GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Institut für Numerische und Angewandte Mathematik

Periodic timetabling with integrated routing: An applicable approach

P. Schiewe, A. Schöbel

Nr. 17

Preprint-Serie des
Instituts für Numerische und Angewandte Mathematik
Lotzestr. 16-18
D - 37083 Göttingen

Periodic timetabling with integrated routing: An applicable approach*

Philine Schiewe[1] and Anita Schöbel[1]

[1]Institute for Numerical and Applied Mathematics,
University of Goettingen, Lotzestr. 16-18, 37083 Göttingen, Germany,
[p.schiewe, schoebel]@math.uni-goettingen.de

June 28, 2018

Abstract

Periodic timetabling is an important yet computationally challenging problem in public transportation planning. The usual objective when designing a timetable is to minimize passengers' travel time. However, in most approaches it is ignored that the routes of the passengers depend on the timetable, so handling their routing separately leads to timetables which are suboptimal for the passengers. This has recently been recognized, but integrating the passengers' routing in the optimization is computationally even harder than solving the classic periodic timetabling problem. In our paper we consider an integer programming model for integrating timetabling and passenger routing for which we develop an exact preprocessing method for reducing the problem size as well as two optimization problems which provide upper and lower bounds on the objective with considerably less computation time compared to solving the exact problem. The bounds are experimentally analyzed on a small benchmark example.

Keywords Timetabling - PESP - Passenger routing - Integrated public transportation planning - Preprocessing

Acknowledgements We thank Florentin Hildebrandt for his help with the experiments.

1 Introduction

The *Periodic Scheduling Problem (PESP)* is a well researched and important problem in mathematical optimization. Its main application deals with periodic timetables in public transportation. Given a line plan, PESP can be used to find a periodic timetable with the goal to minimize the traveling times of the passengers. PESP has been introduced in [25] and shown to be NP complete. An important application of PESP is periodic timetabling for public transport, see e.g., [17, 11, 13, 14]. Advanced integer programming methods as in [18, 13, 4, 10], many of them using cycle bases, were the first approach and are still subject of ongoing research [2, 6]. SAT solvers have been used in [12, 5]. However, PESP is too hard to get optimal solutions even for instances of medium size, hence also heuristics have been developed, among them the modulo simplex [16, 7, 8] or a fast matching approach [19].

*This work was partially supported by DFG under SCHO 1140/8-1.

However, all these approaches are based on fixed passengers' weights and neglect that the route a passenger finally chooses depends on the timetable to be determined. [1] show that fixing passenger paths can lead to arbitrarily bad results compared to routing the passengers during the optimization and that integrated routing may lead to significant improvements on realistic instances. In [5] a SAT approach is used for integrating passengers routes. The importance of integrated routing has also been noted in [23, 24, 22]. However, integrating the routing decisions further increases the complexity of PESP significantly. We hence cannot hope to solve the problem exactly for instances of realistic size.

Based on an integer programming model to integrate passenger routing into timetabling, we propose an exact preprocessing method to reduce the problem size and a heuristic which improves the traveling time for the passengers as well as a second heuristic which provides a lower bound. Together, both heuristics admit an a-posteriori bound for an optimal solution. Both heuristics make use of an exact solution algorithm as subroutine for smaller problem instances. Using an IP solver for these subproblems, they are analyzed on a benchmark instance.

2 Model

For modeling the integrated timetabling and passenger routing problem, we follow the usual approach (see, e.g., [15, 11, 13, 14]). Timetabling requires an *event activity network* (*EAN*) $\mathcal{N}^0 = (\mathcal{E}^0, \mathcal{A}^0)$ consisting of nodes $\mathcal{E}^0$ which are called *events* and of arcs $\mathcal{A}^0$ which are called *activities*. Each activity $a \in \mathcal{A}^0$ has a lower and upper bound $L_a \leq U_a$ which restricts its duration. In PESP we look for a timetable π which assigns a time π_i to each event $i \in \mathcal{E}^0$. The timetable should be feasible in the sense that all durations of the activities lie between their lower and upper bounds. For non-periodic timetabling this condition is given as

$$L_a \leq \pi_j - \pi_i \leq U_a \quad \text{for all } a = (i,j) \in \mathcal{A}^0.$$

However, in PESP we look for a periodic timetable which is repeated every time period T. In this case, the feasibility condition is given as

$$(\pi_j - \pi_i - L_a) \bmod T + L_a \leq U_a \quad \text{for all } a = (i,j) \in \mathcal{A}^0. \tag{1}$$

In public transportation, the event activity network represents a timetable as follows: Given a public transportation network (PTN) (V, E) with stops or stations V and direct connections E, e.g., tracks, between them as well as a set of lines $\mathcal{L}$ that are operated, we start with an event activity network $\mathcal{N}^0 = (\mathcal{E}^0, \mathcal{A}^0)$. Its events represent arrival or departures of lines at stations while the activities represent the driving of vehicles between stations, vehicles waiting at stations or passengers transferring at stations between different lines. For a driving activity $a \in \mathcal{A}^0$ its lower and upper bounds L_a and U_a represent restrictions on the driving time, for trains waiting in a station, the lower bound sets the minimum time the train has to wait to allow passengers to board and alight, and the lower and upper bounds an transfer activities are used to ensure smooth transfers. Often, timetables are periodic with a period length of $T = 60$ minutes. A periodic timetable hence assigns a time $\pi_i \in \{0, \ldots, T-1\}$ to all events $i \in \mathcal{E}^0$ which is then repeated every hour. As in (1) it is feasible if the actual duration of all activities lies between the given upper and lower bounds. In the classic PESP, weights w_a approximating the number of passengers who use activity $a \in \mathcal{A}^0$ are additionally given and the goal is to minimize

$$\sum_{a=(i,j) \in \mathcal{A}^0} w_a \cdot (\pi_j - \pi_i - L_a) \bmod T.$$

The weights w_a can be computed by routing the passengers beforehand, e.g., according to shortest paths with respect to the lower bounds L_a on the activities. However, the routes the passengers really choose depend on the timetable to be computed. This chicken-and-egg causality dilemma has been recognized by many researchers, e.g., [23, 24, 22, 1, 5], hence, as mentioned in the introduction, integration of passengers routes is a topic of ongoing research.

In order to integrate the passenger routing into PESP, let a set of OD pairs $\text{OD} \subseteq V \times V$ be given, representing the pairs of stations between which passengers wish to travel. The event activity network has to be extended to include nodes and arcs representing the beginning or end of the passengers' journeys. The *extended EAN* $\mathcal{N} = (\mathcal{E}, \mathcal{A})$ therefore contains source and target nodes $(v, \text{source}), (v, \text{target})$ for all stations $v \in V$ as well as auxiliary arcs $\mathcal{A}_{\text{aux}}$ linking these new nodes to corresponding departure and arrival events as done in [24, 1, 5]. The upper and lower bounds of the auxiliary arcs are set to zero.

For formally stating the problem definition, we need further notation. To this end, let $\mathcal{N} = (\mathcal{E}, \mathcal{A})$ be the event activity network constructed above and L_a, U_a the the lower and upper bounds on its activities.

