Integrated Line Planning and Vehicle Scheduling for Public Transport*

Philine Schiewe
Technical University Kaiserslautern
Kaiserslautern, Germany
p.schiewe@mathematik.uni-kl.de

Moritz Stinzendörfer
Technical University Kaiserslautern
Kaiserslautern, Germany
stinzendoerfer@mathematik.uni-kl.de

ABSTRACT

In public transport planning, the operational costs are mainly determined by the vehicle schedule. However, in the traditionally sequential planning approach, vehicle scheduling is one of the last problems that is considered. We therefore propose an integrated formulation for line planning and vehicle scheduling problem, which brings an appropriate approximation of the operational costs into the first planning stages. We model the integrated problem as a mixed-integer program and propose a heuristic solution approach. Both approaches are tested on close-to-real-world data sets from the open source software framework LinTim.

1 INTRODUCTION

With growing urban areas, public transport can play an important role in achieving sustainable mobility by consolidating demand and reducing traffic. For being a viable alternative to individual motorized transport modes, public transport has to be attractive for the passengers, e.g., by offering frequent service and short travel times, and economically viable for the operator. The tasks of finding a public transport supply that is attractive for both passengers and operators, is intricate and comprises various subproblems that are closely interrelated. Some of the most important subproblems are line planning, timetabling and vehicle scheduling, three problems that are traditionally solved sequentially and in that order, see [5, 6]. All three problems are extensively studied, see e.g., [2, 9, 15]. In recent years, the optimization potential arising when several subproblems are considered in an integrated manner has been under research, e.g., in [8, 14].

One important aspect of integration is summarized in the concept of the eigenmodel [16], i.e., to change the order in which the subproblems are considered. In this paper, we combine the idea of the eigenmodel and integrate several subproblems by considering line planning and vehicle scheduling simultaneously in order to minimize the operational costs. Therefore, we construct lines, i.e., paths in the infrastructure network that have to be operated by one vehicle end-to-end, and arrange them into vehicle routes. By refraining from using a fixed line pool, we allow for a larger solution space. Additionally, we provide the possibility to create vehicle schedules for a limited number of vehicles and vehicles with a limited range such as electric vehicles. We provide a mixed-integer programming formulation for a restricted version of the integrated line planning and vehicle scheduling problem and propose a fast heuristic to solve close-to real-world instances. Additionally, our approach can be used for creating line pools to serve as a basis for further planning approaches.

Similar approaches in the literature include the transit route network design problem which often comprises determining routes with corresponding frequencies. Here, many of the approaches are (meta-)heuristics or designed for very specific networks, see [7] for an overview. An integrated model for line planning, timetabling and vehicle scheduling is proposed in [12] but due to its size, it can only be used for very small instances. In [11], a heuristic line pool generation procedure from [4] is adapted to generate lines which allow for cost-efficient vehicle schedules for the case of an undirected public transport network and when no depot is considered. A heuristic sequential approach for first creating a vehicle schedule and then lines is presented in [10] where the goal is to maximize the attractiveness for the passengers.

The remainder of the paper is structured as follows. Section 2 gives an overview on the classical sequential approach to public transport planning. In Section 3, we present our model for the integrated line planning and vehicle routing problem as well as a short analysis. The restricted version with the corresponding mixed-integer program and a heuristic solution approach are presented in Section 4 and experimentally evaluated in Section 5. The paper is concluded in Section 6.

2 SEQUENTIAL PLANNING PROCESS

As input we assume that a *public transport network (PTN)*, i.e., a digraph (V, A), is given. Here, the nodes V represent stations and the arcs A direct connections between them, such as roads or tracks. As we optimize the operational costs of the public transport supply, we assume that the passengers' demand is given and their routes in the PTN are fixed. For a simple path P in (V, A) we denote by $A(P)$ the arcs of P and by $\alpha(P), \omega(P)$ the first and last node of P, respectively. Similarly, we call for arcs $a = (u, v)$ the incident nodes $\alpha(a) = u$ and $\omega(a) = v$.

For the classical *line planning* problem as described, e.g., in [15], lower and upper frequency bounds $f_a^{\min} \le f_a^{\max}$, $a \in A$, are given which guarantee that passengers can travel on their routes while safety restrictions are respected. The goal is to find a set of lines, i.e., paths in the PTN which adhere to the given bounds.

Definition 2.1. Let a public transport network (V, A) with upper and lower frequency bounds $f_a^{\min} \le f_a^{\max}$, $a \in A$, and a line pool, i.e., a set of possible lines $\mathcal{L}^0$ be given. The *line planning problem* is to find a subset $\mathcal{L} \subset \mathcal{L}^0$ with frequencies $f(l) \in \mathbb{N}_{>0}$, $l \in \mathcal{L}$, such

---

*This work is partially funded by DFG under grants SCHO 1140/8-2.

that
$$f_a^{\min} \leq \sum_{l \in \mathcal{L}: a \in A(l)} f(l) \leq f_a^{\max}.$$
We call $(\mathcal{L}, f)$ the corresponding *line concept*.

Obviously, the line pool $\mathcal{L}^0$ has a large influence on the line concept and the quality of the public transport supply, especially if it is too restrictive, see [4]. For the remainder of the paper, we assume that lines have to be simple paths and start and end at terminal stations $\bar{V} \subset V$ but do not impose further restrictions.

