in E} 2 \text{length}_e f_{\text{min}}. \]
and
\[
dur(L, R) = \sum_{r \in R} \left( \sum_{\ell \in L} (\text{dur}_\ell + L^{\text{turn}}) \right)_T,
\]
\[
f_{\min} = \left\{ e \in E \mid f^e = 1 \right\} \sum_{e \in E} 2f^e (L^e + L^{\text{turn}})_T = \text{dur} \cdot T.
\]

Hence, for every solution to Model 1 we can construct a solution \((L, R)\) to (cost-opt LTS) such that \(g(L, R) = c_{\text{time}} \cdot \text{dur} \cdot T + c_{\text{length}} \sum_{e \in E} 2L^e \cdot f^e_{\min}\).

Together with Theorem 3 \((L, R)\) is optimal for (cost-opt LTS) and hence Model 1 has the same objective value as (cost-opt LTS).

\[\square\]

**Proof of Theorem 10.** Let \((L, R)\) be some feasible solution to (cost-opt LTS). Then we know that we can set \(f^e_{\min} = |\ell \in L \mid e \in \ell|\) (and \(f_{e,u}\) accordingly) as in the proof of Theorem 3 to some feasible flow which satisfies (16). Furthermore we can enumerate all lines with some bijective mapping \(\varphi : L \rightarrow \|L\|\) such that \(x_{e,\varphi(\ell)} = 1\) iff \(e \in \ell\) for all \(\ell \in L\) and also \(y_{s,e}(l) = 1\) iff \(s \in e\) for some \(e \in \ell\) and \(z_i = 1\) for all \(i \in \|L\|\) and 0 else. Since \(L\) was some feasible line plan all lines are simple paths and hence also constraints (17) to (22) are fulfilled. Now for the objective function it holds that

\[
\text{length}(L, R) = \sum_{\ell \in L} \text{length}_\ell + \sum_{r = (l_1, \ldots, l_r) \in R} \sum_{i=1}^{k_r-1} \text{length}_{l_i, l_{i+1}}
\]
\[
\geq \sum_{\ell \in L} \sum_{e \in \ell} 2\text{length}_e = \sum_{\ell \in L} \sum_{e \in E} 2x_{e,\varphi(\ell)} \text{length}_e = \sum_{l=1}^L \sum_{e \in E} 2x_{e,l} \text{length}_e.
\]

For the duration we get

\[
dur(L, R) = \sum_{r = (l_1, \ldots, l_r) \in R} \left[ \sum_{i=1}^{k_r} (\text{dur}_{l_i} + \text{dur}_{l_i, l_{i+1}}) \right]_T \geq \left[ \sum_{r \in R} \sum_{\ell' \in r} (\text{dur}_{\ell'} + \text{L}^{\text{turn}}) \right]_T
\]
\[
= \left(4\right) \left[ \sum_{\ell' \in L} \left( L^{\text{turn}} - L^{\text{wait}} + \sum_{e \in \ell'} (L^e + L^{\text{wait}}) \right) \right]_T
\]
\[
= \left[ \sum_{l=1}^L \left( 2z_l (L^{\text{turn}} - L^{\text{wait}}) + \sum_{e \in E} 2(L^e + L^{\text{wait}}) \cdot x_{e,l} \right) \right]_T \geq \text{dur} \cdot T.
\]

Hence, by finally setting

\[
dur = \left[ \sum_{l=1}^L \left( 2z_l (L^{\text{turn}} - L^{\text{wait}}) + \sum_{e \in E} 2(L^e + L^{\text{wait}}) \cdot x_{e,l} \right) \right]_T.
\]

20
we conclude that from any feasible solution \((L, R)\) to (cost-opt LTS) we can construct some feasible solution to Model 2 such that

\[
g(L, R) \geq c_{\text{time}} \cdot \text{dur} \cdot T + c_{\text{length}} \sum_{i=1}^{L} \sum_{e \in E} 2x_{e,i} \cdot \text{length}_e,
\]

which means that the objective function value of Model 2 is a lower bound to (cost-opt LTS).

On the other hand every feasible solution to Model 2 is a feasible solution to Model 1. This can be seen by setting the three types of variables, \(f_{\text{min}}, f_{e,u,} \) and \(\text{dur} \), that are contained in both models, to be the same. Hence constraints (11) - (13) are satisfied, and also (10) is satisfied since

\[
\text{dur} \cdot T \geq L \sum_{l=1}^{L} \left( 2z(l) \frac{L_{\text{turn}} - L_{\text{wait}}}{\geq 0} + \sum_{e \in E} 2(L_{e,\text{drive}} + L_{e,\text{wait}}) \cdot x_{e,l} \right) \geq \sum_{e \in E} 2f_{\text{min}}(L_{e,\text{drive}} + L_{e,\text{wait}}).
\]

For the objective functions it additionally holds that

\[
\sum_{l=1}^{L} \sum_{e \in E} 2x_{e,i} \cdot \text{length}_e = \sum_{e \in E} 2f_{\text{min}} \cdot \text{length}_e.
\]

This means that every solution to Model 2 can be projected to a solution of Model 1 with smaller objective value in Model 1, meaning that Model 2 is an upper bound to Model 1.

