

A Model for Strategic Ridepooling and Its Integration with Line Planning

Lena Dittrich¹  

Department of Mathematics, RPTU University Kaiserslautern-Landau, Germany

Michael Rihlmann  

Fraunhofer Institute for Industrial Mathematics ITWM, Kaiserslautern, Germany

Sarah Roth  

Department of Mathematics, RPTU University Kaiserslautern-Landau, Germany

Anita Schöbel  

Department of Mathematics, RPTU University Kaiserslautern-Landau, Germany

Fraunhofer Institute for Industrial Mathematics ITWM, Kaiserslautern, Germany

Abstract

Ridepooling becomes more and more popular and providing comfortable and easy-to-use transportation (nearly as taxi rides) is known to motivate passengers to use public transport. In this paper we develop a model for strategic planning of ridepooling. Here we decide in which regions ridepooling should be offered and what capacities are needed, neglecting the operational details of dial-a-ride planning. We use this model for integrating ridepooling and line planning, and analyze the integrated model theoretically and numerically. Our experiments show the potential of the approach.

2012 ACM Subject Classification Applied computing → Transportation; Applied computing → Decision analysis

Keywords and phrases Multi-modal planning, Line plan, Ridepooling, Integrated models

Digital Object Identifier 10.4230/OASICS.ATMOS.2025.16

Funding *Lena Dittrich*: Funded by BMBF Project number 05M22UKB (SynphOnie).

Sarah Roth: Funded by BMBF Project number 05M22UKB (SynphOnie).

Acknowledgements We want to thank Prof. Dr.-Ing. Markus Friedrich and Julian Zimmer from Institut für Straßen- und Verkehrswesen at University of Stuttgart for providing us with some estimations for realistic input parameters of our models. We also want to thank Ricardo Reicherz who developed and implemented a heuristic for the classic dial-a-ride problem.

1 Motivation

Line planning is an important step in public transport optimization and numerous papers on the topic exist in the literature, ranging from the now 100 years old paper of [15] to very recent surveys [21, 22] summarizing the various models that have been developed for line generation, line selection, frequency setting, their integration, the different underlying routing models and robustness issues. Typically, in line planning, the task is to choose a set of lines and assign frequencies to the lines. Determining a timetable which is then implemented in practice is usually a separate problem.

In regional areas or outside the main traffic hours, schedule-based transportation with large buses is often replaced by smaller vehicles operating on a demand-based and more flexible basis. Such systems are often called *ridepooling*. Passengers can request a ridepooling

¹ corresponding author



© Lena Dittrich, Michael Rihlmann, Sarah Roth, and Anita Schöbel;
licensed under Creative Commons License CC-BY 4.0

25th Symposium on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2025).

Editors: Jonas Sauer and Marie Schmidt; Article No. 16; pp. 16:1–16:20



OpenAccess Series in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

vehicle for a trip from one location to another at a specific time. An operator decides about the routes of the ridepooling vehicles and schedules the specific requests of the passengers. There is a growing offer of ridepooling services across Europe. In Italy, for example, there are about ten bigger projects of on-demand services in the area between Milan and Florence. In Germany, the operator MOIA services cities like Hamburg and Hannover (c.f. [14]). In the UK, there are, e.g., the West Midlands Bus on Demand service [28] and the On-Demand Rideshare in Birmingham [1].

The operational details of scheduling the requests in ridepooling have been researched extensively. The underlying model is called the (online) dial-a-ride problem, which has already become popular in the 80s of the previous century often with an application to scheduling trips of elderly persons (e.g. [3], [12], see [25] for a survey), but is nowadays researched again (see e.g. [9], [11], and [29] for a survey), also with the perspective of future self-driving cars in mind. However, the strategic aspects of ridepooling, i.e., to identify where in the network ridepooling areas are needed and with how many vehicles they should be operated have to the best of our knowledge not been modeled and treated algorithmically so far. This is in particular important when designing a multi-modal transport system in which both, lines and ridepooling vehicles, operate. Then it is important to understand whether the two modes (regular bus lines and ridepooling) compete or complement each other.

The goal of this paper is to design a public transport system which consists of regularly operated bus lines as well as ridepooling services. This means we want to treat line planning and ridepooling simultaneously. As a first step towards this goal, we need a model for *strategic* ridepooling in which we disregard operational details, similarly to how line planning is a preparatory step to determining a timetable to implement. Our contribution is hence the following.

1. We develop a model for *strategic* ridepooling: Given some demand, which ridepooling areas are needed and with how many vehicles should they be operated?
2. We integrate line planning and strategic ridepooling to plan lines and ridepooling areas simultaneously.
3. We analyze the resulting model, its complexity status and give bounds on its objective function value.
4. Our experiments are promising: Problems can be solved for small and medium instances and show reasonable results. We discuss how optimal solutions and the runtime of the model vary depending on the input parameters.

The remainder of this paper is structured as follows: Based on the known cost model of line planning (described in Section 2.1) we develop a model for strategic ridepooling in Section 2.2. The ridepooling model requires a set of potential ridepooling areas with given vehicle frequencies. In Section 2.3 the construction of this input data is discussed.

The integrated ridepooling and line planning model is stated and analyzed in Section 3. In Section 4, results from numerical experiments are presented. We conclude in Section 5.

2 Strategic Planning of Ridepooling Areas

2.1 Line Planning

Line planning is well known in the literature, see [6] for a survey. Nevertheless, there are two reasons why we briefly introduce line planning here. First, the model we propose for strategic ridepooling is analogous to the basic line planning model. Second, in Section 3 we integrate the two tasks: planning lines and planning ridepooling services.

Let $\text{PTN} = (V, E)$ be a public transport network where V is a set of stops and the set of edges E consists of the direct links between pairs of stops. A *line* is a path in the PTN. Usually, it is assumed that a set of potential lines, the so-called *line pool* $\mathcal{L}^{\text{Pool}}$ is given. The goal is to choose a set $\mathcal{L}$ of lines from the line pool (*line selection*) and to assign frequencies to them such that every edge in the PTN on which passengers wish to travel is covered by a line. The frequency f_l is defined as the number of runs of line $l \in \mathcal{L}$ within a given planning interval T . The chosen lines $\mathcal{L}$ are called a *line plan*. The lines $l \in \mathcal{L}$ together with their frequencies f_l are called a *line concept*.

The line planning problem may involve many different constraints, and many different approximations of a reasonable objective function exist. For our investigations we use the basic version of the *cost model* of [5], see also [23]: Let $\mathcal{L}^{\text{Pool}}$ be a given set of potential lines. We assume that for every line $l \in \mathcal{L}^{\text{Pool}}$ the costs l-cost_l for operating it once from its first to its last station are known. These costs depend on the length of the line and on the time needed to drive from its first to its last station. As decision variables we use the frequencies f_l for all $l \in \mathcal{L}^{\text{Pool}}$. We require them to be non-negative and integral. A frequency of $f_l = 0$ means that line l is not selected in the line plan. The cost model of line planning reads as

$$\begin{aligned}
 \min \quad & \sum_{l \in \mathcal{L}^{\text{Pool}}} \text{l-cost}_l \cdot f_l \\
 \text{s.t.} \quad & L_e \leq \sum_{l \in \mathcal{L}^{\text{Pool}}: e \in l} f_l \leq U_e & \forall e \in E \\
 & f_l \in \mathbb{N}_0 & \forall l \in \mathcal{L}^{\text{Pool}}.
 \end{aligned} \tag{1}$$

It minimizes the sum of costs for operating the selected lines. The constraints ensure that there is the right amount of “frequency” along every edge. The lower bounds L_e usually stem from a passenger-oriented consideration: In a first step, passengers are routed along shortest paths in the PTN. This gives an amount of $w_e^{\min}$ passengers wishing to travel along edge e . Assuming that all vehicles have the same capacity l-cap the values L_e are chosen as $L_e := \left\lceil \frac{w_e^{\min}}{\text{l-cap}} \right\rceil$, where $w_e^{\min}$ is the number of passengers who wish to travel along this edge. The upper bounds U_e can represent capacity constraints to restrict the number of vehicles that pass along an edge in the planning period. The cost model as stated above has been used, e.g., in [27, 7, 17, 18].