During the *timetabling* stage, (periodic) times are assigned to arrivals and departures at each station of each line, i.e., the events are repeated periodically with a fixed period length T. The dependencies between events are modeled as activities, which represent, e.g., vehicles driving or dwelling, and transfers of passengers. Each activity imposes a lower and an upper bound on the difference of the corresponding event times. The resulting problem is the well-known NP-hard periodic event scheduling problem, see [17].

However, for the remainder of the paper, we do not consider headway constraints and assume that transfers between lines have no restricting upper bound. Therefore, a feasible timetable can be constructed easily by considering each line separately.

For a given timetable, a vehicle schedule is constructed with the objective of minimizing the operational costs. We consider a periodic vehicle scheduling problem where all vehicles start from the same depot dep.

Definition 2.2. Let a public transport network (V, A), a line concept $(\mathcal{L}, f)$ and a periodic timetable π with period length T be given. A *periodic vehicle schedule* is a multiset of paths or *(vehicle) routes* $\mathcal{R}$ such that

- each route $r \in \mathcal{R}$ is a concatenation of pairwise disjoint lines $l \in \mathcal{L}$ and
- each line $l \in \mathcal{L}$ is contained in exactly $f(l)$ vehicle routes.

The distance covered by a vehicle route $r \in \mathcal{R}$ is
$$\begin{aligned}\text{dist}(r) = &\sum_{l \in r} \sum_{a \in A(l)} \text{dist}(a) + \text{dist}(\text{dep}, \alpha(r)) + \text{dist}(\omega(r), \text{dep}) \\ &+ \sum_{\substack{l, l' \text{consecutive} \\ \text{lines in } r}} \text{dist}(\omega(l), \alpha(l'))\end{aligned}$$
where $\text{dist}(a)$ is the distance from $\alpha(a)$ to $\omega(a)$. We call
$$\text{dist}(\mathcal{R}) = \sum_{r \in \mathcal{R}} \text{dist}(r)$$
the distance of the vehicle schedule.

Similarly, the duration of a vehicle route $r \in \mathcal{R}$ for timetable π is
$$\begin{aligned}\text{dur}(r, \pi) := &\sum_{l \in r} \sum_{a \in A(l)} \text{dur}(a, l, \pi) \\ &+ \text{dur}(\text{dep}, \alpha(r)) + \text{dur}(\omega(r), \text{dep}) + \text{dur}(\text{dep}) \\ &+ \sum_{\substack{l, l' \text{consecutive} \\ \text{lines in } r}} \text{dur}(\omega(l), \alpha(l'))\end{aligned}$$
where $\text{dur}(a, l, \pi)$ is the duration between the time scheduled for the departure of line l at $\alpha(a)$ and at $\omega(a)$ and $\text{dur}(\omega(l), \alpha(l'))$ the duration of relocating between lines l and l'. Note that the duration for getting from the depot to the first station of r and from the last station of r to the depot does not depend on the timetable. The minimal turn-over time at the depot is represented by dur(dep). We call
$$\text{dur}(\mathcal{R}, \pi) = \sum_{r \in \mathcal{R}} \text{dur}(r, \pi)$$
the duration of the vehicle schedule.

For each route $r \in \mathcal{R}$, the number of vehicles needed to operate it is
$$\text{veh}(r, \pi) = \left\lceil \frac{\text{dur}(r, \pi)}{T} \right\rceil.$$
The total number of vehicles needed to operate $\mathcal{R}$ is
$$\text{veh}(\mathcal{R}, \pi) = \sum_{r \in \mathcal{R}} \text{veh}(r, \pi).$$

For parameters $(\lambda, \mu, \kappa) \in \mathbb{R}^3_{\geq 0}$, we define the *costs* of vehicle schedule $\mathcal{R}$ for timetable π as
$$\text{cost}(\mathcal{R}, \pi) := \lambda \cdot \text{dist}(\mathcal{R}) + \mu \cdot \text{dur}(\mathcal{R}, \pi) + \kappa \cdot \text{veh}(\mathcal{R}, \pi).$$

In the basic definition, there are no restrictions on the vehicle routes. However, restricting the duration of a vehicle route might be important, especially if electric vehicles are considered. In this case, restricting the duration of a route according to the battery capacity and choosing dur(dep) such that the battery can be reloaded guarantees that the vehicle schedule can be operated by electric vehicles.

3 MODELING THE LINE PLANNING AND VEHICLE SCHEDULING PROBLEM

As in the sequential planning process a vehicle schedule is constructed for a given line plan and timetable, we have to adapt our notation for defining the integrated problem.