Definition 1. For an activity $a = (i, j) \in \mathcal{A}^0$ we define its duration w.r.t a timetable π as

$$d_a(\pi) = (\pi_j - \pi_i - L_a) \bmod T + L_a.$$

The duration of auxiliary activities $a \in \mathcal{A}_{\text{aux}}$ is defined as $d_a(\pi) = 0$ for all timetables π. We abbreviate the vector of durations as $d(\pi) = (d_a(\pi))_{a \in \mathcal{A}}$ and the vector of lower/upper bounds as $L = (L_a)_{a \in \mathcal{A}}$ and $U = (U_a)_{a \in \mathcal{A}}$, respectively.

For every feasible timetable π, the duration of an activity is larger than its lower and smaller than its upper bound, i.e.,

$$L_a \leq d_a(\pi) \leq U_a. \tag{2}$$

Let $\beta = (\beta_a)_{a \in \mathcal{A}}$ be a vector of activity lengths, then a shortest path for OD pair (u, v) from (u, source) to (v, target) according to these activity lengths is defined as $\text{SP}_{u,v}(\beta)$. The length of a path P according to the activity lengths β is defined as

$$\text{len}(P, \beta) = \sum_{a \in P} \beta_a,$$

i.e., $\text{len}(\text{SP}_{u,v}(\beta), \beta)$ is the length of the shortest path with respect to the activity lengths β while $\text{len}(\text{SP}_{u,v}(\beta), \gamma)$ describes the length of $\text{SP}_{u,v}(\beta)$ with respect to *other* activity lengths γ. Note that due to (2) the length of a shortest path w.r.t the lower bounds L_a on the activities is a lower bound on the duration for every feasible timetable π:

$$\text{len}(\text{SP}_{u,v}(L), L) \leq \text{len}(\text{SP}_{u,v}(d(\pi)), L) \leq \text{len}(\text{SP}_{u,v}(d(\pi)), d(\pi)). \tag{3}$$

Let $C_{u,v}$ be the number of passengers of OD pair $(u, v) \in \text{OD}$. Then we can state the *integrated timetabling and passenger routing problem.*

Integrated timetabling and passenger routing problem (P)
Find a feasible periodic timetable π with period length T minimizing the total travel time *on shortest paths w.r.t* $d(\pi)$ for all passengers, i.e.,

$$\min \quad \mathcal{R}_{\text{SP}}(\pi) = \sum_{(u,v) \in \text{OD}} C_{u,v} \cdot \text{len}(\text{SP}_{u,v}(d(\pi)), d(\pi))$$

such that π is a feasible periodic timetable with period length T.

$\mathcal{R}_{\text{SP}}(\pi)$ evaluates a timetable π w.r.t *shortest path routing.* The problem is called *integrated* because the paths for the OD pairs are not precomputed but determined within the optimization. Note that this is not the case in the classical PESP in which the OD pairs are routed beforehand, usually on shortest paths with respect to the lower bounds L_a on the activities. This means, in the classical PESP we only look for the timetable, but not for the paths. In view of the notations here, the (classical) PESP can be given as follows:

Periodic event scheduling problem ($PESP$)
Find a feasible periodic timetable π with period length T minimizing the total travel time *on fixed shortest paths w.r.t the lower bounds* L for all passengers, i.e.,

$$\min \quad \mathcal{R}_{\text{LB}}(\pi) = \sum_{(u,v)\in\text{OD}} C_{u,v} \cdot \text{len}(\text{SP}_{u,v}(L), d(\pi))$$

such that π is a feasible periodic timetable with period length T.

$\mathcal{R}_{\text{LB}}(\pi)$ evaluates a timetable π w.r.t *lower bound routing.*

As we are solving (P) with integer programming methods, we give the IP model (along the lines of [22, 1, 5]) combining PESP constraints for the timetable on the EAN with a passenger flow model on the extended EAN. π_i are variables representing the (periodic) time of the events while z_a are so-called *modulo parameters* for the activities modeling the periodicity of the timetable. The passenger flow is represented by the flow variables $p_a^{u,v}$.

$$(P) \qquad \min \sum_{(u,v)\in\text{OD}} C_{u,v} \cdot \sum_{a=(i,j)\in\mathcal{A}^0} p_a^{u,v} \cdot (\pi_j - \pi_i + z_a \cdot T) \tag{4}$$

$$\pi_j - \pi_i + z_a \cdot T \geq L_a \qquad a = (i,j) \in \mathcal{A}^0 \tag{5}$$

$$\pi_j - \pi_i + z_a \cdot T \leq U_a \qquad a = (i,j) \in \mathcal{A}^0 \tag{6}$$

$$A \cdot (p_a^{u,v})_{a\in\mathcal{A}} = b^{u,v} \qquad (u,v) \in \text{OD} \tag{7}$$

$$\pi_i \in \{0, \ldots, T-1\} \qquad i \in \mathcal{E}^0$$

$$z_a \in \mathbb{Z} \qquad \forall a \in \mathcal{A}^0$$

$$p_a^{u,v} \in \{0,1\} \qquad (u,v) \in \text{OD}, a \in \mathcal{A}$$

Note that the objective function can easily be linearized which is omitted here. While constraints (5) and (6) are the standard PESP constraints modeling the timetable, the passenger flow is handled in constraint (7) accounting for the majority of the constraints. The passenger flow is modeled separately for each OD pair $(u,v) \in \text{OD}$ using flow variables $p_a^{u,v}$ where A is the node-arc-incidence matrix of the extended EAN $\mathcal{N}$ and the vector $b^{u,v}$ ensures that the flow starts at the source node belonging to the OD pair and ends at the corresponding target node.

The standard PESP constraints (5) and (6) can be substituted by cycle base PESP constraints, see e.g., [15, 20] leading to a significant decrease in runtime for the classical PESP. In Section 4 we experimentally see that this also holds for the integrated formulation (P).

3 Two approaches for reducing the problem size

In this section we describe the two main ideas that make the integrated timetable and passenger routing problem (P) tractable. These ideas can also be applied to any other problem in which OD pairs should be routed along shortest paths, e.g., network design problems and problems in telecommunication.

3.1 Combining shortest path routing with routing along fixed paths

The major factor for the size of the problem is the number of OD pairs as for each OD pair and each activity a new variable has to be created. The first idea for reducing the problem size therefore is to only route a subset $\mathrm{OD}_{\mathrm{route}} \subset \mathrm{OD}$ of the passengers as described in Heuristic LB. To this end, we split OD into two disjoint sets $\mathrm{OD} = \mathrm{OD}_{\mathrm{route}} \sqcup \mathrm{OD}_{\mathrm{fix}}$.

Heuristic LB
Find a feasible periodic timetable π with period length T minimizing the total travel time *on shortest paths w.r.t* $d(\pi)$ for all passengers in $\mathrm{OD}_{\mathrm{route}}$, i.e.,

$$\min \quad \mathcal{R}_{\mathrm{SP}}(\mathrm{OD}_{\mathrm{route}}, \pi) = \sum_{(u,v)\in \mathrm{OD}_{\mathrm{route}}} C_{u,v} \cdot \mathrm{len}(\mathrm{SP}_{u,v}(d(\pi)), d(\pi))$$

such that π is a feasible periodic timetable with period length T.