For the usage in the current paper, we transform constraint (1) that currently bounds the frequency directly to instead bound the total capacity for transporting passengers on each edge. I.e., we use the passenger data $w_e^{\min}$ as lower bound instead of a lower bound on the frequencies. Instead of rounding the number of vehicles needed on every edge, we now require directly that all passengers are transported, i.e.,

$$\sum_{l \in \mathcal{L}^{\text{Pool}}: e \in l} \text{l-cap} \cdot f_l \geq w_e^{\min}$$

for every edge $e \in E$. Defining $w_e^{\max} := U_e \cdot \text{l-cap}$ hence results in the (equivalent) line planning model used in this paper:

Given a PTN with lower and upper bounds $w_e^{\min}, w_e^{\max}$ for every edge $e \in E$, a line pool $\mathcal{L}^{\text{Pool}}$ with costs l-cost_l for every line $l \in \mathcal{L}^{\text{Pool}}$, and a capacity l-cap for a vehicle operating on the lines, the *line planning model* is the following.

Line Planning (LC) (Finding a line concept)

$$\begin{aligned}
 \min \quad & \sum_{l \in \mathcal{L}^{\text{Pool}}} \text{l-cost}_l \cdot f_l & (\text{LC}) \\
 \text{s.t.} \quad & w_e^{\min} \leq \sum_{l \in \mathcal{L}^{\text{Pool}}: e \in l} \text{l-cap} \cdot f_l \leq w_e^{\max} & \forall e \in E \\
 & f_l \in \mathbb{N}_0 & \forall l \in \mathcal{L}^{\text{Pool}}.
 \end{aligned}$$

2.2 A new Model for Planning Ridepooling Areas

In order to combine line planning and ridepooling decisions on a strategic level, we need a model for ridepooling that leaves out the operational details such as which exact route which ridepooling vehicle travels at which time to cover which request. While these operational details are changing scenario-dependently from hour to hour and from day to day, we are interested in average passenger numbers as they are also used for line planning. In this section we develop a ridepooling model analogous to the cost model (LC) of line planning described in the previous section.

Let us again assume that a PTN= (V, E) is given and that we want to cover all the demand. However, here we want to cover the demand by ridepooling and not by bus lines. The main idea is to introduce *ridepooling areas* in which the same ridepooling provider offers a transportation service. This is in many regions common practice: only if the origin and the destination of a trip belong to a pre-specified area, a passenger is allowed to request a ridepooling vehicle for this trip. Such ridepooling areas may have different sizes and shapes. In a first step, we construct a pool $\mathcal{R}^{\text{Pool}}$ of such ridepooling areas. This is analogous to the idea of using a line pool $\mathcal{L}^{\text{Pool}}$ in the line planning problem. In line planning, the frequency f_l of a selected line l must be chosen so that there is enough capacity on every edge to transport all passengers. For ridepooling we proceed analogously and allow to adapt the supply to the demand. More precisely, for each ridepooling area in the pool we choose the number of ridepooling vehicles v_r that serve this area such that the demand along every edge is covered.

More formally, we define a *ridepooling area* r as a set of edges $r \subseteq E$. A ridepooling area r is called connected, if its induced graph $G_r := (V(r), r)$ is a connected graph where $V(r) := \{v \in V : v \text{ is incident with an edge } e \in r\}$ contains the endpoints of the edges in r . While connectedness is usually required in practice, it is technically not needed for our model. For each ridepooling area we determine the number of vehicles for this ridepooling area.

Note that a line (i.e. a path in the PTN) can be interpreted as a special case of a ridepooling area in which the vehicles move along pre-determined trips instead of moving around freely within their assigned area. However, there is a difference: If a *line* l runs with a specific frequency, say, $f_l = 3$, all edges $e \in l$ are visited three times per planning period. A *ridepooling vehicle*, on the other hand, does not need to visit all edges with the same frequency, but could visit a specific edge more often than others. Hence, a ridepooling area has a higher flexibility than a line. We take this into account by introducing a new parameter

$\alpha_{r,e}$ as the *vehicle frequency* of edge $e \in E$

for ridepooling area $r \in \mathcal{R}^{\text{Pool}}$. The vehicle frequency $\alpha_{r,e}$ says how often a single vehicle from ridepooling area $r \in \mathcal{R}^{\text{Pool}}$ visits edge $e \in E$ on average within the planning interval T . Since the demand changes from day to day, $\alpha_{r,e}$ can only be an approximation. The product $\alpha_{r,e} \cdot v_r$ gives us the number of times edge e is served within the planning period T . This is hence analogous to the frequency for the edges of a line.

■ **Table 1** Comparison of models for line planning and strategic ridepooling.

	Line planning	Ridepooling
Pool	Pool of lines $\mathcal{L}^{\text{Pool}}$	Pool of ridepooling areas $\mathcal{R}^{\text{Pool}}$
Supply	Frequencies of lines f_l	Number of vehicles in ridepooling areas v_r
Distribution on edges	Frequency f_l is the same for all edges $e \in l$	Vehicle frequencies $\alpha_{r,e}$ allow edge-dependent frequencies for edges $e \in r$
Cost	Depends linearly on the frequency of a line	Depends linearly on the number of vehicles

Given a PTN with lower and upper bounds $w_e^{\min}, w_e^{\max}$ for every edge $e \in E$, a pool of ridepooling areas $\mathcal{R}^{\text{Pool}}$ with costs r-cost_r for every ridepooling area $r \in \mathcal{R}^{\text{Pool}}$, values $\alpha_{r,e}$ for every $r \in \mathcal{R}^{\text{Pool}}, e \in E$ and a capacity r-cap for the ridepooling vehicles, the resulting strategic ridepooling model (RP) is the following.

Strategic Ridepooling (RP)

$$\begin{aligned}
 \min \quad & \sum_{r \in \mathcal{R}^{\text{Pool}}} \text{r-cost}_r \cdot v_r & (\text{RP}) \\
 \text{s.t.} \quad & w_e^{\min} \leq \sum_{r \in \mathcal{R}^{\text{Pool}}: e \in r} \alpha_{r,e} \cdot \text{r-cap} \cdot v_r \leq w_e^{\max} & \forall e \in E \\
 & v_r \in \mathbb{N}_0 & \forall r \in \mathcal{R}^{\text{Pool}}.
 \end{aligned}$$

Since lines can be seen as special ridepooling areas, (LC) is a special case of (RP):

► **Lemma 1.** *Model (LC) for line planning is a special case of the model for strategic ridepooling (RP).*

This directly clarifies the complexity status of (RP) (see Appendix C.1 for a formal proof).

► **Corollary 2.** *(RP) is NP-hard, even if we require that all ridepooling areas are simple paths.*

Table 1 compares the setting for line planning and strategic ridepooling.

2.3 Vehicle Frequencies

(RP) needs not only ridepooling areas as input but also vehicle frequencies $\alpha_{r,e}$. We now discuss how reasonable values for $\alpha_{r,e}$ can be found. Recall that for each ridepooling vehicle assigned to the ridepooling area $r \in \mathcal{R}^{\text{Pool}}$, $\alpha_{r,e}$ says how many times, on average, it traverses the edge $e \in r$ within the planning period T . Let r be a ridepooling area. We want to find values $\alpha_{r,e}$ such that:

- The values reflect the demand, i.e., edges with high demand should be visited more often (on average) than edges with low demand. The goal is to be as proportional to the traffic loads as possible.
- The values should also be realizable by the ridepooling vehicles in the following sense: There exists a set of tours, one for each vehicle, such that the number of visits of a vehicle at an edge is proportional to $\alpha_{r,e}$.

16:6 A Model for Strategic Ridepooling and Its Integration with Line Planning

In this section we show how values $\alpha_{r,e}$ which are realizable and rather proportional to the demand can be found. Since we do not know a priori how many vehicles will be assigned to ridepooling area r , the idea is to construct *one* feasible tour for the (average) vehicle which visits every edge approximately proportional to its traffic load.

Let d_e be the time needed to drive along the edge $e \in E$. Due to the definition of $\alpha_{r,e}$, namely the average number of visits of edge e in one period T , we receive

$$\sum_{e \in r} \alpha_{r,e} \cdot d_e \leq T \text{ for every ridepooling area } r \in \mathcal{R}^{\text{Pool}}.$$

An “optimal” set of values for $\alpha_{r,e}$ would be proportional to the number L_e of visits needed to transport all passengers, i.e., to

$$L_e := \left\lceil \frac{w_e^{\min}}{\text{r-cap}} \right\rceil.$$

Let us assume that a tour $\mathcal{C}$ that visits each edge e *exactly* L_e times exists. Then $\mathcal{C}$ has a duration D_r^L of

$$D_r^L = \sum_{e \in r} L_e \cdot d_e.$$

A single vehicle can drive tour $\mathcal{C}$ at most $\frac{T}{D_r^L}$ times within the planning period T . We hence have that every edge is traversed $\alpha_{r,e}^* := L_e \cdot \frac{T}{D_r^L}$ times within period T , i.e., the values $\alpha_{r,e}^*$ are proportional to L_e and are realizable if all vehicles of the ridepooling area always drive along the tour $\mathcal{C}$. This is hence the best possible case.