Definition 3.1. Let a public transport network (V, A) with minimal and maximal frequency $f_a^{\min} \leq f_a^{\max}$, $a \in A$, arc duration $\text{dur}(a)$, $a \in A$, minimal turn-over time dur(dep) at the depot and set of terminal stations $\bar{V} \subset V$ as well as a maximal line duration K and maximal number of routes R be given. The *line planning and vehicle scheduling problem* (LVP) is to find a multiset of simple paths $\mathcal{R}$, i.e., vehicle routes, such that

- for all arcs $a \in A$
$$f_a^{\min} \leq |\{r \in \mathcal{R} : a \in A(r)\}| \leq f_a^{\max},$$
i.e., each arc a is contained in at least $f_a^{\min}$ and at most $f_a^{\max}$ vehicle routes,
- for all routes $r \in \mathcal{R}$
$$\begin{aligned}\text{dur}(r) := &\sum_{a \in A(r)} \text{dur}(a) + \text{dur}(\text{dep}) \\ &+ \text{dur}(\text{dep}, \alpha(r)) + \text{dur}(\omega(r), \text{dep}) \\ \leq &K,\end{aligned}$$
i.e., the duration of each route r does not exceed K,
- the number of vehicle routes $|\mathcal{R}| \leq R$,
- $\alpha(r), \omega(r) \in \bar{V}$, i.e., the first and the last node of each vehicle route r are contained in the set of terminal stations and
- for parameter set $(\lambda, \mu, \kappa) \in \mathbb{R}^3_{\geq 0}$ the approximated costs
$$\text{cost}(\mathcal{R}) := \lambda \cdot \text{dist}(\mathcal{R}) + \mu \cdot \sum_{r \in \mathcal{R}} \text{dur}(r) + \kappa \cdot \sum_{r \in \mathcal{R}} \left\lceil \frac{\text{dur}(r)}{T} \right\rceil$$

are minimized.

The corresponding set of lines $\mathcal{L}$ consists of the paths from $\mathcal{R}$ where the multiplicity of $r \in \mathcal{R}$ corresponds to the frequency $f(r)$.

Note that for a feasible solution of (LVP) the vehicle schedule $\mathcal{R}$ and the line concept $(\mathcal{L}, f)$ are feasible by construction. As each vehicle route consists of only one line, we do not have to consider relocating between lines.

Unfortunately, (LVP) is NP-hard even when no restrictions on K and R are imposed.

Theorem 3.2. *(LVP) is NP-hard, even if* $K = R = \infty$ *and* $\bar{V} = V$.

Proof. In this setting, (LVP) is equivalent to finding a costs-minimal line concept from the set of all simple paths. Setting $\lambda = \kappa = 0$ and $\text{dur}(\text{dep}, v) = \text{dur}(v, \text{dep}) = 0$ for all $v \in V$ leads to the same cost structure as in [4] where this problem is shown to be NP-hard. □

The maximal route duration K and the maximal number of routes R can be used to influence the structure of the resulting line concept and vehicle schedule. R restricts the number of vehicle routes and therefore the number of lines such that there are not too many - possible very short - lines which would be undesirable from a passengers' point of view. K restricts the duration of a vehicle route, which is beneficial from a robustness viewpoint as long vehicle routes tend to propagate delays.

Together, they can also be used to bound the number of operated vehicles.

Lemma 3.3. *Let π be a feasible timetable with* $\text{dur}(a) \geq \text{dur}(a, l, \pi)$ *for all* $a \in A(l), l \in \mathcal{L}$. *Then*

$$R \cdot \left\lceil \frac{K}{T} \right\rceil \geq |\mathcal{R}| \cdot \left\lceil \frac{K}{T} \right\rceil \geq \sum_{r \in \mathcal{R}} \left\lceil \frac{\text{dur}(r)}{T} \right\rceil \geq \text{veh}(\mathcal{R}, \pi).$$

Proof. The first inequality follows directly from $|\mathcal{R}| \leq R$, the second from $\text{dur}(r) \leq K$ and the last from $\text{dur}(a) \geq \text{dur}(a, l, \pi)$ for all $a \in A(l), l \in \mathcal{L}$. □

Note that we can choose $\text{dur}(a) \geq \text{dur}(a, l, \pi)$ a priori when the construction of the bounds for timetabling is known. Additionally, we can bound the costs $\text{cost}(\mathcal{R}, \pi)$ of the vehicle schedule for a feasible timetable π.

Lemma 3.4. *Let π be a feasible timetable with* $\text{dur}(a) \geq \text{dur}(a, l, \pi)$ *for all* $a \in A(l), l \in \mathcal{L}$. *Then*

$$\text{cost}(\mathcal{R}) \geq \text{cost}(\mathcal{R}, \pi).$$

Proof. This follows directly from Lemma 3.3, as $\text{dist}(\mathcal{R})$ is independent of the timetable and $\text{dur}(a) \geq \text{dur}(a, l, \pi)$ for all $a \in A(l)$, $l \in \mathcal{L}$. □

4 SOLVING THE RESTRICTED LINE PLANNING AND VEHICLE SCHEDULING PROBLEM

To solve the integrated line planning and vehicle scheduling problem, we consider the following restriction (rLVP): For each arc $a \in A$, we suppose that an arc frequency $f(a) \in \{f_a^{\min}, \ldots, f_a^{\max}\}$ is given and we have to find a solution to (LVP) such that

- for all arcs $a \in A$
$$f(a) = |\{r \in \mathcal{R} \colon a \in r\}|,$$
i.e., each arc a is contained in exactly $f(a)$ vehicle routes and
- for parameter set $(\lambda, \mu, \kappa) \in \mathbb{R}^3_{\geq 0}$ the approximated costs
$$\tilde{\text{cost}}(\mathcal{R}) := \lambda \cdot \text{dist}(\mathcal{R}) + \mu \cdot \sum_{r \in \mathcal{R}} \text{dur}(r) + \kappa \cdot |\mathcal{R}| \cdot \left\lceil \frac{K}{T} \right\rceil$$
are minimized.