In $\mathcal{R}_{\mathrm{SP}}(\mathrm{OD}_{\mathrm{route}}, \pi)$, the evaluation based on shortest path routing is restricted to OD pairs in $\mathrm{OD}_{\mathrm{route}}$. In order to better compare the solution found by Heuristic LB with other solutions, we add lower bounds for every OD pair from $\mathrm{OD}_{\mathrm{fix}}$ to the value of $\mathcal{R}_{\mathrm{SP}}(\mathrm{OD}_{\mathrm{route}}, \pi)$. For $\overline{\mathrm{OD}} \subseteq \mathrm{OD}$ we define

$$\tilde{L}(\overline{\mathrm{OD}}) = \sum_{(u,v)\in \overline{\mathrm{OD}}} C_{u,v} \cdot \mathrm{len}(\mathrm{SP}_{u,v}(L), L)$$

summing for all OD pairs in $\overline{\mathrm{OD}}$ the length of a shortest path w.r.t the lower bounds of the activities measured by these lower bounds. Due to (3) we have that

$$\tilde{L}(\overline{\mathrm{OD}}) \leq \mathcal{R}_{\mathrm{SP}}(\overline{\mathrm{OD}}, \pi) \tag{8}$$

for any feasible timetable π. We define

$$h(\mathrm{OD}_{\mathrm{route}}, \pi) = \mathcal{R}_{\mathrm{SP}}(\mathrm{OD}_{\mathrm{route}}, \pi) + \tilde{L}(\mathrm{OD}_{\mathrm{fix}})$$

which yields a better approximation of $\mathcal{R}_{\mathrm{SP}}(\pi)$. However, completely disregarding the passengers not in $\mathrm{OD}_{\mathrm{route}}$ in the optimization can lead to objective values that are even worse than the values of the original PESP as can be seen in the experiments in Section 4.3. The heuristic is hence useless for generating a good solution, but can nevertheless be used as a good lower bound on (P) as we will see in Theorem 3.

The better idea is to include also the OD pairs in $\mathrm{OD}_{\mathrm{fix}}$ in the optimization by routing them beforehand on shortest paths w.r.t the lower bounds of the activities and add their travel times as weights to the objective function. This is described in Heuristic UB.

Heuristic UB
Find a feasible periodic timetable π with period length T minimizing the sum of the total travel time *on shortest paths w.r.t* $d(\pi)$ for all passengers in $\mathrm{OD}_{\mathrm{route}}$ and the total travel time *on fixed shortest paths w.r.t the lower bounds* L for all passengers in $\mathrm{OD}_{\mathrm{fix}}$, i.e.,

$$\begin{aligned}\min \quad f(\mathrm{OD}_{\mathrm{route}}, \pi) = & \sum_{(u,v)\in \mathrm{OD}_{\mathrm{route}}} C_{u,v} \cdot \mathrm{len}(\mathrm{SP}_{u,v}(d(\pi)), d(\pi)) \\ & + \sum_{(u,v)\in \mathrm{OD}_{\mathrm{fix}}} C_{u,v} \cdot \mathrm{len}(\mathrm{SP}_{u,v}(L), d(\pi))\end{aligned}$$

a such that π is a feasible periodic timetable with period length T.

For $\overline{\text{OD}} \subseteq \text{OD}$ define

$$\mathcal{R}_{\text{LB}}(\overline{\text{OD}}, \pi) = \sum_{(u,v)\in\overline{\text{OD}}} C_{u,v} \cdot \text{len}(\text{SP}_{u,v}(L), d(\pi)),$$

i.e., we evaluate the timetable π w.r.t *lower bound routing* restricted to OD pairs in $\overline{\text{OD}}$. We then receive

$$f(\text{OD}_{\text{route}}, \pi) = \mathcal{R}_{\text{SP}}(\text{OD}_{\text{route}}, \pi) + \mathcal{R}_{\text{LB}}(\text{OD}_{\text{fix}}, \pi), \tag{9}$$

i.e., Heuristic UB combines shortest path routing for the OD pairs in OD_{route} with fixed routing w.r.t the lower-bounds for the OD pairs in OD_{fix}. Although Heuristic LB may provide better solutions than Heuristic UB, our experiments show that the latter performs significantly better. For $\text{OD}_{\text{route}} = \text{OD}$ we get problem (P) in which all passengers are routed during the optimization and for $\text{OD}_{\text{fix}} = \text{OD}$ we get the classical PESP. In Section 4.1 we have a look at different strategies for choosing OD_{route}.

We can easily see that both, Heuristic LB and Heuristic UB, behave monotonously when the set OD_{route} is extended.

Lemma 2. *Let* OD^1, OD^2 *with* $\text{OD}^1 \subset \text{OD}^2$ *be two sets of OD pairs.*

1. *Let* $\tilde{\pi}^1, \tilde{\pi}^2$ *be solutions for Heuristic LB w.r.t* OD^1 *and* OD^2.
   Then $h(\text{OD}^1, \tilde{\pi}^1) \leq h(\text{OD}^2, \tilde{\pi}^2)$.
2. *Let* π^1, π^2 *be solutions for Heuristic UB w.r.t* OD^1 *and* OD^2.
   Then $f(\text{OD}^1, \pi^1) \geq f(\text{OD}^2, \pi^2)$.

Proof.

1. Since $\tilde{\pi}^1$ is optimal for OD^1 we compute

$$\begin{aligned} h(\text{OD}^1, \tilde{\pi}^1) &\leq h(\text{OD}^1, \tilde{\pi}^2) \\ &= \mathcal{R}_{\text{SP}}(\text{OD}^1, \tilde{\pi}^2) + \tilde{L}(\text{OD}^2 \setminus \text{OD}^1, \tilde{\pi}^2) + \tilde{L}(\text{OD} \setminus \text{OD}^2) \\ &\overset{(8)}{\leq} \mathcal{R}_{\text{SP}}(\text{OD}^1, \tilde{\pi}^2) + \mathcal{R}_{\text{SP}}(\text{OD}^2 \setminus \text{OD}^1, \tilde{\pi}^2) + \tilde{L}(\text{OD} \setminus \text{OD}^2) \\ &= h(\text{OD}^2, \tilde{\pi}^2). \end{aligned}$$

2. Here, π^2 is optimal for OD^2 and we obtain

$$\begin{aligned} f(\text{OD}^2, \pi^2) &\leq f(\text{OD}^2, \pi^1) \\ &= \mathcal{R}_{\text{SP}}(\text{OD}^1, \pi^1) + \mathcal{R}_{\text{SP}}(\text{OD}^2 \setminus \text{OD}^1, \pi^1) + \mathcal{R}_{\text{LB}}(\text{OD} \setminus \text{OD}^2, \pi^1) \\ &\overset{(*)}{\leq} \mathcal{R}_{\text{SP}}(\text{OD}^1, \pi^1) + \mathcal{R}_{\text{LB}}(\text{OD}^2 \setminus \text{OD}^1, \pi^1) + \mathcal{R}_{\text{LB}}(\text{OD} \setminus \text{OD}^2, \pi^1) \\ &= f(\text{OD}^1, \pi^1) \end{aligned}$$

   where (*) holds since $\mathcal{R}_{\text{SP}}(\overline{\text{OD}}, \pi) \leq \mathcal{R}_{\text{LB}}(\overline{\text{OD}}, \pi)$ for all $\overline{\text{OD}} \subseteq \text{OD}$ and all feasible π.

□

Lemma 2 is the main ingredient for the following theorem which shows that h and f are lower and upper bounds on the optimal objective value of (P) and improve when OD_{route} is extended.