Let us first answer the question when such a tour that visits each edge *exactly* L_e times exists. To this end we look for a Euler cycle. We do not use the graph G_r of the ridepooling area r , but we build a multigraph G'_r which reflects the given demand structure. G'_r has the same node set as G_r , but the number of edges between two nodes corresponds to the minimum number of times a vehicle needs to drive between them to cover the passenger demand. Formally, let $G_r = (V(r), r)$ be given. Then we construct $G'_r = (V(r), E(r))$ with the same set of nodes as G_r and for every edge $e = (u, v) \in r$ of G_r we introduce $L_e := \left\lceil \frac{w_e^{\min}}{\text{r-cap}} \right\rceil$ edges between u and v in G'_r . Note that the values L_e are often assumed to be given right from the start in line planning. We receive the edge set

$$E(r) := \bigcup_{e=(u,v) \in r} \{(u,v)_1, \dots, (u,v)_{L_e}\}.$$

► **Lemma 3.** *Let $r \in \mathcal{R}^{\text{Pool}}$. Then there exists a tour which visits every edge $e \in r$ exactly $\alpha_{r,e}^*$ times if and only if every node in G'_r has even node degree.*

Proof. This is due to the well-known fact that an Euler tour in a graph exists if and only if all node degrees are even. ◀

Even node degrees need not always be the case in G'_r . We hence add further edges to the multigraph so that every node in the graph has an even node degree. This results in N_e edges per original edge $e \in r$. One (heuristic) way to determine N_e is to replace every edge $e \in r$ not by L_e but N_e edges, where

$$N_e := \begin{cases} L_e & \text{if } L_e \text{ is even,} \\ L_e + 1 & \text{if } L_e \text{ is odd.} \end{cases}$$

We call the resulting multigraph with even node degrees G''_r .

The multiplicities of the edges can then be transformed to the vehicle frequencies of ridepooling area r : The graph G_r'' hence contains an Euler cycle whose duration is

$$D_r^N = \sum_{e \in r} N_e \cdot d_e.$$

Within the planning period T a vehicle can make this trip $\frac{T}{D_r^N}$ times. This means that every edge $e \in r$ is traversed $N_e \cdot \frac{T}{D_r^N}$ times. Hence, for all $e \in r$ define $\alpha_{r,e} := N_e \cdot \frac{T}{D_r^N}$.

We summarize what we have obtained, for the proof see Appendix C.2.

- **Lemma 4.** *The values $\alpha_{r,e} := N_e \cdot \frac{T}{D_r^N}$ for all $r \in \mathcal{R}^{\text{Pool}}$, $e \in r$*
- *can be computed in polynomial time,*
 - *can be realized, i.e., there exists a set of tours for the vehicles such that the values $\alpha_{r,e}$ equal the number of average visits of a vehicle on the edge e in ridepooling area r within period T ,*
 - *and for all $e \in r$ they satisfy that*
 - $\alpha_{r,e} - \alpha_{r,e}^* \leq T \cdot L_e \cdot \frac{\sum_{e \in r: L_e \text{ odd}} d_e}{(D_r^L)^2 + D_r^L \cdot \sum_{e \in r: L_e \text{ odd}} d_e} \leq \alpha_{r,e}^*$
 - $\alpha_{r,e} - \alpha_{r,e}^* \geq \begin{cases} -\frac{T}{D_r^L} & \text{if } L_e \text{ is odd,} \\ 0 & \text{if } L_e \text{ is even.} \end{cases}$

Part 3 of Lemma 4 shows that the gap between $\alpha_{r,e}$ and $\alpha_{r,e}^*$ is bounded. The lower and upper bounds on the gap have an intuitive interpretation: The smallest possible gap for an edge e depends on whether or not the constructed multigraph G_r'' has “too many”, i.e., more than L_e edges corresponding to e . The upper bound on the gap depends on how many edges, overall throughout the whole ridepooling area r have too many corresponding edges in the multigraph. From the upper bound we additionally get that $\alpha_{r,e} \leq 2\alpha_{r,e}^*$.

The resulting values $\alpha_{r,e}$ have been evaluated and tested compared to solutions for the classic dial-a-ride problem showing very promising results: A simple insertion heuristic for the classic dial-a-ride problem has been implemented in LinTim ([20]). A more detailed description of the heuristic can be found in Appendix B.

The heuristic and the vehicle frequencies were tested on a five by five grid network for randomly generated demand and a maximum detour factor of 2, i.e., passengers might have to make detours, but these are restricted by the (travel) length of a shortest path, meaning the length of their trip is bounded by twice the length of a shortest path. The number of vehicles resulting from the vehicle frequencies $\alpha_{r,e}$ is only slightly lower as the number of vehicles required by the insertion heuristic. Assuming that the heuristic solutions are not always optimal, i.e., it is probably possible to serve the passengers with slightly fewer vehicles, the number of vehicles from the optimal solution of (RP) appears to be a good estimation.

This procedure for defining the vehicle frequencies $\alpha_{r,e}$ can be further generalized by, instead of replacing each edge $e \in r$ of a ridepooling area $r \in \mathcal{R}^{\text{Pool}}$ by N_e edges, replacing it by $M_e \in \mathbb{N}$ edges such that the resulting multigraph G_r' is connected and every node has even degree. Then, the resulting vehicle frequencies for all $e \in r$ are $\alpha_{r,e} = M_e \cdot \frac{T}{D_r^M}$ where $D_r^M = \sum_{e \in r} M_e \cdot d_e$. Also completely other methods to obtain values for the vehicle frequencies are possible, e.g., simulations.

Note that the direction of the passenger demand has been disregarded in the construction of the vehicle frequencies $\alpha_{r,e}$. However, passengers tend to have a preferred direction for their journey. The strategic ridepooling model (RP) can also be considered with a directed graph as the underlying network. The construction of the vehicle frequencies $\alpha_{r,e}$ works

analogously to the undirected case described above, using the well-known fact that a directed graph has a Euler cycle if and only if it is strongly connected and for each vertex the in-degree equals the out-degree.

We remark that it is allowed to include the same ridepooling area multiple times with different vehicle frequencies in the ridepooling pool $\mathcal{R}^{\text{Pool}}$. This offers more flexibility for the model and potentially improves the resulting solutions. However, it also increases the size of the problem, making it more difficult to solve.

3 Integrating Line planning and Strategic Planning of Ridepooling Areas

In Section 2, key differences between line-based and demand responsive public transport have been discussed. Ideally, when designing a public transport system, both modes are established in a way that utilizes their strengths and advantages such that they complement each other. We now propose a model for the integrated planning of a line network and ridepooling areas by combining (LC) with the strategic ridepooling model (RP).

In this combination we are again not interested in the specific routes the ridepooling vehicles should drive for a specific scenario, but we have the strategic aspects in mind. Let us mention that there exist a few papers integrating line planning with ridepooling in the operational planning in the following sense: Passengers' requests can be covered not only by ridepooling vehicles but passengers can also be routed with legs using the existing (and fixed) public transport system. This problem is called *Integrated Dial-A-Ride Problem* (IDARP), see [16].

In contrast to (IDARP) we do not assume the public transport system as fixed, but we aim at planning the lines and their frequencies simultaneously with setting up the ridepooling system. As already said we are furthermore not interested in the operational details, but plan strategically.

In our integrated model we minimize the overall costs, which are the sum of costs of the line network and the ridepooling vehicles. The demand may be covered by lines and by ridepooling areas. The combined problem (LC+RP) hence aims at determining the frequencies f_l and the number of vehicles v_r for all $l \in \mathcal{L}^{\text{Pool}}$ and all $r \in \mathcal{R}^{\text{Pool}}$.

$$\begin{aligned}
 \min \quad & \sum_{l \in \mathcal{L}^{\text{Pool}}} \text{l-cost}_l \cdot f_l + \sum_{r \in \mathcal{R}^{\text{Pool}}} \text{r-cost}_r \cdot v_r & (\text{LC+RP}) \\
 \text{s.t.} \quad & w_e^{\min} \leq \sum_{\substack{l \in \mathcal{L}^{\text{Pool}}: \\ e \in l}} \text{l-cap} \cdot f_l + \sum_{\substack{r \in \mathcal{R}^{\text{Pool}}: \\ e \in r}} \text{r-cap} \cdot \alpha_{r,e} \cdot v_r \leq w_e^{\max} \quad \forall e \in E \\
 & f_l, v_r \in \mathbb{N}_0 & \forall l \in \mathcal{L}^{\text{Pool}}, r \in \mathcal{R}^{\text{Pool}}.
 \end{aligned}$$

Both separate problems (LC) and (RP) are NP-complete, which shows the complexity of the integrated model (LC+RP).