From Lemma 3.3, we know that the optimal objective value of (rLVP) is an upper bound on the optimal objective value of (LVP).

Corollary 4.1. *Let $\mathcal{R}$ be an optimal solution of* (LVP) *and $\bar{\mathcal{R}}$ an optimal solution of* (rLVP). *Then*

$$\tilde{\text{cost}}(\bar{\mathcal{R}}) \geq \text{cost}(\bar{\mathcal{R}}) \geq \text{cost}(\mathcal{R}).$$

Proof. The first inequality follows directly from Lemma 3.3 while the second inequality follows as $\bar{\mathcal{R}}$ is a feasible solution of (LVP). □

4.1 Graph Construction for Vehicle Routing Formulation

We can model (rLVP) as a slightly modified capacitated vehicle routing problem on the following digraph $\tilde{G} = (\tilde{V}, \tilde{A})$. The idea is that the nodes $\tilde{V}$ represent the arcs of PTN (V, A). By setting the demand of each node in $\tilde{V}$ to the duration of the arc in A and the capacity of the vehicle routing problem to $K - \text{dur}(\text{dep})$, we get a direct correspondence between the vehicle routes in both graphs.

An example of the construction is given in Figure 1.

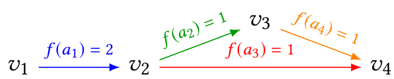
We set

$$\tilde{V} := \{v_a^i : a \in A, i \in \{1, ..., f(a)\}\} \cup \{\tilde{v}_0\},$$

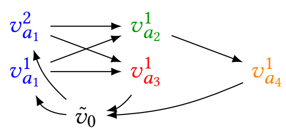
i.e., we create $|f(a)|$ nodes for each arc $a \in A$ in the public transport network and an artificial depot node $\tilde{v}_0$.

To indicate the corresponding arc $a = (v, w) \in A$ of the created nodes $\tilde{v} \in \tilde{V}$ in the vehicle routing graph, we use $\tilde{v}' := a$ if $\tilde{v} = v_a^i$ for $i \in \{1, ..., f(a)\}$.

For each node $v \in V$ and each incoming arc $a \in \delta^-(v)$ and outgoing arc $b \in \delta^+(v)$ in PTN (V, A) we create arcs (v_a^i, v_b^j), $i \in \{1, ..., f(a)\}, j \in \{1, ..., f(b)\}$, in the vehicle routing graph $\tilde{G}$.



(a) PTN (V, A)**.**



(b) Vehicle routing graph $\tilde{G} = (\tilde{V}, \tilde{A})$**.**

Figure 1: Transformation of the public transport network (V, A) **to the vehicle routing graph** $\tilde{G}$**. Terminal stations are** $\bar{V} = \{v_1, v_4\}$**.**

To ensure that each line begins and ends at a terminal station, we have arcs from the artificial depot node $\tilde{v}_0$ to each node $\tilde{v}$ whose corresponding left station $\alpha(\tilde{v}')$ is a terminal station and the other way round if the corresponding right station $\omega(\tilde{v}')$ is a terminal station. Formalized we have

$$\begin{aligned}\tilde{A} :=&\{(v_a^i, v_b^j): && a, b \in A, i \in \{1, \dots, f(a)\}, j \in \{1, \dots, f(b)\},\\ & && \omega(a) = v = \alpha(b) \text{ for } v \in V\}\\ &\cup \{(\tilde{v}_0, v_a^i): && a \in A, i \in \{1, \dots, f(a)\}, \alpha(a) = v \text{ for } v \in \bar{V}\}\\ &\cup \{(v_a^i, \tilde{v}_0): && a \in A, i \in \{1, \dots, f(a)\}, \omega(a) = v \text{ for } v \in \bar{V}\}.\end{aligned}$$

The demand of each node $\tilde{v} \in \tilde{V}$ is defined as $d(\tilde{v}) := \text{dur}(a)$ for $\tilde{v}' = a$, whereas the capacity of the vehicles is set to $C := K - \text{dur}(\text{dep})$ and the maximal number of vehicles is set to R.

For a feasible solution to the vehicle routing problem in $\tilde{G}$ we know

- there are at most R tours,
- the demand of all nodes in a tour does not exceed K and
- all nodes $\tilde{V}$ are covered by exactly one tour.

We can translate such a tour $(\tilde{v}_0, \tilde{v}_1, \dots, \tilde{v}_n, \tilde{v}_0)$ in $\tilde{G}$ back to a not necessarily simple path $(\tilde{v}_1', \dots, \tilde{v}_n')$ in (V, A). By construction, the resulting set of paths covers all arcs according to their frequencies $f(a)$, there are at most R paths and the duration of each path including the turn-over time does not exceed K.