Theorem 3. *Let* $\mathrm{OD}^1, \mathrm{OD}^2$ *with* $\mathrm{OD}^1 \subset \mathrm{OD}^2$ *be two sets of routed OD pairs,* $\tilde{\pi}^1, \tilde{\pi}^2$ *the respective solutions for Heuristic LB and* π^1, π^2 *the respective solutions for Heuristic UB. Let* π^* *be an optimal solution for the integrated timetabling and passenger routing problem* (P)*.*

1. $h(\mathrm{OD}^i, \tilde{\pi}^i)$ *is a lower bound on* $\mathcal{R}_{\mathrm{SP}}(\pi^*)$ *for* $i \in \{1,2\}$*.*
2. $f(\mathrm{OD}^i, \pi^i)$ *is an upper bound on* $\mathcal{R}_{\mathrm{SP}}(\pi^*)$ *for* $i \in \{1,2\}$*.*
3. $h(\mathrm{OD}^1, \tilde{\pi}^1) \leq h(\mathrm{OD}^2, \tilde{\pi}^2) \leq \mathcal{R}_{\mathrm{SP}}(\pi^*) \leq f(\mathrm{OD}^2, \pi^2) \leq f(\mathrm{OD}^1, \pi^1)$*.*

Proof. Part **1** and **2** are clear from Lemma 2 as for $\mathrm{OD}_{\mathrm{route}} = \mathrm{OD}$ both Heuristic LB and UB compute the optimal solution of (P). Part **3** is a direct implication of parts **1** and **2** and Lemma 2. □

We can also approximate the resulting gap as the following corollary shows.

Corollary 4. *Let* $\tilde{\pi}$ *be a solution for Heuristic LB and* π *a solution for Heuristic UB for* $\mathrm{OD}_{\mathrm{route}} \subset \mathrm{OD}$ *and* $\mathrm{OD}_{\mathrm{route}} \sqcup \mathrm{OD}_{\mathrm{fix}} = \mathrm{OD}$*. Let* π^* *be an optimal solution for the integrated timetabling and passenger routing problem* (P)*. Then the optimality gap can be bounded by*

$$\mathcal{R}_{\mathrm{SP}}(\pi) - \mathcal{R}_{\mathrm{SP}}(\pi^*) \leq \sum_{(u,v)\in \mathrm{OD}_{\mathrm{fix}}} C_{u,v} \cdot \mathrm{len}(\mathrm{SP}_{u,v}(L), U - L).$$

Proof.

$$\begin{aligned}
\mathcal{R}_{\mathrm{SP}}(\pi) - \mathcal{R}_{\mathrm{SP}}(\pi^*) &\leq f(\mathrm{OD}_{\mathrm{route}}, \pi) - h(\mathrm{OD}_{\mathrm{route}}, \tilde{\pi}) \leq f(\mathrm{OD}_{\mathrm{route}}, \tilde{\pi}) - h(\mathrm{OD}_{\mathrm{route}}, \tilde{\pi}) \\
&= \mathcal{R}_{\mathrm{SP}}(\mathrm{OD}_{\mathrm{route}}, \tilde{\pi}) + \mathcal{R}_{\mathrm{LB}}(\mathrm{OD}_{\mathrm{fix}}, \tilde{\pi}) - \mathcal{R}_{\mathrm{SP}}(\mathrm{OD}_{\mathrm{route}}, \tilde{\pi}) - \tilde{L}(\mathrm{OD}_{\mathrm{fix}}) \\
&= \sum_{(u,v)\in \mathrm{OD}_{\mathrm{fix}}} C_{u,v} \cdot \mathrm{len}(\mathrm{SP}_{u,v}(L), d(\tilde{\pi}) - L) \\
&\overset{(2)}{\leq} \sum_{(u,v)\in \mathrm{OD}_{\mathrm{fix}}} C_{u,v} \cdot \mathrm{len}(\mathrm{SP}_{u,v}(L), U - L).
\end{aligned}$$

□

For example, if U_a and L_a differ only by a fixed percentage p, i.e., if $U_a = L_a(1 + \frac{p}{100})$, the optimality gap for the solution computed by Heuristic UB is at most

$$\tfrac{p}{100} \sum_{(u,v)\in \mathrm{OD}_{\mathrm{fix}}} C_{u,v} \cdot \mathrm{len}(\mathrm{SP}_{u,v}(L), L).$$

3.2 Preprocessing

When integrating passenger routing into timetabling, we create flow variables for all passengers and *all* activities. However, if OD pairs use shortest paths, usually some activities can be sorted out beforehand. E.g., an OD pair traveling from l'Aquila to Rome is unlikely to pass through Milano on a shortest path no matter which timetable we choose. We try to find for each OD pair a small subset of activities such that no matter what timetable is chosen in the end, this subset contains a shortest path for the OD pair. This means that it suffices to generate flow variables for this OD pair only for this subset of activities instead of all activities.

For each activity $a \in \mathcal{A}$ we know that it is not part of a shortest path from s to t for any timetable if the best case shortest path from s to t via a is longer than the worst case shortest path from s directly to t. Here, the best and worst case depend on the timetable, i.e., the best case shortest path is the shortest $s - a - t$ path for the timetable which is best for this OD pair while the worst case shortest path is the shortest $s - t$ paths in the worst possible timetable.

But as finding a feasible timetable is NP complete, see [25], the construction of a best or worst case timetable is difficult as well. Instead, we use the lower and upper bounds on the activities as bounds on the length of the best/worst case shortest paths.

Theorem 5. *Let $(u, v) \in \mathrm{OD}$ be an OD pair and let $a = (i, j) \in \mathcal{A}$ be an activity, such that*

$$\mathrm{len}(\mathrm{SP}_{u,v}(U), U) < \mathrm{len}(\mathrm{SP}_{u,i}(L), L) + L_a + \mathrm{len}(\mathrm{SP}_{j,v}(L), L).$$

Then, for any timetable π no shortest path $\mathrm{SP}_{u,v}(d(\pi))$ w.r.t π contains activity a.

Proof. Let π be any timetable and $\mathrm{SP}_{u,v}(d(\pi))$ be a shortest path w.r.t π. Then its length satisfies

$$\begin{aligned}
\mathrm{len}(\mathrm{SP}_{u,v}(d(\pi)), d(\pi)) &\leq \mathrm{len}(\mathrm{SP}_{u,v}(U), d(\pi)) \leq \mathrm{len}(\mathrm{SP}_{u,v}(U), U) \\
&< \mathrm{len}(\mathrm{SP}_{u,i}(L), L) + L_a + \mathrm{len}(\mathrm{SP}_{j,v}(L), L) \text{ for } a = (i, j) \\
&\leq \mathrm{len}(P, L) \text{ for any path } P \text{ containing activity } a \\
&\leq \mathrm{len}(P, d(\pi)) \text{ for any path } P \text{ containing activity } a.
\end{aligned}$$

□

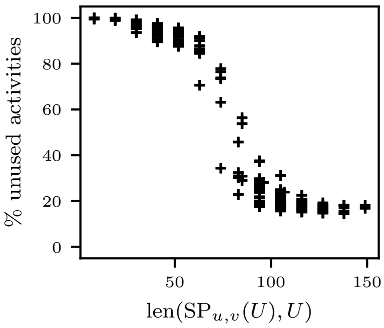
Based on Theorem 5 we propose the following algorithm.