► **Corollary 5.** (LC+RP) is NP-complete.

In the following we present an analysis of (LC+RP) including bounds and valid inequalities. We also state that by variable fixing we receive (LC) and (RP) again. These results lay the basis for future research on algorithms for solving (LC+RP). The proofs of the following results can all be found in Appendix C.3.

Since (LC+RP) is a combination of (LC) and (RP), feasible solutions for the separate problems immediately yield feasible solutions for the integrated problems, which, in turn, yield upper bounds for the optimal objective function value of (LC+RP).

► **Theorem 6.** *The optimal objective function values $z^{(LC)}$ and $z^{(RP)}$ of (LC) and (RP) yield upper bounds for the optimal objective function value $z^{(LC+RP)}$ of LC+RP:*

$$z^{(LC+RP)} \leq \min\{z^{(LC)}, z^{(RP)}\}$$

The demand on all edges needs to be covered sufficiently by lines or ridepooling areas. In some cases and for some edges, this yields a valid inequality.

► **Theorem 7.** *Let $e \in E$.*

- *If there exists a ridepooling area $r \in \mathcal{R}^{\text{Pool}}$ such that $e \in r$ and for all $r' \in \mathcal{R}^{\text{Pool}} \setminus \{r\}$ we have $e \notin r'$ and for all $l \in \mathcal{L}^{\text{Pool}}$ we have $e \notin l$, then in any feasible solution for (LC+RP) it holds that $v_r \geq \left\lceil \frac{w_e^{\min}}{r\text{-cap} \cdot \alpha_{r,e}} \right\rceil$.*
- *If there exists a line $l \in \mathcal{L}^{\text{Pool}}$ such that $e \in l$ and for all $l' \in \mathcal{L}^{\text{Pool}} \setminus \{l\}$ we have $e \notin l'$ and for all $r \in \mathcal{R}^{\text{Pool}}$ we have $e \notin r$, then in any feasible solution for (LC+RP) it holds that $f_l \geq \left\lceil \frac{w_e^{\min}}{l\text{-cap}} \right\rceil$.*

Finally, fixing part of a solution for (LC+RP) leads, again, to an instance of (LC+RP).

► **Theorem 8.** *Let $\mathcal{L}_{\text{fixed}}^{\text{Pool}} \subseteq \mathcal{L}^{\text{Pool}}$ and $f_l \in \mathbb{N}$ for all $l \in \mathcal{L}_{\text{fixed}}^{\text{Pool}}$, $\mathcal{R}_{\text{fixed}}^{\text{Pool}} \subseteq \mathcal{R}^{\text{Pool}}$ and $v_r \in \mathbb{N}$ for all $r \in \mathcal{R}_{\text{fixed}}^{\text{Pool}}$, such that for all $e \in E$ it holds that*

$$\sum_{l \in \mathcal{L}_{\text{fixed}}^{\text{Pool}} : e \in l} l\text{-cap} \cdot f_l + \sum_{r \in \mathcal{R}_{\text{fixed}}^{\text{Pool}} : e \in r} r\text{-cap} \cdot \alpha_{r,e} \cdot v_r \leq w_e^{\max}.$$

Then, the remaining problem of optimizing f_l and v_r for $l \in \mathcal{L}^{\text{Pool}} \setminus \mathcal{L}_{\text{fixed}}^{\text{Pool}}$ and $r \in \mathcal{R}^{\text{Pool}} \setminus \mathcal{R}_{\text{fixed}}^{\text{Pool}}$ is an instance of (LC+RP).

This result yields that fixing f_l leads to a problem of type (RP) and fixing v_r leads to a problem of type (LC). This means, iterative solution approaches are possible.

► **Corollary 9.** *The following holds in the same setting as Theorem 8:*

- *Fixing the ridepooling variables v_r for all $r \in \mathcal{R}^{\text{Pool}}$ such that $\sum_{r \in \mathcal{R}^{\text{Pool}} : e \in r} r\text{-cap} \cdot \alpha_{r,e} \cdot v_r \leq w_e^{\max}$ leads to an instance of (LC).*
- *Fixing the line frequency variables f_l for all $l \in \mathcal{L}^{\text{Pool}}$ such that $\sum_{l \in \mathcal{L}^{\text{Pool}} : e \in l} l\text{-cap} \cdot f_l \leq w_e^{\max}$ leads to an instance of (RP).*

4 Numerical Experiments

The model (LC+RP) has been implemented in the software toolbox LinTim ([20]) and tested on different instances and for different input parameters. The MIP formulations were solved using Gurobi 11.0.1 [10]. In this section, first an example is introduced and its solution is described in Section 4.1. Then interpretations and observations about properties of solutions are discussed in Section 4.2. Different experiments and results regarding the runtime of the model are presented in Section 4.3.

We test our model on three different instances of varying size. All three networks are still smaller than real-life instances, and further tests on larger networks are conceivable.

■ **Table 2** Input parameters for the example in Section 4.1.

l-cost_{fixed}	l-cost_{dist}	r-cost	l-cap	r-cap
8	4	60	42	6

For our experiments, we need a method for line pool generation. Line Pool generation is its own area of research (see [8] and references therein) and the underlying line pool can have a big impact on the quality of the resulting solution. For our experiments we want a line pool that is large enough to offer sufficient flexibility for the model but also not so large as to render the model too complex to solve. In all experiments we used the same method for line pool generation which we describe next: First, a basic line pool is computed with the LinTim-method `k_shortest_paths` [20, p.33]: For each OD-pair we compute k shortest paths with $k = 3$. Afterwards only those paths are added to the line pool, that are not already contained in other paths. When planning a line service and a ridepooling service simultaneously it can be beneficial to also have shorter versions of the lines in the line pool, to allow peripheral areas to be covered more by ridepooling areas. Therefore, in Sections 4.1 and 4.2 in which we provide an example solution and an analysis of properties of solutions, we added for each line also the copies of the line to the line pool, where the first and the last edge are trimmed on each of the two ends separately and also on both ends. In Section 4.3, where we give an analysis of the runtime of (LC+RP), copies of lines where not just one but up to two edges have been trimmed on one or both ends of the lines have been added to the line pool of `k_shortest_paths`, resulting in an even larger line pool. The cost of each line is computed by an affine-linear function:

$$\text{l-cost}_l = \text{l-cost}_{fixed} + \sum_{e \in l} \text{l-cost}_{dist} \cdot d_e.$$

4.1 Example

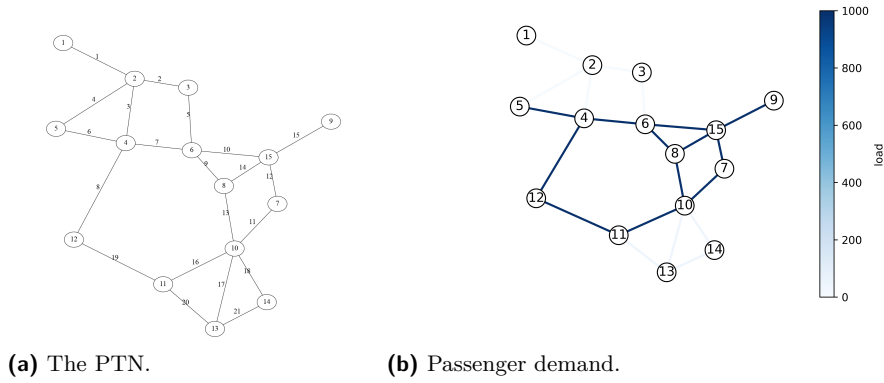
For this example we use the LinTim-dataset `Mandl`, based on [13], which is depicted in Figure 1a.

The ridepooling pool $\mathcal{R}^{\text{Pool}}$ for this example contains all connected subgraphs with at most five edges. The ridepooling vehicle costs are the same for each ridepooling area, i.e., $\text{r-cost}_r = \text{r-cost}$ for all $r \in \mathcal{R}^{\text{Pool}}$.