Before we consider how simple paths can be constructed, we define the costs of arcs $\tilde{A}$ to correspond to $\bar{\text{cost}}$.

$$\text{cost}(v_a^i, v_b^j) = \lambda \cdot \text{dist}(a) + \mu \cdot \text{dur}(a),$$

if $v_a^i \neq \tilde{v}_0 \neq v_b^j$,

$$\text{cost}(\tilde{v}_0, v_b^j) = \lambda \cdot \text{dist}(\text{dep}, v) + \mu \cdot (\text{dur}(\text{dep}, v) + \text{dur}(\text{dep})) + \kappa \cdot \left\lceil \frac{K}{T} \right\rceil,$$

if $\alpha(b) = v$ and

$$\text{cost}(v_a^j, \tilde{v}_0) = \lambda \cdot \text{dist}(w, \text{dep}) + \mu \cdot \text{dur}(w, \text{dep}),$$

if $\omega(a) = w$.

4.2 MIP Formulation for Simple Lines

We formulate the mixed-integer program as a modified capacitated vehicle routing problem in (1)-(13).

The first part up to constraints (7) is equal to the formulation of the capacitated vehicle routing problem as in [18].

The variable $x_{\tilde{a}}$ indicates if the corresponding arc $\tilde{a} \in \tilde{A}$ is used and $u_{\tilde{v}}$ describes the summed up demands of the nodes on the corresponding tour starting at $\tilde{v}_0$ up to $\tilde{v} \in \tilde{V}$. Constraints (6) - (7) ensure that the capacity C is not exceeded. Constraints (4)-(5) ensure that no more than R vehicles are used.

By construction, we can have multiple nodes in the transformed vehicle routing graph corresponding to the same arc in the public transport network. As a consequence, it is possible to obtain non-simple paths after translating the vehicle routing solution back to a line concept, i.e., multiple identical arcs on the same line.

Therefore we adapt the previously mentioned capacity constraints to ensure that all lines are simple paths.

$$\min \sum_{\tilde{a} \in \tilde{A}} x_{\tilde{a}} \cdot \text{cost}(\tilde{a}) \tag{1}$$

$$\text{s.t.} \sum_{\tilde{a} \in \delta^-(\tilde{v})} x_{\tilde{a}} = 1 \qquad \tilde{v} \in \tilde{V} \setminus \{\tilde{v}_0\} \tag{2}$$

$$\sum_{\tilde{a} \in \delta^+(\tilde{v})} x_{\tilde{a}} = 1 \qquad \tilde{v} \in \tilde{V} \setminus \{\tilde{v}_0\} \tag{3}$$

$$\sum_{\tilde{a} \in \delta^-(\tilde{v}_0)} x_{\tilde{a}} \leq R \tag{4}$$

$$\sum_{\tilde{a} \in \delta^+(\tilde{v}_0)} x_{\tilde{a}} \leq R \tag{5}$$

$$d(\tilde{v}) \leq u_{\tilde{v}} \leq C \qquad \tilde{v} \in \tilde{V} \setminus \{\tilde{v}_0\} \tag{6}$$

$$u_{\tilde{v}} + d(\tilde{w}) - u_{\tilde{w}} \leq (1 - x_{(\tilde{v}, \tilde{w})}) \cdot (C + d(\tilde{v})) \qquad (\tilde{v}, \tilde{w}) \in \tilde{A} \tag{7}$$

$$u_{p,\tilde{v}}^a = 1 \qquad p \in A, \tilde{v} \in \tilde{V} \setminus \{\tilde{v}_0\}, p = \tilde{v}' \tag{8}$$

$$u_{p,\tilde{v}}^a - u_{p,\tilde{w}}^a \leq 1 - x_{(\tilde{v}, \tilde{w})} \qquad p \in A, (\tilde{v}, \tilde{w}) \in \tilde{A}, \tilde{v} \neq \tilde{v}_0 \neq \tilde{w} \tag{9}$$

$$u_{\tilde{w}', \tilde{v}}^a \leq 1 - x_{(\tilde{v}, \tilde{w})} \qquad (\tilde{v}, \tilde{w}) \in \tilde{A}, \tilde{v} \neq \tilde{v}_0 \neq \tilde{w} \tag{10}$$

$$x_{\tilde{a}} \in \{0, 1\} \qquad \tilde{a} \in \tilde{A} \tag{11}$$

$$u_{\tilde{v}} \geq 0 \qquad \tilde{v} \in \tilde{V} \setminus \{\tilde{v}_0\} \tag{12}$$

$$0 \leq u_{p,\tilde{v}}^a \leq 1 \qquad p \in A, \tilde{v} \in \tilde{V} \setminus \{\tilde{v}_0\} \tag{13}$$

We introduce the variable $u_{p,\tilde{v}}^a$ that indicates if the tour that contains node $\tilde{v} \in \tilde{V}$ also contains a node $\tilde{w} \in \tilde{V}$ with $\tilde{w}' = p \in A$, i.e., the arc $p \in A$ of the public transport network is already covered by the vehicle routing tour containing node $\tilde{v}$. Obviously, $u_{p,\tilde{v}}^a = 1$ if $p = \tilde{v}'$ (see constraints (8)).

If we use arc $(\tilde{v}, \tilde{w})$, i.e., $x_{(\tilde{v}, \tilde{w})} = 1$, we copy the value of $u_{p,\tilde{v}}^a$ to $u_{p,\tilde{w}}^a$ for all $p \in A$ in (9). Additionally, this tour may not have covered the public transport network arc $\tilde{w}' \in A$ before node $\tilde{w} \in \tilde{V}$, i.e., $u_{\tilde{w}', \tilde{v}}^a = 0$ and is ensured in constraints (10).