Algorithm 1 Preprocessing for integrated timetabling and passenger routing

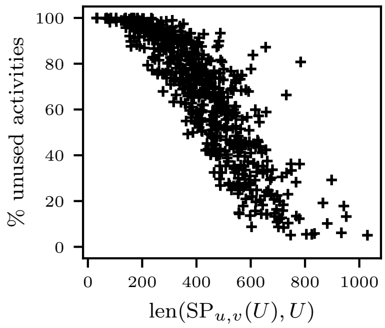
Input: EAN $\mathcal{N} = (\mathcal{E}, \mathcal{A})$, interval $[L_a, U_a]$ of possible arc length for all activities $a \in \mathcal{A}$, starting event s, ending event t.
Output: list of activities $\bar{\mathcal{A}}$ which are not needed.
Initialize $\bar{\mathcal{A}} = \emptyset$.
Compute $\beta := \mathrm{len}(\mathrm{SP}_{s,t}(U), U)$.
for event $i \in \mathcal{E}$ **do**
    Compute $\gamma_i := \mathrm{len}(\mathrm{SP}_{s,i}(L), L)$.
    Compute $\delta_i := \mathrm{len}(\mathrm{SP}_{i,t}(L), L)$.
end for
for activity $a = (i, j) \in \mathcal{A}$ **do**
    if $\gamma_i + L_a + \delta_j > \beta$ **then**
        $\bar{\mathcal{A}} = \bar{\mathcal{A}} \cup \{a\}$
    end if
end for

As shown in Figure 1 the number of activities which are not needed highly depends on the length of the worst case shortest path. Especially if origin and destination are close, almost all other activities can be discarded when looking for a shortest path.

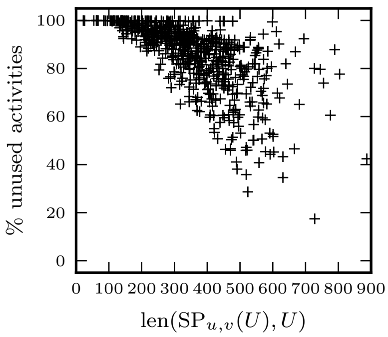
Another important factor is the variability of the path length. In many publications on periodic timetabling, it is assumed that the durations of the driving and waiting activities are fixed, e.g. in [1, 19, 12]. We analyzed the effect of this assumption on the preprocessing step: Figures 1b and 1c show the same event activity network differing in the bounds on the duration of the activities. In this case we used a close-to-real world network based on the long-distance train network of Germany. While in Figure 1b all activity durations are allowed to be in a given interval, in Figure 1c most activity durations are fixed and only the durations of transfer activities are variable (restricted to intervals). This decreases the problem size a lot and additionally increases the effect of the preprocessing algorithm as the difference between best and worst case paths decreases. We conclude that preprocessing is significantly more effective for fixed durations of the waiting and driving activities.



(a) All OD pairs of dataset `grid`.



(b) 10 % of OD pairs of dataset `long-distance`.



(c) 10 % of OD pairs of dataset `long-distance` with fixed durations of drive and wait activities.

Figure 1: Percentage of unused activities depending on the length of the worst case shortest path. Dataset `grid` is the benchmark example used in Section 4, datset `long-distance` is derived from the German long-distance train network.

4 Computational results

We show the influence of passenger routing as well as the benefits of preprocessing on two datasets from the software framework LinTim, see [21]. The evaluations in Sections 4.1 to 4.4 are done on the benchmark dataset `grid` which was introduced in [3]. It models an urban traffic network which consists of 25 stations arranged as a 5×5 grid and 40 edges, resulting in an event activity network with 392 events and 2382 activities. There are 567 OD pairs with a total of 2546 passengers. Note that contrasting to many other publications on periodic timetabling, for dataset `grid` we do *not* assume that the duration of drive and wait activities is fixed, yielding a larger but more realistic problem.

To demonstrate some technical aspects of the bounds, we use a very small artificial dataset `toy` with 8 stations and 8 edges with two different line plans resulting in a an event activity network with 32 events and 44 activities for `toy-1` and 156 events and 1088 activities for `toy-2`.

We use Gurobi 8 [9] to solve the IP model presented in Section 2 and an IP formulation of the cycle base variant on a computer with 6 CPUs at 3.06 GHz and 132 GB RAM.

For all datasets we use a 4 hour limit on the solver time which is never reached for datasets `toy-1` and `toy-2`.

Our experiments investigate what happens if the number of OD pairs to be routed is increased. To this end we determine sets OD^k which contain k OD pairs to be routed. This is done as follows: We sort the OD pairs according to some given rule as described in Section 4.1. OD^k is then defined as the set of the first k OD pairs according to the sorting. We run the two heuristics with the sets $\mathrm{OD}_{\mathrm{route}} = \mathrm{OD}^k$, $k = 0, \ldots, |\mathrm{OD}|$. The timetables resulting from Heuristic UB are called π^k and the timetables resulting from Heuristic LB are denoted by $\tilde{\pi}^k$. The optimal timetable for (P) minimizing $\mathcal{R}_{\mathrm{SP}}(\pi)$ is denoted by π^*.

4.1 Which OD pairs should be routed?

In order to determine which OD pairs should be routed during the optimization, we compare different methods to choose the OD pairs in OD^k for dataset `grid`, namely routing the k largest or the k smallest OD pairs, k random OD pairs, the k OD pairs with the largest Euclidean distance between origin and destination or choose k OD pairs according to Corollary 4. Therefore, we take the k OD pairs for which the differences $U_a - L_a$ between the upper and lower bounds on their shortest paths (w.r.t lower bound routing) weighted by the number of passengers C_{uv} are largest. Figure 2 shows that, especially when few OD pairs are routed, the choice of the routed OD pairs strongly impacts the solution quality. Routing the "wrong" set of OD pairs can even lead to solutions which are worse than not routing any OD pairs. When many OD pairs are routed the influence of the method to choose routed OD pairs diminishes. Choosing OD pairs for $\mathrm{OD}_{\mathrm{route}}$ according to Corollary 4 yields by far the best results leading to an improvement in the objective by only routing 5 OD pairs which is not matched by most of the other methods when 150 OD pairs are routed.

4.2 Influence of preprocessing and chosen IP formulation

We now have a look at the influence of the preprocessing method presented in Section 3.2 and of the different IP formulations on the runtime of our algorithm.

Figure 3a shows that, as in the classical PESP, the IP model for PESP with integrated routing has a shorter runtime if the cycle base formulation is used instead of the standard IP formulation. The positive effect of the preprocessing method can be better observed in Figure 3b as in Figure 3a the solver time limits are hit earlier.

As the preprocessing method does not change the optimal solution of (P) and both IP formulations are equivalent, we only consider the version of the algorithm which uses the cycle base IP model and the preprocessing method in the following sections. Note that the runtime of Heuristic LB is considerably shorter than the runtime of Heuristic UB. Section 4.3 shows the large differences in solution quality between them.