**Proof of Theorem 13.** Let \(L, R\) be some line-pure feasible solution to (cost-opt LTS). For the objective value of \((L, R)\) we know that

\[
\text{length}(L, R) = \sum_{r=(l_1, \ldots, l_k) \in R} \sum_{i=1}^{k_r} \text{length}_{l_i} + \text{length}_{l_i, l_{i+1}} = \sum_{l \in L} \sum_{e \in E} 2 \text{length}_e = \sum_{l \in L} \sum_{e \in E} \text{length}_e,
\]

and that

\[
\text{dur}(L, R) = \sum_{r \in R} \sum_{l' \in r} (\text{dur}_{l'} + L_{\text{turn}}) = \sum_{l \in L} \left( 2(\text{dur}_l + L_{\text{turn}}) \right) T = \sum_{l \in L} \left( 2(L_{\text{turn}} - L_{\text{wait}}) + \sum_{e \in E : e \in l} 2(L_{e,\text{drive}} + L_{e,\text{wait}}) \right) T.
\]

We can extend the line plan \(L\) to some feasible solution to Model 2* by again defining a bijective mapping \(\varphi : L \rightarrow \|L\|\) such that \(x_{e,\varphi(l)} = 1\) if \(e \in l\) for \(l \in L\) for all \(e \in E\). Analogously a solution \(x_{e,l}\) can be transformed into some feasible line plan \(L\) by defining a line \(l\) to contain exactly all edges \(e \in E\) if \(x_{e,l} = 1\). Thus there exists a bijection between the set of feasible solutions between (cost-opt LTS) and Model 2* as well as the same objective function for both problems since

\[
\sum_{l \in L} \sum_{e \in E} 2 \text{length}_e = \sum_{l \in L} \sum_{e \in E} 2x_{e,\varphi(l)} \cdot \text{length}_e = \sum_{l=1}^{L} \sum_{e \in E} 2x_{e,i} \cdot \text{length}_e
\]
\[
\sum_{l \in L} \left[ 2(L_{\text{turn}} - L_{\text{wait}}) + \sum_{e \in E : e \in l} 2(L_e^{\text{drive}} + L_e^{\text{wait}}) \right]_T = \sum_{l=1}^L \left[ 2z_l(L_{\text{turn}} - L_{\text{wait}}) + \sum_{e \in E} 2x_{e,l} \text{length}_e(L_e^{\text{drive}} + L_e^{\text{wait}}) \right]_T = \sum_{l=1}^L d_l.
\]

Hence their optimal objective values coincide. \(\square\)

**References**


<table>
<thead>
<tr>
<th>Number</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018 - 1</td>
<td>Brimberg, J.; Schöbel, A.</td>
<td>When closest is not always the best: The distributed p-median problem</td>
</tr>
<tr>
<td>2018 - 3</td>
<td>Schmidt, M., Schöbel, A., Thom, L.</td>
<td>Min-ordering and max-ordering scalarization methods for multi-objective robust optimization</td>
</tr>
<tr>
<td>2018 - 4</td>
<td>Krüger, C., Castellani, F., Gelledmann, J., Schöbel, A.</td>
<td>Peat and Pots: An application of robust multi-objective optimization to a mixing problem in agriculture</td>
</tr>
<tr>
<td>2018 - 5</td>
<td>Krüger, C.</td>
<td>Peat and pots: Analysis of robust solutions for a biobjective problem in agriculture</td>
</tr>
<tr>
<td>2018 - 7</td>
<td>Lehrenfeld, C., Olshanskii, M.A.</td>
<td>An Eulerian Finite Element Method for PDEs in Time-dependent Domains</td>
</tr>
<tr>
<td>2018 - 9</td>
<td>Lederer, P.L. Lehrenfeld, C. Schöberl, J.</td>
<td>HYBRID DISCONTINUOUS GALERKIN METHODS WITH RELAXED H(DIV)-CONFORMITY FOR INCOMPRESSIBLE FLOWS. PART II</td>
</tr>
<tr>
<td>Number</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------</td>
<td>--------------------------------------------------------------</td>
</tr>
<tr>
<td>2018 - 10</td>
<td>Lehrenfeld, C.; Rave, S.</td>
<td>Mass Conservative Reduced Order Modeling of a Free Boundary Osmotic Cell Swelling Problem</td>
</tr>
<tr>
<td>2018 - 12</td>
<td>M. Lübbecke, C. Puchert, P. Schiewe, A. Schöbel</td>
<td>Integrating line planning, timetabling and vehicle scheduling</td>
</tr>
<tr>
<td>2018 - 13</td>
<td>M. Friedrich, M. Müller-Hannemann, R. Rückert, A. Schiewe, A. Schöbel</td>
<td>Robustness as a Third Dimension for Evaluating Public Transport Plans</td>
</tr>
<tr>
<td>2018 - 14</td>
<td>J. Pätzold, A. Schiewe, A. Schöbel</td>
<td>Cost-Minimal Public Transport Planning</td>
</tr>
</tbody>
</table>