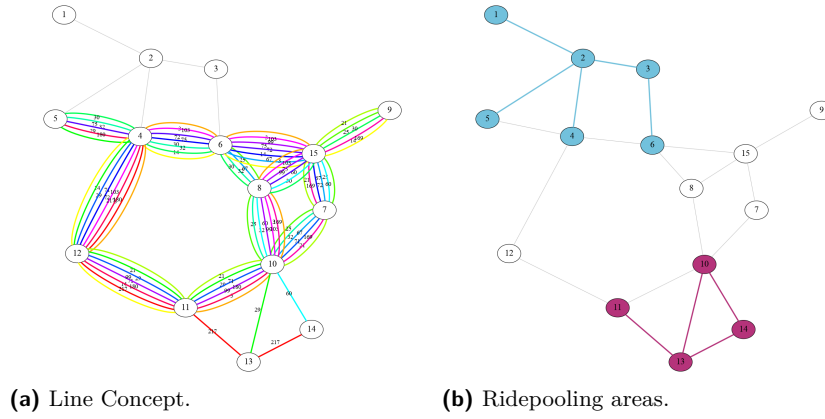
Realistic input parameters, especially for the costs, can be difficult to estimate, since they can vary greatly based on assumptions about the vehicles utilized in practice: electric vehicles have lower operating costs than those with a combustion engine but are, currently, more expensive. In the future, it is likely that public transport vehicles will drive autonomously, rendering the driver and hence also their salary unnecessary. The input parameters used for the example in this section can be seen in Table 2.

The passenger demand is represented by edge loads depicted in Figure 1b. There are 10 passengers who wish to travel on the edges 1 through 5 and 50 passengers for the edges 17, 18, 20 and 21. The remaining edges have a much higher passenger demand of 1000 passengers per edge. Such a structure of passenger demand, while artificially created and rather simple, is not very far from many realistic scenarios: Large cities often have a high passenger volume while neighboring suburban areas typically have lower passenger demand.

The resulting optimal solution is shown in Figure 2. The area of the network with very high demand is covered exclusively by lines, while the two smaller areas with lower demand are supplied by ridepooling services. One of the areas is covered only by a ridepooling area,



■ **Figure 1** The network Mandl for the example in Section 4.1.



■ **Figure 2** Optimal solution of (LC+RP) on the Mandl dataset.

while the other also has some line service. It shows that, while line-based public transport is very useful for areas with high passenger demand, the smaller, cheaper and more flexible ridepooling vehicles are a useful alternative wherever there are not enough passengers to fill a whole line vehicle.

4.2 Ridepooling Percentage

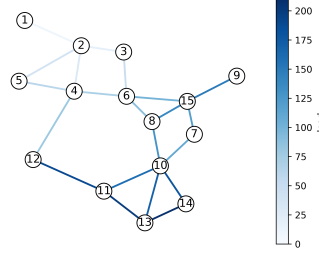
We want to evaluate solutions of (LC+RP), in particular we are interested to visualize in which parts of the network ridepooling is offered instead of classic lines. To this end, we need a measure for the amount of capacity offered by ridepooling. We introduce the *ridepooling percentage* for a solution (f_l, v_r) of (LC+RP).

First, for each edge $e \in E$ there is a certain amount of capacity in the chosen lines and ridepooling areas that contain the edge:

- Line-Cap $_e := \sum_{l: e \in l} f_l \cdot \text{l-cap}$
- RP-Cap $_e := \sum_{r: e \in r} \alpha_{re} \cdot v_r \cdot \text{r-cap}$

Then the ridepooling percentage on the edge is defined as follows:

$$\text{RP-percentage}_e := \frac{\text{RP-Cap}_e}{\text{Line-Cap}_e + \text{RP-Cap}_e}$$



■ **Figure 3** Edge Loads.

For the whole network, the ridepooling percentage is defined analogously:

$$\text{RP-percentage} = \frac{\sum_{e \in E} \text{RP-Cap}_e}{\sum_{e \in E} \text{Line-Cap}_e + \text{RP-Cap}_e}$$

The ridepooling percentage can be used to visualize (optimal) solutions computed for different input parameters. We observe that the ridepooling percentage and hence the shape of the optimal solution strongly depends on the ratio of the costs of the ridepooling vehicles and the lines.

Figure 4 shows the ridepooling percentage on each edge of the network. The cost of the ridepooling vehicles varies, while all other input parameters stay the same. The ridepooling pool contains all connected subgraphs of the network with at most five edges, the line pool, as in Section 4.1, was computed using the `k_shortest_paths` method and then adapted by adding subpaths of already existing lines. The passenger demand is given by edge loads, which are shown in Figure 3. The line costs are computed using $\text{l-cost}_{\text{fixed}} = 10$ and $\text{l-cost}_{\text{dist}} = 5$. The line vehicle capacity is $\text{l-cap} = 60$ and the ridepooling vehicle capacity is $\text{r-cap} = 5$. We vary the costs for the ridepooling vehicle using 10, 20, 30, and 40 as input parameters.

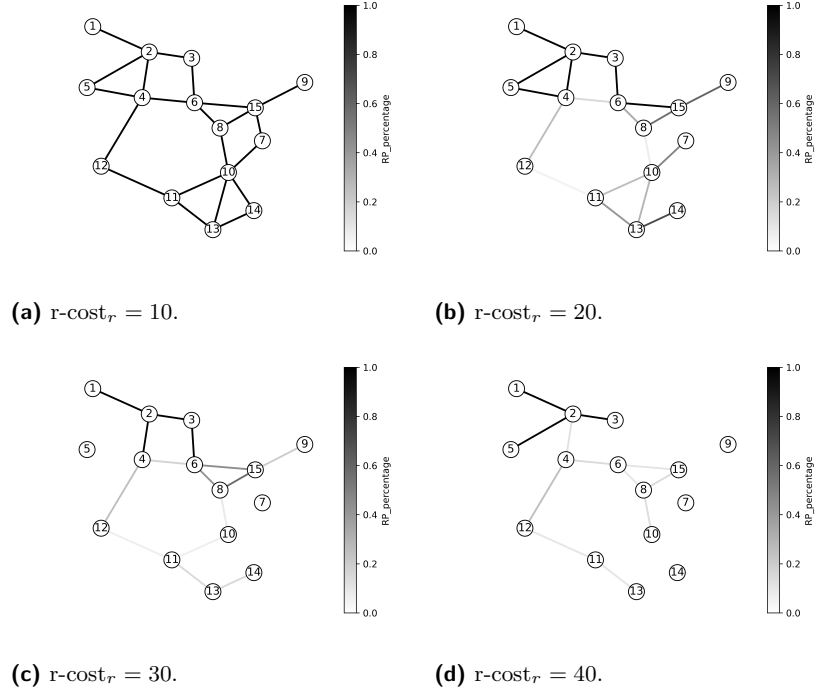
All solutions depicted in Figure 4 are either optimal or have an optimality gap of less than 3%.

For increasing the cost of the ridepooling vehicles, we observe that the ridepooling percentage overall decreases and fewer edges are covered by ridepooling. Typically, ridepooling vehicles are smaller than line vehicles such as buses. As the cost of the ridepooling vehicles increases, only areas of the network with low demand are covered by ridepooling areas: wherever there is not enough passenger demand to justify introducing a large and expensive line vehicle, on-demand transport with its smaller and cheaper vehicles is a good alternative.

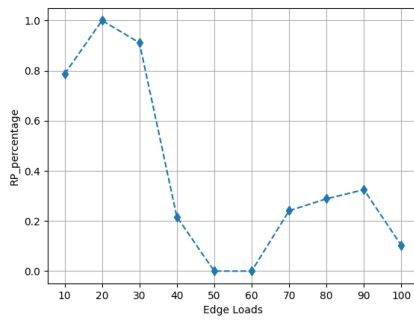
Ridepooling can also be useful as an addition in areas with high passenger demand. For instance, if there already is a line network but in some areas the passenger demand is exceptionally high, then ridepooling can be introduced to supplement the line based public transport system. This effect can be observed in the solution with $\text{r-cost} = 20$ in Figure 4.

Generally, the more passengers there are the more useful large line vehicles become, since the cost per passenger is small if the demand is high enough. Figure 5 investigates what happens when we increase the demand. It shows the ridepooling percentage when increasing the edge loads, where we assume the same amount of passenger demand on each edge of the Mandl network (see Figure 7b), and the same input parameters as for the solutions in Figure 4. Not all depicted solutions are optimal, but all have an optimality gap below 5%.