4.3 Comparing Heuristic LB and Heuristic UB

Figure 4 shows the solution quality of the Heuristics LB and UB when both are evaluated w.r.t shortest path routing for all passengers, i.e., when evaluating the objective function $\mathcal{R}_{\mathrm{SP}}(\pi^k)$

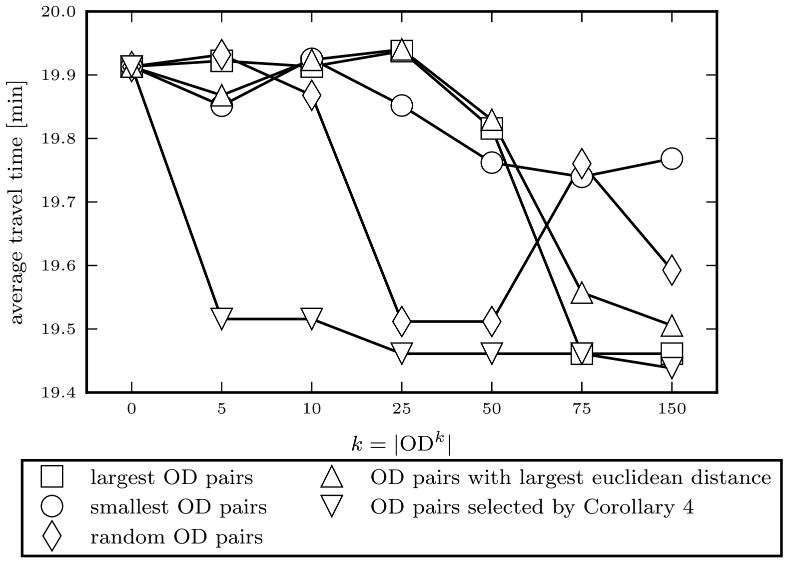
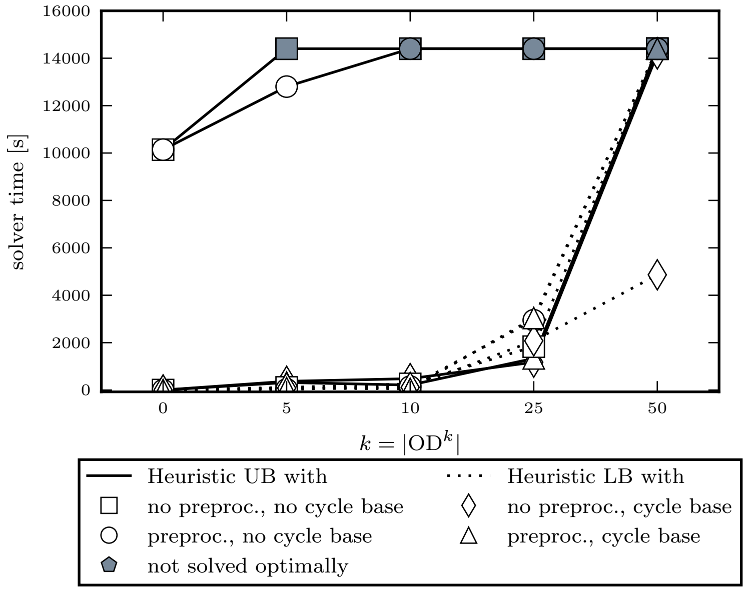


Figure 2: Comparison of different methods to choose OD^k for dataset `grid`.

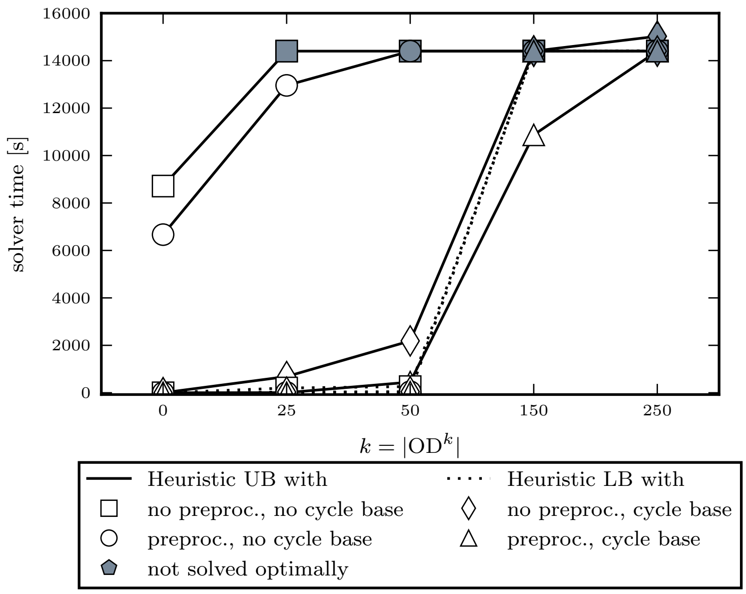
and $\mathcal{R}_{\mathrm{SP}}(\tilde{\pi}^k)$ of the original problem. We see that the solution quality for Heuristic UB is better with a maximal difference of 9.3%. Additionally, the quality of solutions found by Heuristic UB improves monotonously while the quality of solutions found by Heuristic LB fluctuates a lot. As we have seen in Section 4.1 this can also be the case for solutions found by Heuristic UB depending on the routing method. The more OD pairs are routed the better Heuristic LB becomes as the influence of the neglected OD pairs in $\mathrm{OD}_{\mathrm{fix}}$ diminishes. For 150 routed OD pairs the difference between both heuristics is only 1%.

4.4 Best configuration with bounds

Sections 4.2 and 4.3 show that using Heuristic UB with cycle bases and preprocessing is the fastest way to get good solutions for the integrated timetabling and passenger routing problem. We test this approach for dataset `grid` to find solutions when more (and even all) OD pairs are routed. This is shown in Figure 5 together with the behavior of the resulting bounds $f(\mathrm{OD}^k, \pi^k)$ and $h(\mathrm{OD}^k, \tilde{\pi}^k)$ where both heuristics are run for increasing sizes of $\mathrm{OD}_{\mathrm{route}} = \mathrm{OD}^k$. When at least 50 OD pairs are routed, the time limit of 4 hours does not suffice to solve the problem optimally such that the bounds also are only approximations. The lower bound $h(\mathrm{OD}^k, \tilde{\pi}^k)$ is adjusted according to the gap such that it is still a lower bound on $\mathcal{R}_{\mathrm{SP}}(\pi^*)$ where π^* is an optimal timetable for (P) while the upper bound $f(\mathrm{OD}^k, \pi^k)$ does not need to be adjusted. When routing all OD pairs the gap of the IP solver is with 3% so large that the solution is slightly worse than the solution for routing 400 passengers. Compared to the solution of the classical PESP the excess travel time, i.e., the travel time that is needed additionally to the lower bound $\sum_{(u,v)\in \mathrm{OD}} C_{u,v} \cdot \mathrm{SP}_{u,v}(L, L)$ when all passengers travel on shortest paths and these paths are all realized with the lower bounds, is reduced by 51.8% when 400 OD pairs are routed. Note that the lower bound does not change from the trivial lower bound no matter how many OD pairs



(a) Routing k OD pairs according to Corollary 4.



(b) Routing the k largest OD pairs.