4.3 MIP Formulation for Elementary Lines

In a similar way, we can ensure that we only obtain elementary lines after translating the vehicle routing solution back to a line concept, i.e., we get lines with no repeating nodes.

Therefore, we use nearly the same constraints as (8) - (10) to exclude repeating source nodes (see constraints (14) - (17)) and target nodes (see constraints (18) - (21)) in the lines translated back from the vehicle routing tours.

Again, we have the variable $u_{p,\tilde{v}}^\alpha$ that indicates if the tour that contains node $\tilde{v} \in \tilde{V}$ also contains a node $\tilde{w} \in \tilde{V}$ with $\alpha(\tilde{w}') = p \in V$, i.e., the node $p \in V$ of the public transport network is already covered (as source node) by the vehicle routing tour containing node $\tilde{v}$. The same applies for the variables $u_{p,\tilde{v}}^\omega$ and the target nodes. By this construction, it is possible that a line begins and ends at the same node.

Note that if we ensure elementary lines, we do not have to ensure simple lines.

$$u^{\alpha}_{p,\tilde{v}} = 1 \qquad p \in V, \tilde{v} \in \tilde{V} \setminus \{\tilde{v}_0\}, p = \alpha(\tilde{v}') \quad (14)$$

$$u^{\alpha}_{p,\tilde{v}} - u^{\alpha}_{p,\tilde{w}} \leq 1 - x_{(\tilde{v},\tilde{w})} \qquad p \in V, (\tilde{v},\tilde{w}) \in \tilde{A}, \tilde{v} \neq \tilde{v}_0 \neq \tilde{w} \quad (15)$$

$$u^{\alpha}_{\alpha(\tilde{w}'),\tilde{v}} \leq 1 - x_{(\tilde{v},\tilde{w})} \qquad (\tilde{v},\tilde{w}) \in \tilde{A}, \tilde{v} \neq \tilde{v}_0 \neq \tilde{w} \quad (16)$$

$$0 \leq u^{\alpha}_{p,\tilde{v}} \leq 1 \qquad p \in V, \tilde{v} \in \tilde{V} \setminus \{\tilde{v}_0\} \quad (17)$$

$$u^{\omega}_{p,\tilde{v}} = 1 \qquad p \in V, \tilde{v} \in \tilde{V} \setminus \{\tilde{v}_0\}, p = \omega(\tilde{v}') \quad (18)$$

$$u^{\omega}_{p,\tilde{v}} - u^{\omega}_{p,\tilde{w}} \leq 1 - x_{(\tilde{v},\tilde{w})} \qquad p \in V, (\tilde{v},\tilde{w}) \in \tilde{A}, \tilde{v} \neq \tilde{v}_0 \neq \tilde{w} \quad (19)$$

$$u^{\omega}_{\omega(\tilde{w}'),\tilde{v}} \leq 1 - x_{(\tilde{v},\tilde{w})} \qquad (\tilde{v},\tilde{w}) \in \tilde{A}, \tilde{w} \neq \tilde{v}_0 \neq \tilde{w} \quad (20)$$

$$0 \leq u^{\omega}_{p,\tilde{v}} \leq 1 \qquad p \in V, \tilde{v} \in \tilde{V} \setminus \{\tilde{v}_0\} \quad (21)$$

4.4 Heuristic Solution Approach

In addition to solving the MIP directly, we can use modifications of know heuristics for the capacitated vehicle routing problem to solve (rLVP). In particular, we tested a modification of the savings algorithm by Clarke and Wright [3].

The algorithm is initialized with $|\tilde{V} - 1|$ tours $(\tilde{v}_0, \tilde{v}, \tilde{v}_0)$ for all $\tilde{v} \in \tilde{V}$. After that, the saving

$$s(\tilde{v}_i, \tilde{v}_j) = \text{cost}(\tilde{v}_i, \tilde{v}_0) + \text{cost}(\tilde{v}_0, \tilde{v}_j) - \text{cost}(\tilde{v}_i, \tilde{v}_j)$$

is calculated for all $\tilde{v}_i, \tilde{v}_j \in \tilde{V}$ with $\tilde{v}_i \neq \tilde{v}_0 \neq \tilde{v}_j$ and sorted in non-increasing order.

Now, the first unused saving $s(\tilde{v}_i, \tilde{v}_j)$ is taken and the tours T_i, T_j corresponding to node $\tilde{v}_i$ and $\tilde{v}_j$ are merged by removing arcs $(\tilde{v}_i, \tilde{v}_0), (\tilde{v}_0, \tilde{v}_j)$ and adding $(\tilde{v}_i, \tilde{v}_j)$, if the following conditions are fulfilled:

- $T_i \neq T_j$
- there exists arc $(\tilde{v}_i, \tilde{v}_0)$ in T_i and $(\tilde{v}_0, \tilde{v}_j)$ in T_j
- the summed up demand of both tours T_i and T_j does not exceed the capacity C.

Subsequently, the next unused saving is taken and this step is repeated until R tours are left.