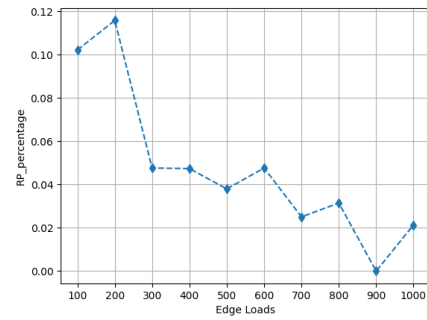
Basically, we observe that a higher amount of passenger demand leads to a lower the overall ridepooling percentage. However, this is not a strict rule. The capacity of the line vehicles is $\text{l-cap} = 60$, which could explain some of the non-monotonicity of the graphs in



■ **Figure 4** The effect of the cost of ridepooling vehicles on the ridepooling percentage of each edge.



(a) Ridepooling percentage with respect to edge loads of 10, 20, $\dots$, 100.



(b) Ridepooling percentage with respect to edge loads of 100, 200, $\dots$, 1000.

■ **Figure 5** RP-percentage for constant edge loads.

Figure 5: whenever dividing the edge loads by the line vehicle capacity $l\text{-cap}$ leaves a large remainder, the ridepooling percentage is smaller. When the remainder is small, then instead of using line vehicles with empty seats, the passengers can be transported using smaller ridepooling vehicles instead and the ridepooling percentage becomes larger. This effect is known as step-fixed costs.

4.3 Runtime

To analyze the impact of ridepooling (RP) to the integrated line planning and ridepooling problem (LC+RP) we use three instances varying in size and three different sized ridepooling pools. The dataset **Toy** is a small artificial instance consisting of 8 stops, while the larger datasets **Mandl** [13] and **Sioux-Falls** [26] are based on real-world data with 15 stops and 24 stops, respectively. The PTNs of the three instances are shown in Figure 7 in Appendix A.

With each dataset we performed four computations: As a benchmark, we solve the line planning problem (LC). The integrated line planning and ridepooling problem (LC+RP) is solved with a small, a medium-sized and a large pool of ridepooling areas $\mathcal{R}^{\text{Pool}}$. We use the same line pool $\mathcal{L}^{\text{Pool}}$ throughout the four runs.

The ridepooling pool includes connected subgraphs of the PTN induced by at least l_N and at most u_N nodes. Table 3 shows the values of l_N and u_N for the three considered PTNs. For each node v of the PTN one induced connected subgraph containing v with exactly i nodes for $l_N \leq i \leq u_N$ is chosen randomly. The ridepooling pools for a fixed dataset are constructed such that smaller ridepooling pools are subsets of the larger pools. For a more detailed description of the algorithms for the generation of potential lines and potential ridepooling areas in these experiments, see our Software Library **LinTim** [20].

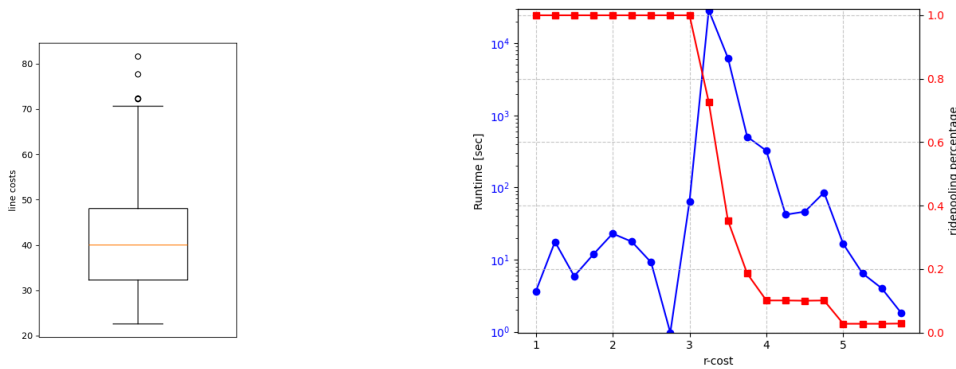
■ **Table 3** Number of nodes of the induced subgraphs used as areas in the different sized ridepooling pools.

	small		medium		large	
	l_N	u_N	l_N	u_N	l_N	u_N
Toy	3	3	3	4	2	4
Mandl	3	5	3	7	3	10
Sioux-Falls	3	8	3	10	2	12

The runtime experiments were performed on an Intel(R) Xeon(R) Gold 6240R CPU @ 2.40GHz with 380GB memory. Table 4 summarizes the runtime results together with the sizes of the used line pool and ridepooling pool.

■ **Table 4** Runtime results in seconds for different sized datasets and ridepooling pools. The largest instance was not solved to optimality within the time limit of 8 hours. The table shows the remaining optimality gap.

	Toy			Mandl			Sioux-Falls		
	$ \mathcal{L}^{\text{Pool}} $	$ \mathcal{R}^{\text{Pool}} $	time	$ \mathcal{L}^{\text{Pool}} $	$ \mathcal{R}^{\text{Pool}} $	time	$ \mathcal{L}^{\text{Pool}} $	$ \mathcal{R}^{\text{Pool}} $	time
(LC)	43	—	0.17	625	—	0.41	1955	—	6.51
(LC+RP) small		6	0.19		41	52.43		135	2699.0
(LC+RP) med		12	1.41		67	75.97		183	9226.25
(LC+RP) large		19	1.42		107	912.14		249	0.297%



(a) Boxplot of the line costs in the line pool.

(b) Runtimes and ridepooling percentages for varying costs of a ridepooling vehicle.

■ **Figure 6** The runtime of (LC+RP) depends on the ratio of line and ridepooling costs.

We observe that runtime increases significantly when integrating ridepooling to the line planning problem. Considering larger ridepooling pools leads to longer runtimes. The integrated problem with the large ridepooling pool on the *Sioux-Falls* datasets was not solved to optimality within the time limit of 8 hours. Table 4 shows the remaining optimality gap. Note that the cardinalities of the ridepooling pools and the line pools created by the methods described above do not scale with the same rate, i.e. the ratio of $|\mathcal{R}^{\text{Pool}}|$ and $|\mathcal{L}^{\text{Pool}}|$ decreases for larger datasets. However, considering larger ridepooling pools would lead to even longer runtimes.

We also found that the runtime of the integrated line planning and ridepooling problem highly depends on the relation of the costs of lines determined by $\text{l-cost}_{\text{fixed}}$ and $\text{l-cost}_{\text{dist}}$ to the costs of a ridepooling vehicle r-cost . To demonstrate the impact, we did multiple runs on the dataset *Mandl* with the medium-sized pool and varying costs of ridepooling vehicles. The line pool and the costs of the lines were the same throughout all runs. Figure 6 shows on the left a boxplot of the line costs in the line pool and on the right a plot of the runtime and the ridepooling percentage. Note that the runtime is depicted on a logarithmic scaled axis. The run for $\text{r-cost} = 3.25$ could not be solved within the time limit of 8 hours. The remaining gap was 0.0267%.

For low values of r-cost the model was solved very fast and all demand was covered by ridepooling. High values of r-cost also lead to fast runtimes and solutions with a ridepooling percentage of nearly 0. In both cases, the demand is covered (almost) only with one of the two service types. Those types of solutions seem to be found very fast by the model. If the ratio of the ridepooling and line costs is balanced, there are a lot more possible combinations of the services to discover and thus the solution process of (LC+RP) is very time consuming. This seems to be the reason for the curve starting with a small runtime, increasing up to a maximum and then decreasing again. The overall curve is not concave due to local disturbances which most likely stem from the effect of step-fixed costs already described at the end of Section 4.2.

5 Conclusion and further research

In this paper we developed a model for strategic ridepooling which then could be combined with line planning. We analyze the integrated model theoretically and experimentally. The results are promising. As future research we mention the following topics:

First, there is plenty of literature to generate line pools, e.g., [8]. For (RP) and for (LC+RP) we need a pool of ridepooling areas. A first step of future research is to develop criteria for good ridepooling areas and use them to design algorithms that construct such a pool which is reasonable in its size but contains promising areas.

Second, the model (LC+RP) relies on the cost model of line planning. In this model origin-destination (OD)-pairs are not used, but only their resulting traffic loads $w_e^{\min}$. This comes with drawbacks: First, transfers cannot be counted, and second, passengers' paths are considered as fixed although they depend on the lines and the ridepooling areas. There exist extensions to more realistic line planning models in which transfers are accounted for (e.g., the direct travelers approach [4]) and in which routing of passengers is integrated (see [2, 24, 19]). In future research we plan to use also such models for integrating line planning and ridepooling.