Figure 3: Influence of preprocessing and of the choice of the IP formulation on the runtime for dataset `grid` for Heuristic LB and Heuristic UB.

are routed. This is due to the fact that for routing many OD pairs the problem was not solved optimally and the solver did not find a better lower bound during the optimization process. In

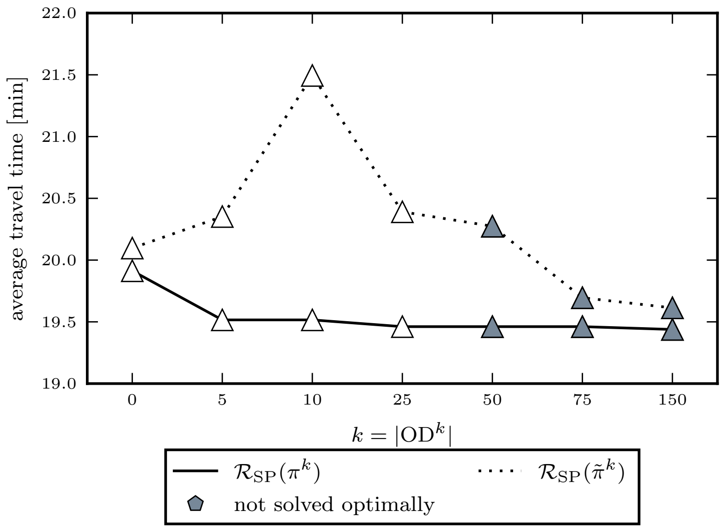


Figure 4: Comparing Heuristic LB and Heuristic UB for dataset `grid`. Here, π^k,$\tilde{\pi}^k$ are the optimal timetables for Heuristic UB and Heuristic LB routing the OD pairs in OD^k.

Section 4.5 we show that the lower bound does change depending on the number of OD pairs routed and can be used to prove optimality even when not all OD pairs are routed.

4.5 Bounds for instance `toy`

Although for the presented dataset `grid` the lower bound could not be improved from the trivial lower bound, this is not always the case. Figure 6 shows the value of the lower bound $h(\mathrm{OD}^k, \tilde{\pi}^k)$ as in Figure 6a the actual gap can be bounded early on and in Figure 6b it helps to show that the solution found by routing 5 OD pairs is already optimal.

5 Outlook

The heuristics and the preprocessing method presented in this paper can also make use of other solution approaches to the integrated timetabling and passenger routing problem. Replacing the IP solver for example by the column generation approach used in [1] might yield even better solutions. Also, selecting the OD pairs to be routed is ongoing research. This may even be combined with iterative approaches: Start with finding a timetable for $\mathrm{OD}_{\mathrm{route}} = \emptyset$. Determine the OD pairs whose rerouted path does not coincide with the path determined by the lower bound routing, and add them to $\mathrm{OD}_{\mathrm{route}}$ in the next step.

It is subject of current research to integrate not only timetabling and routing, but also timetabling, line planning, and routing. In an adapted form the two heuristics presented in this paper can also be applied to this extension of the integrated timetabling and passenger routing problem.

Since the ideas presented here are of general nature, they can also be applied to other problems where routing is a part of the objective, e.g., in location planning or in telecommunication.

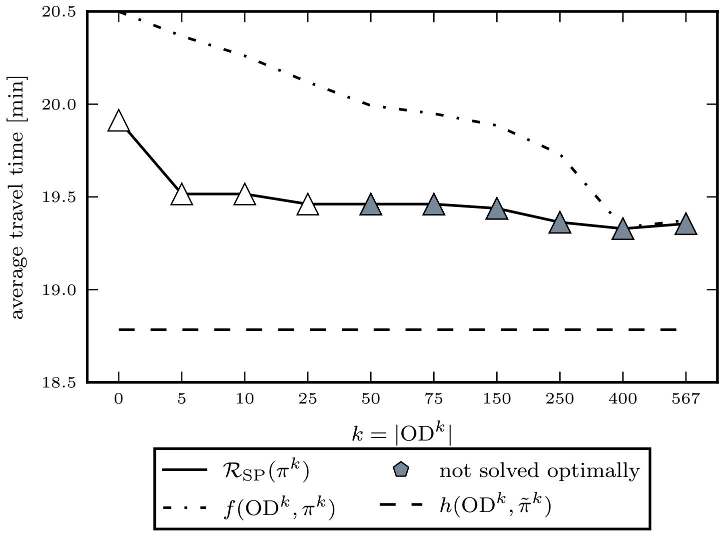
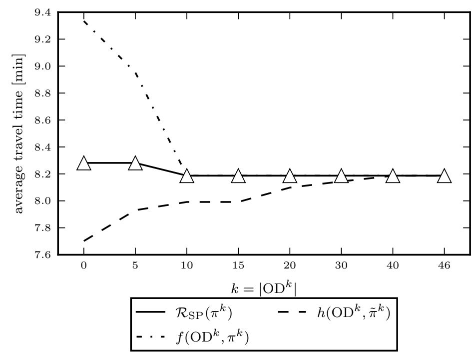
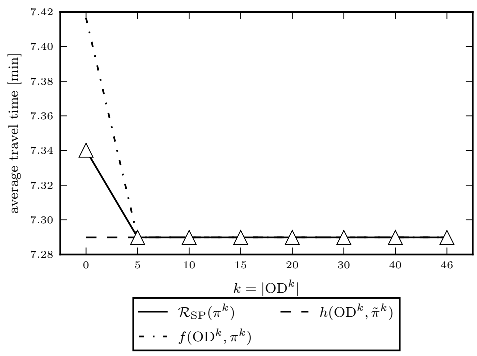


Figure 5: Evaluation of Heuristic UB and Heuristic LB with cycle based IP formulation and preprocessing for routing k OD pairs according to Corollary 4. We depict the lower and upper bound on the optimal objective value $\mathcal{R}_{\mathrm{SP}}(\pi^*)$ for dataset `grid` together with the objective values $\mathcal{R}_{\mathrm{SP}}(\pi^k)$ for the solutions π^k obtained by Heuristic UB.



(a) Dataset `toy-1`.



(b) Dataset `toy-2`.

Figure 6: Evaluation of Heuristic UB and Heuristic LB with cycle based IP formulation and preprocessing for routing k OD pairs according to Corollary 4. We depict the lower and upper bound on the optimal objective value $\mathcal{R}_{\mathrm{SP}}(\pi^*)$ together with the objective values $\mathcal{R}_{\mathrm{SP}}(\pi^k)$ for the solutions π^k obtained by Heuristic UB.

References

[1] R. Borndörfer, H. Hoppmann, and M. Karbstein. Passenger routing for periodic timetable optimization. *Public Transport*, 2016.

[2] R. Borndörfer, H. Hoppmann, and M. Karbstein. Separation of Cycle Inequalities for the Periodic Timetabling Problem. In Piotr Sankowski and Christos Zaroliagis, editors, *24th Annual European Symposium on Algorithms (ESA 2016)*, volume 57 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 21:1–21:13, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[3] M. Friedrich, M. Hartl, A. Schiewe, and A. Schöbel. Angebotsplanung im öffentlichen Verkehr - planerische und algorithmische Lösungen. In *Heureka'17*, 2017.

[4] L. Galli and S. Stiller. Strong Formulations for the Multi-module PESP and a Quadratic Algorithm for Graphical Diophantine Equation Systems. In *Proc. ESA*, 2010.