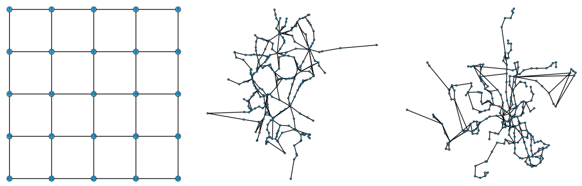
As in the MIP, by adding further conditions we can ensure simple and elementary lines, respectively. To guarantee the former, we only merge two tours T_i, T_j, if the corresponding public transport network arcs of the T_i are not the same as those in T_j.

In a similar way we proceed with the target and source nodes of the corresponding public transport network arcs of tour T_i and T_j to ensure elementary lines.

In some cases, the algorithm may terminate even though there are still more than R tours. This is due to the significantly increased number of conditions for merging two tours.

5 EXPERIMENTAL EVALUATION

We test the two solution approaches for (rLVP) on three data sets from the open source software framework LinTim [13], `grid`, `long-distance` and `goettingen`, see Figure 2, and compare them to the traditional sequential solution approach. The data sets represent an artificial benchmark instance, the long-distance train network in Germany and the bus network in Göttingen, respectively.



(a) **grid.** (b) **long-distance.** (c) **goettingen.**

Figure 2: Public transport networks for various data sets.

Table 1: Instance size and mean solver time (in seconds) and gap of the MIP formulation as well as the mean run time the heuristic approach in seconds. Note that for the MIP solution approach a time limit of 60 minutes is applied.

	PTN		Heu.	MIP	
Data Set	$\lvert V\rvert$	$\lvert A\rvert$	Time	Time	Gap
`grid`	25	80	0	3600	100%
`long-distance`	250	652	2	3600	99%
`goettingen`	257	548	12	3600	66%

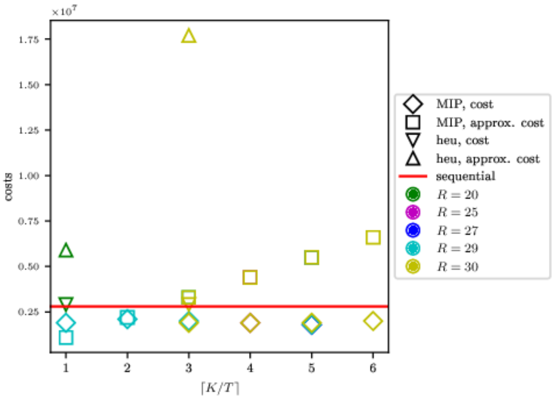
For each data set, we evaluate the approximated costs $\bar{\text{cost}}(\mathcal{R})$ and the actual costs $\text{cost}(\mathcal{R}, \pi)$ for various settings of K and R for the MIP formulation and the heuristic solution approach of (rLVP) for simple lines. For the MIP formulation, we use Gurobi 8.1.1 [1] and report the best solution found within a time limit of 60 minutes. These solutions are compared to the traditional approach of sequentially finding a line plan for a given pool, a timetable and a vehicle schedule.

Note that the runtime for the heuristic is considerably smaller than for the MIP-formulation, especially on the largest data set `goettingen`. The mean runtimes for all settings of K and R are reported in Table 1.

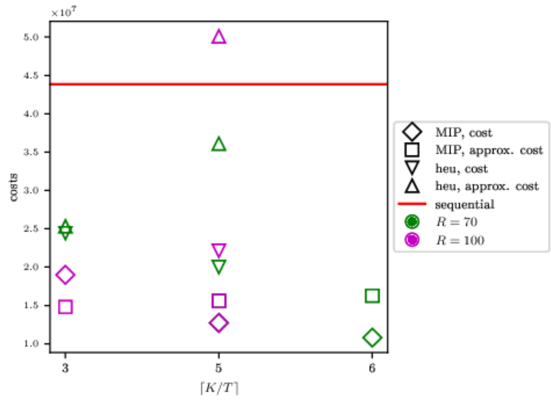
The results are depicted in Figure 3. We make the following observations:

- By using an integrated approach to line planning and vehicle scheduling, it is possible to find solutions that are much cheaper than by using the sequential solution approach, even when the duration of a vehicle route is restricted. For data sets `goettingen`, `grid` and `long-distance`, the costs can be reduced by up to 18%, 36% and 75%, respectively.
- For smaller K, i.e., for shorter vehicle routes, more routes R have to be allowed to find feasible solutions.
- With the heuristic approach, it is not possible to find feasible solutions for all given combinations of K and R.
- For larger K, the error of using $\bar{\text{cost}}(\mathcal{R})$ instead of $\text{cost}(\mathcal{R}, \pi)$ increases.
- While the costs $\text{cost}(\mathcal{R}, \pi)$ do increase slightly for smaller, i.e., more restrictive K, the difference is much smaller than suggest by $\bar{\text{cost}}(\mathcal{R})$.