More experiments on larger and even real-life instances are a topic of ongoing research. We are also currently developing a column generation approach for LC+RP to generate both lines and ridepooling areas within the solution process. Other heuristic approaches for larger instances could be developed. Furthermore, similarly to how a line concept is used to determine a timetable in a following step, the operational aspects of a ridepooling service as planned by LC+RP could be evaluated by applying algorithms for the classic dial-a-ride problem.

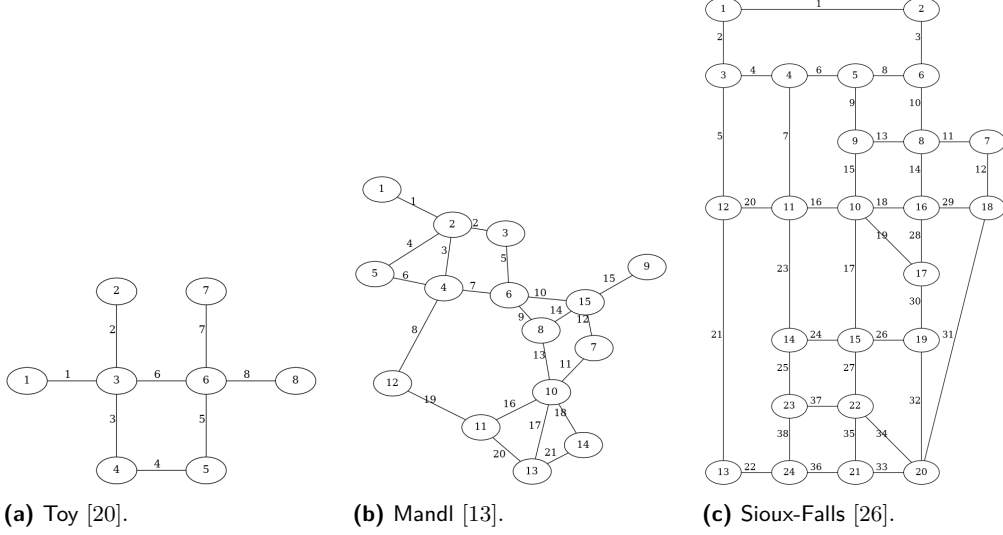
Finally, a further aspect of future research is to consider not only the two modes *bus transportation* and *ridepooling* but also other modes such as metro transportation, car transportation or even bikes, see [30] for some first ideas on this.

References

- 1 Birmingham on Demand. Ride for less with Birmingham On-Demand. <https://city.ridewithvia.com/birmingham>. Accessed: 2025-07-01.
- 2 R. Borndörfer, M. Grötschel, and M. E. Pfetsch. A column generation approach to line planning in public transport. *Transportation Science*, 41:123–132, 2007. doi:10.1287/TRSC.1060.0161.
- 3 Ralf Borndörfer, Martin Grötschel, Fridolin Klostermeier, and Christian Küttner. Berliner Telebussystem bietet Mobilität für Behinderte. *Der Nahverkehr*, 1:20–22, 1997.
- 4 M.R. Bussieck, P. Kreuzer, and U.T. Zimmermann. Optimal lines for railway systems. *European Journal of Operational Research*, 96(1):54–63, 1997.
- 5 M.T. Claessens, N.M. van Dijk, and P.J. Zwaneveld. Cost optimal allocation of rail passenger lines. *European Journal on Operational Research*, 110:474–489, 1998. doi:10.1016/S0377-2217(97)00271-3.
- 6 Javier Durán-Micco and Pieter Vansteenwegen. A survey on the transit network design and frequency setting problem. *Public Transport*, 14(1):155–190, 2022. doi:10.1007/S12469-021-00284-Y.
- 7 M. Friedrich, M. Hartl, A. Schiewe, and A. Schöbel. Integrating Passengers' Assignment in Cost-Optimal Line Planning. In *ATMOS 2017*, volume 59 of *OpenAccess Series in Informatics (OASICS)*, pages 1–16, 2017. doi:10.4230/OASICS.ATMOS.2017.5.
- 8 P. Gattermann, J. Harbering, and A. Schöbel. Line pool generation. *Public Transport*, 9(1-2):7–32, 2017. doi:10.1007/s12469-016-0127-x.
- 9 Konstantinos Gkiotsalitis and A. Nikolopoulou. The multi-vehicle dial-a-ride problem with interchange and perceived passenger travel times. *Transportation research part C: emerging technologies*, 156:104353, 2023.
- 10 Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2024. URL: <https://www.gurobi.com>.
- 11 Andy Ham. Dial-a-ride problem: mixed integer programming revisited and constraint programming proposed. *Engineering Optimization*, 55(2):257–270, 2023.

- 12 Jang-Jei Jaw, Amedeo R. Odoni, Harilaos N. Psaraftis, and Nigel H.M. Wilson. A heuristic algorithm for the multi-vehicle advance request dial-a-ride problem with time windows. *Transportation Research Part B: Methodological*, 20(3):243–257, 1986.
- 13 C.E. Mandl. *Applied Network Optimization*. Academic Press, London, UK, 1979.
- 14 Moia. Ridepooling in europe. <https://www.moia.io/en/blog/ridepooling-in-europe>. Accessed: 2025-07-01.
- 15 A. Patz. Die richtige Auswahl von Verkehrslinien bei großen Strassenbahnnetzen. *Verkehrstechnik*, 50/51, 1925. (in German).
- 16 Marcus Posada, Henrik Andersson, and Carl H. Häll. The integrated dial-a-ride problem with timetabled fixed route service. *Public Transport*, 9(1-2):217–241, November 2017. doi:10.1007/s12469-016-0128-9.
- 17 Güvenç Şahin, Amin Ahmadi Digehehara, Ralf Borndörfer, and Thomas Schlechte. Multi-period line planning with resource transfers. *Transportation Research Part C: Emerging Technologies*, 119:102726, 2020.
- 18 A. Schiewe, A. Schöbel, and L. Sieber. Line planning for different demand periods. *Operations Research Forum*, 4:92, 2023. doi:10.1007/s43069-023-00268-7.
- 19 P. Schiewe. *Integrated Optimization in Public Transport Planning*, volume 160 of *Optimization and Its Applications*. Springer, 2020. doi:10.1007/978-3-030-46270-3.
- 20 Philine Schiewe, Anita Schöbel, Sven Jäger, Sebastian Albert, Christine Biedinger, Thorsten Dahlheimer, Vera Grafe, Olli Herrala, Klara Hoffmann, Sarah Roth, Alexander Schiewe, Moritz Stinzendörfer, and Reena Urban. Documentation for lintim 2024.12, 2024. URL: <https://nbn-resolving.de/urn:nbn:de:hbz:386-kluedo-85839>.
- 21 M. Schmidt and A. Schöbel. Modeling and optimizing transit lines. In S. Parragh and T.V. Woensel, editors, *Handbook on Transport Modeling*, Research Handbooks in Transportation Studies, chapter 16. Edward Elgar Publishing, 2025.
- 22 M. Schmidt and A. Schöbel. Passenger-oriented and robust transit line planning. In S. Parragh and T.V. Woensel, editors, *Handbook on Transport Modeling*, Research Handbooks in Transportation Studies, chapter 17. Edward Elgar Publishing, 2025.
- 23 A. Schöbel. Line planning in public transportation: models and methods. *OR Spectrum*, 34(3):491–510, 2012. doi:10.1007/S00291-011-0251-6.
- 24 A. Schöbel and S. Scholl. Line planning with minimal travel time. In *5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS'05)*, Open Access Series in Informatics (OASICS), pages 1–16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2006. doi:10.4230/OASICS.ATMOS.2005.660.
- 25 Marius M. Solomon and Jacques Desrosiers. Survey paper – time window constrained routing and scheduling problems. *Transportation science*, 22(1):1–13, 1988. doi:10.1287/TRSC.22.1.1.
- 26 Ben Stabler. Sioux Falls – Github, 2018. URL: <https://github.com/bstabler/TransportationNetworks/tree/master/SiouxFalls>.
- 27 L. M. Torres, R. Torres, R. Borndörfer, and M. E. Pfetsch. Line planning on paths and tree networks with applications to the quito trolebús system. In Matteo Fischetti and Peter Widmayer, editors, *8th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'08)*, Open Access Series in Informatics (OASICS), pages 1–13, Dagstuhl, Germany, 2008. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/OASICS.ATMOS.2008.1583.
- 28 Transport for West Midlands. West midlands bus on demand. <https://www.tfwm.org.uk/plan-your-journey/ways-to-travel/buses-in-the-west-midlands/on-demand-buses-in-the-west-midlands/>. Accessed: 2025-07-01.
- 29 P. Vansteenwegen, L. Melis, D. Aktaş, B. D. G. Montenegro, F. S. Vieira, and K. Sörensen. A survey on demand-responsive public bus systems. *Transportation Research Part C: Emerging Technologies*, 137:103573, 2022.
- 30 J. Zimmer, M. Friedrich, and A. Schöbel. Suitability of different transport means as a function of travel demand: Pareto-optimal solutions (in german). *Straßenverkehrstechnik*, pages 796–805, 2024. doi:10.53184/SVT10-2024-3.