[5] P. Gattermann, P. Großmann, K. Nachtigall, and A. Schöbel. Integrating Passengers' Routes in Periodic Timetabling: A SAT approach. In Marc Goerigk and Renato Werneck, editors, *16th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2016)*, volume 54 of *OpenAccess Series in Informatics (OASIcs)*, pages 1–15, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[6] M. Goerigk and C. Liebchen. An Improved Algorithm for the Periodic Timetabling Problem. In Gianlorenzo D'Angelo and Twan Dollevoet, editors, *17th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2017)*, volume 59 of *OpenAccess Series in Informatics (OASIcs)*, pages 12:1–12:14, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[7] M. Goerigk and A. Schöbel. Engineering the modulo network simplex heuristic for the periodic timetabling problem. In Panos Pardalos and Steffen Rebennack, editors, *Proceedings of the 10th International Symposium on Experimental Algorithms (SEA)*, volume 6630 of *Lecture Notes in Computer Science*, pages 181–192. Springer, 2011.

[8] M. Goerigk and A. Schöbel. Improving the modulo simplex algorithm for large-scale periodic timetabling. *Computers & Operations Research*, 40(5):1363–1370, 2013.

[9] Gurobi Optimizer. see http://www.gurobi.com/, 2018. Gurobi Optimizer Version 8.0, Houston, Texas: Gurobi Optimization, Inc.

[10] T. Kavitha, C. Liebchen, K. Mehlhorn, D. Michails, T. Ueckerdt, and K.A. Zweig. Cycle bases in graph: Characterization, algorithms, complexity, and applications. *Computer Science Review*, 3(4):199–243, 2009.

[11] L. Kroon and L. Peeters. A variable trip time model for cyclic railway timetabling. *Transportation Science*, 37(2):198–212, 2003.

[12] M. Kümmling, P. Großmann, K. Nachtigall, J. Opitz, and R. Weiß. A state-of-the-art realization of cyclic railway timetable computation. *Public Transport*, 7(3):281–293, 2015.

[13] C. Liebchen. Finding short integral cycle bases for cyclic timetabling. In *Proceedings of European Symposium on Algorithms (ESA) 2003*, pages 715–726, 2003.

[14] C. Liebchen. *Periodic Timetable Optimization in Public Transport.* dissertation.de – Verlag im Internet, Berlin, 2006.

[15] K. Nachtigall. *Periodic Network Optimization and Fixed Interval Timetables.* PhD thesis, University of Hildesheim, 1998.

[16] K. Nachtigall and J. Opitz. Solving Periodic Timetable Optimisation Problems by Modulo Simplex Calculations. In Matteo Fischetti and Peter Widmayer, editors, *8th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'08)*, volume 9 of *OpenAccess Series in Informatics (OASIcs)*, Dagstuhl, Germany, 2008. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[17] K. Nachtigall and S. Voget. A genetic algorithm approach to periodic railway synchronization. *Computers & Operations Research*, 23(5):453–463, 1996.

[18] M. Odijk. A constraint generation algorithm for the construction of periodic railway timetables. *Transportation Research Part B: Methodological*, 30(6):455–464, 1996.

[19] J. Pätzold and A. Schöbel. A Matching Approach for Periodic Timetabling. In Marc Goerigk and Renato Werneck, editors, *16th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2016)*, volume 54 of *OpenAccess Series in Informatics (OASIcs)*, pages 1–15, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[20] L. Peeters and L. Kroon. A cycle based optimization model for the cyclic railway timetabling problem. In S. Voß and J. Daduna, editors, *Computer-Aided Transit Scheduling*, volume 505 of *Lecture Notes in Economics and Mathematical systems*, pages 275–296. Springer, 2001.

[21] A. Schiewe, S. Albert, J. Pätzold, P. Schiewe, A. Schöbel, and J. Schulz. LinTim: An integrated environment for mathematical public transport optimization. Documentation. Technical Report 2018-08, Preprint-Reihe, Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen, http://lintim.math.uni-goettingen.de/, 2018.

[22] M. Schmidt. *Integrating Routing Decisions in Public Transportation Problems*, volume 89 of *Optimization and Its Applications.* Springer, 2014.

[23] M. Schmidt and A. Schöbel. The complexity of integrating routing decisions in public transportation models. In Thomas Erlebach and Marco Lübbecke, editors, *Proceedings of ATMOS10*, volume 14 of *OpenAccess Series in Informatics (OASIcs)*, pages 156–169, Dagstuhl, Germany, 2010. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[24] M. Schmidt and A. Schöbel. Timetabling with passenger routing. *OR Spectrum*, 37:75–97, 2015.

[25] P. Serafini and W. Ukovich. A mathematical model for periodic scheduling problems. *SIAM Journal on Discrete Mathematics*, 2(4):550–581, 1989.

Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
D - 37083 Göttingen

Telefon: 0551/394512
Telefax: 0551/393944

Email: trapp@math.uni-goettingen.de URL: http://www.num.math.uni-goettingen.de

Verzeichnis der erschienenen Preprints 2018

Number	Authors	Title
2018 - 1	Brimberg, J.; Schöbel, A.	When closest is not always the best: The distributed p-median problem
2018 - 2	Drezner, T.; Drezner, Z.; Schöbel, A.	The Weber Obnoxious Facility Location Model: A Big Arc Small Arc Approach
2018 - 3	Schmidt, M., Schöbel, A., Thom, L.	Min-ordering and max-ordering scalarization methods for multi-objective robust optimization
2018 - 4	Krüger, C., Castellani, F., Geldermann, J., Schöbel, A.	Peat and Pots: An application of robust multiobjective optimization to a mixing problem in agriculture
2018 - 5	Krüger, C.	Peat and pots: Analysis of robust solutions for a biobjective problem in agriculture
2018 - 6	Schroeder, P.W., John, V., Lederer, P.L., Lehrenfeld, C., Lube, G., Schöberl, J.	On reference solutions and the sensitivity of the 2d Kelvin–Helmholtz instability problem
2018 - 7	Lehrenfeld, C., Olshanskii, M.A.	An Eulerian Finite Element Method for PDEs in Time-dependent Domains
2018 - 8	Schöbel, A., Schiewe, A., Albert, S., Pätzold, J., Schiewe, P., Schulz, J.	LINTIM : An integrated environment for mathematical public transport optimization. Documentation
2018 - 9	Lederer, P.L. Lehrenfeld, C. Schöberl, J.	HYBRID DISCONTINUOUS GALERKIN METHODS WITH RELAXED H(DIV)-CONFORMITY FOR INCOMPRESSIBLE FLOWS. PART II

Number	Authors	Title
2018 - 10	Lehrenfeld, C.; Rave, S.	Mass Conservative Reduced Order Modeling of a Free Boundary Osmotic Cell Swelling Problem
2018 - 11	M. Friedrich, M. Hartl, A. Schiewe, A. Schöbel	System Headways in Line Planning
2018 - 12	M. Lübbecke, C. Puchert, P. Schiewe, A. Schöbel	Integrating line planning, timetabling and vehicle scheduling
2018 - 13	M. Friedrich, M. Müller-Hannemann, R. Rückert, A. Schiewe, A. Schöbel	Robustness as a Third Dimension for Evaluating Public Transport Plans
2018 - 14	J. Pätzold, A. Schiewe, A. Schöbel	Cost-Minimal Public Transport Planning
2018 - 15	S. Jäger, A. Schöbel	The blockwise coordinate descent method for integer programs
2018 - 16	A. Schiewe, P. Schiewe, M. Schmidt	The line planning routing game
2018 - 17	P. Schiewe, A. Schöbel	Periodic timetabling with integrated routing: An applicable approach