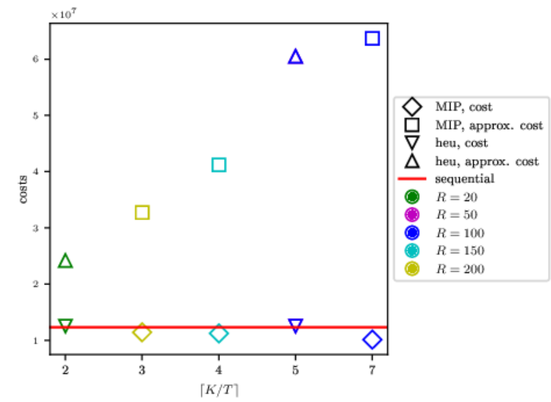
For data set `grid`, we additionally compared using simple paths for the vehicle route to restricting the routes to elementary paths. Here, the costs of the heuristic solutions increased by about 8%



(a) `grid`.



(b) `long-distance`.



(c) `goettingen`.

Figure 3: Costs for the different solution approaches for various data sets.

compared to the simple paths. Additionally, there are more infeasible combinations of K and R when the routes are restricted further. However, for some applications it might be necessary to restrict the set of lines to elementary instead of simple lines.

6 CONCLUSION AND OUTLOOK

In this paper, we modeled the integrated line planning and vehicle scheduling problem and proposed a solution approach for fixed arc frequencies. While the heuristic solution approach is especially fast and therefore can be used for larger data sets, the MIP-based solution approach outperforms the classical sequential solution approach even when restricting the duration of vehicle routes.

To better incorporate the passengers' perspective, it is also possible to use (rLVP) to generate a *line pool* instead of a line concept by choosing higher values for $f(a)$, $a \in A$. This allows for creating a larger set of potential lines which can be operated within the duration restriction K from which a line concept can be chosen separately. Evaluating these line pools in comparison to the approach from [4] and in combination with other integrated solution approaches such as integrated line planning, timetabling and passenger routing may lead to interesting new solution approaches.

REFERENCES

[1] 2019. Gurobi Optimizer. http://www.gurobi.com/. Gurobi Optimizer Version 8.1.1, Houston, Texas: Gurobi Optimization, Inc.
[2] S. Bunte and N. Kliewer. 2009. An overview on vehicle scheduling models. *Public Transport* 1, 4 (2009), 299–317.
[3] G. Clarke and J. W. Wright. 1964. Scheduling of vehicles from a central depot to a number of delivery points. *Operations research* 12, 4 (1964), 568–581.
[4] P. Gattermann, J. Harbering, and A. Schöbel. 2017. Line pool generation. *Public Transport* 9, 1-2 (2017), 7–32.
[5] V. Guihaire and J. Hao. 2008. Transit network design and scheduling: A global review. *Transportation Research Part A: Policy and Practice* 42, 10 (2008), 1251–1273.
[6] D. Huisman, L. Kroon, R. Lentink, and M. Vromans. 2005. Operations research in passenger railway transportation. *Statistica Neerlandica* 59, 4 (2005), 467–497.
[7] K. Kepaptsoglou and M. Karlaftis. 2009. Transit route network design problem. *Journal of transportation engineering* 135, 8 (2009), 491–505.
[8] R. N. van Lieshout. 2021. Integrated Periodic Timetabling and Vehicle Circulation Scheduling. *Transportation Science* 55, 3 (2021), 768–790.
[9] R. Lusby, J. Larsen, M. Ehrgott, and D. Ryan. 2011. Railway track allocation: models and methods. *OR spectrum* 33, 4 (2011), 843–883.
[10] M. Michaelis and A. Schöbel. 2009. Integrating Line Planning, Timetabling, and Vehicle Scheduling: A customer-oriented approach. *Public Transport* 1, 3 (2009), 211–232.
[11] J. Pätzold, A. Schiewe, P. Schiewe, and A. Schöbel. 2017. Look-Ahead Approaches for Integrated Planning in Public Transportation. In *17th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2017) (OpenAccess Series in Informatics (OASIcs))*, Vol. 59. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 17:1–17:16.
[12] J. Pätzold, A. Schiewe, and A. Schöbel. 2018. Cost-Minimal Public Transport Planning. In *18th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2018) (OpenAccess Series in Informatics (OASIcs))*, Vol. 65. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 8:1–8:22.
[13] A. Schiewe, S. Albert, P. Schiewe, A. Schöbel, and F. Spühler. 2020. *LinTim: An integrated environment for mathematical public transport optimization. Documentation for version 2020.12.* Technical Report. TU Kaiserslautern. https://nbn-resolving.org/urn:nbn:de:hbz:386-kluedo-62025
[14] P. Schiewe. 2020. *Integrated Optimization in Public Transport Planning.* Optimization and Its Applications, Vol. 160. Springer. https://doi.org/10.1007/978-3-030-46270-3
[15] A. Schöbel. 2012. Line planning in public transportation: models and methods. *OR spectrum* 34, 3 (2012), 491–510.
[16] A. Schöbel. 2017. An eigenmodel for iterative line planning, timetabling and vehicle scheduling in public transportation. *Transportation Research Part C: Emerging Technologies* 74 (2017), 348–365.
[17] P. Serafini and W. Ukovich. 1989. A mathematical model for periodic scheduling problems. *SIAM Journal on Discrete Mathematics* 2, 4 (1989), 550–581.
[18] P. Toth and D. Vigo. 2002. Models, relaxations and exact approaches for the capacitated vehicle routing problem. *Discrete Applied Mathematics* 123, 1-3 (2002), 487–512.