A Graphics



■ **Figure 7** PTNs of the LinTim instances Toy, Mandl, and Sioux-Falls.

B Dial-a-ride Heuristic

In a network, passenger requests are given by their origin, destination and desired departure time. The algorithm then goes through the passenger requests one by one and checks for every already operating vehicle if the request can be inserted into its tour to serve the corresponding passengers. A new request can be inserted into the trip of a vehicle if this insertion does not violate any given constraints, such as the maximum detour factor or waiting time for the passengers or the capacity of the vehicle. If a vehicle is found where the new request can be inserted, then this is done. Otherwise, a new vehicle is added. This insertion heuristic must not lead to optimal solutions, but provides an estimate for the number of vehicles necessary to serve a given amount of passenger demand.

The constants $\alpha_{r,e}$, determined by the procedure described in Section 2.3, have been tested in comparison to this insertion heuristic as follows: The passenger requests have been transformed into so-called OD-data, disregarding the desired departure times and simply computing for each pair of nodes $u, v \in V$ the number of passengers who wish to travel from u to v . Traffic loads $w_e^{\min}$ for the edges $e \in E$ are then generated by routing all passengers along shortest paths in the underlying network. Then, having computed $\alpha_{r,e}$ as described before with a single ridepooling area $r = E$, the optimal solution for (RP) is the following:

$$v_r = \max_{e \in E} \left\lceil \frac{w_e^{\min}}{\alpha_{r,e} \cdot \text{r-cap}} \right\rceil.$$

The resulting number of vehicles can be compared to the number of vehicles required in the solution from the insertion-heuristic.

C Proofs

C.1 From Section 2.2

Proof of Lemma 1. Given an instance of the line planning problem, we set

- $\mathcal{R}^{\text{Pool}} := \mathcal{L}^{\text{Pool}}$ since every line $l \in \mathcal{L}^{\text{Pool}}$ can be interpreted as a (connected) ridepooling area,
- $\text{r-cost}_r := \text{l-cost}_r$ for all $r \in \mathcal{R}^{\text{Pool}}$,
- $\text{r-cap} := \text{l-cap}$, and
- $\alpha_{e,r} := 1$ for all $e \in E$, $r \in \mathcal{R}^{\text{Pool}}$.

We receive an instance of the ridepooling problem (RP) in which the variables v_r correspond to the frequencies f_r . I.e., the resulting program of type (RP) is exactly the line planning problem (LC). ◀

C.2 From Section 2.3

Proof of Lemma 4.

1. The values $\alpha_{r,e}$ can be immediately computed using the procedure described in Section 2.3. The number of constants that need to be computed is $\sum_{r \in \mathcal{R}^{\text{Pool}}} |r|$ which has the following polynomial upper bound:

$$\sum_{r \in \mathcal{R}^{\text{Pool}}} |r| \leq \sum_{r \in \mathcal{R}^{\text{Pool}}} |E| = |\mathcal{R}^{\text{Pool}}| |E|.$$

2. For every vehicle of ridepooling area $r \in \mathcal{R}^{\text{Pool}}$ we know that an Euler tour in G_r'' exists which realizes exactly the values $\alpha_{r,e}$. If all vehicles drive this tour, the result follows.
3. Clearly, $0 \leq N_e - L_e \leq 1$, i.e., $D_r^N \geq D_r^L$. Note that:

$$D_r^L \cdot D_r^N = D_r^L \cdot \left(D_r^L + \sum_{e \in r: L_e \text{ odd}} d_e \right) = (D_r^L)^2 + D_r^L \cdot \sum_{e \in r: L_e \text{ odd}} d_e$$

Now, consider the gap between $\alpha_{r,e}$ and $\alpha_{r,e}^*$:

$$\alpha_{r,e} - \alpha_{r,e}^* = L_e \cdot \frac{T}{D_r^L} - N_e \cdot \frac{T}{D_r^N} = T \cdot \frac{L_e \cdot D_r^N - N_e \cdot D_r^L}{D_r^N \cdot D_r^L}$$

On the one hand that gives

$$\begin{aligned} \alpha_{r,e} - \alpha_{r,e}^* &\leq T \cdot \frac{L_e \cdot D_r^N - L_e \cdot D_r^L}{D_r^N \cdot D_r^L} \\ &= T \cdot L_e \cdot \frac{D_r^N - D_r^L}{D_r^N \cdot D_r^L} \\ &= T \cdot L_e \cdot \frac{\sum_{e \in r} N_e \cdot d_e - L_e \cdot d_e}{D_r^L \cdot D_r^L} \\ &= T \cdot L_e \cdot \frac{\sum_{e \in r: L_e \text{ odd}} d_e}{(D_r^L)^2 + D_r^L \cdot \sum_{e \in r: L_e \text{ odd}} d_e} \leq \frac{T}{D_r^L} L_e = \alpha_{r,e}^* \end{aligned}$$

and on the other hand we get

$$\begin{aligned} \alpha_{r,e} - \alpha_{r,e}^* &\geq T \cdot \frac{L_e \cdot D_r^N - N_e \cdot D_r^L}{D_r^N \cdot D_r^L} \\ &= T \cdot \frac{L_e - N_e}{D_r^L} = \begin{cases} -\frac{T}{D_r^L} & \text{if } L_e \text{ is odd,} \\ 0 & \text{if } L_e \text{ is even.} \end{cases} \end{aligned}$$

C.3 From Section 3

Proof of Theorem 6. Feasible solutions for LC and (RP) can be transformed into feasible solutions for (LC+RP):

- Let $(f_l)_{l \in \mathcal{L}^{\text{Pool}}}$ be a feasible solution for (LC), then $((f_l)_{l \in \mathcal{L}^{\text{Pool}}}, (0)_{r \in \mathcal{R}^{\text{Pool}}})$ is feasible for (LC+RP) with the same objective function value.
- Let $(v_r)_{r \in \mathcal{R}^{\text{Pool}}}$ be a feasible solution for (RP), then $((0)_{l \in \mathcal{L}^{\text{Pool}}}, (v_r)_{r \in \mathcal{R}^{\text{Pool}}})$ is also feasible for (LC+RP) with the same objective function value. ◀

Proof of Theorem 7. We only show the first case, the proof of the second case is analogous.

As r is the only ridepooling area or line that contains e , it is required that $\text{r-cap} \cdot \alpha_{r,e} v_r \geq w_e^{\min}$ satisfies the constraint for the edge e .

$$\Rightarrow v_r \geq \frac{w_e^{\min}}{\text{r-cap} \cdot \alpha_{r,e}}$$

$\Rightarrow$ Since v_r needs to be integer, we can conclude that $v_r \geq \left\lceil \frac{w_e^{\min}}{\text{r-cap} \cdot \alpha_{r,e}} \right\rceil$. ◀

Proof of Theorem 8. Obtain a new instance of (LC+RP) by defining new lower and upper bounds $w_e^{\hat{\min}}$ and $w_e^{\hat{\max}}$:

$$\begin{aligned} w_e^{\hat{\min}} &:= w_e^{\min} - \left(\sum_{l \in \mathcal{L}_{\text{fixed}}^{\text{Pool}} : e \in l} \text{l-cap} \cdot f_l + \sum_{r \in \mathcal{R}_{\text{fixed}}^{\text{Pool}} : e \in r} \text{r-cap} \cdot \alpha_{r,e} \cdot v_r \right) \\ w_e^{\hat{\max}} &:= w_e^{\max} - \left(\sum_{l \in \mathcal{L}_{\text{fixed}}^{\text{Pool}} : e \in l} \text{l-cap} \cdot f_l + \sum_{r \in \mathcal{R}_{\text{fixed}}^{\text{Pool}} : e \in r} \text{r-cap} \cdot \alpha_{r,e} \cdot v_r \right) \end{aligned} \quad \blacktriangleleft$$