

The line planning routing game

Philine Gattermann¹, Alexander Schiewe¹, and Marie Schmidt²

¹: Institute for Numerical and Applied Mathematics, University of Göttingen

²: Rotterdam School of Management, Erasmus University
Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands,
phone: +31 10 4082199, email: schmidt2@rsm.nl

3rd December 2014

Abstract

In this paper, we propose a novel algorithmic approach to solve line planning problems. To this end, we model the line planning problem as a game where the passengers are players which aim at minimizing individual objective functions composed of travel time, transfer penalties, and a share of the overall cost of the solution. To find equilibria of this routing game, we use a best-response algorithm. We investigate, under which conditions on the line planning model a passenger's best-response can be calculated efficiently and which properties are needed to guarantee convergence of the best-response algorithm. Furthermore, we determine the price of anarchy which bounds the objective value of an equilibrium with respect to a system-optimal solution of the line planning problem. For problems where best-responses cannot be found efficiently, we propose heuristic methods. We demonstrate our findings on some small computational examples.

Keywords: Transportation, Game Theory, Routing, Line Planning, Routing Game

1 Introduction

Due to the high complexity of public transportation planning, the planning process is normally subdivided in subsequent steps, such as network design, line planning, timetabling, vehicle scheduling, etc. The *line planning problem* aims at determining the routes which are served regularly by a vehicle and the frequencies of these services. When evaluating a line concept, both the *costs* for implementing the line concept and the *quality* of the line concept from the passengers' perspective have to be taken into account. Possible ways to measure the quality of a line concept \mathcal{L} from the point of view of a passenger p are the *travel time* and the *number of transfers* on the route that passenger p would choose given the line concept \mathcal{L} . Hence, to evaluate a line concept \mathcal{L} from a quality perspective, passengers' routes for the chosen line concept have to be determined.

For this reason, recent publications on passenger-oriented line planning solve the problem of finding a line concept \mathcal{L} and a set of passengers' routes \mathcal{R} simultaneously, by modeling and solving it as linear integer programs, see e.g., [SS06a, BGP07, BN10, BK12, Sch14].

In order to achieve problem formulations which can be solved for practical instances, these models use several simplifications. Often, transfer times are assumed to be independent of line frequencies (see, e.g., [SS06a, BN10, BK12, Sch14]) or not taken into account at all [BGP07]. [GvHK06, GvHK04] use a model that allows to adjust transfer times to frequencies, but make a different restriction: for each passenger, the path in the network on which he travels is fixed beforehand (even if the exact connection, i.e., the sequence of trains used on this path, is not).

Furthermore, the cited approaches determine a system-optimum with respect of the cumulated objective functions of all passengers. In order to achieve a system-optimal solution, single passengers may be assigned routes which are significantly worse than their individually optimal route.

One possible approach to overcome these shortcomings is to iterate between planning and routing steps, as proposed in [Sch14]. However, this heuristic allows only to *remove* lines from the initial solutions, but not to include new lines, so it can get stuck fast in a local optimum.

In this paper, we propose a novel algorithmic approach to solve line planning problems. Instead of integrating planning and routing steps or iterating between both as done in the approaches described above, we regard only the routing step and include all planning decisions in this step. To this end, we define an individual objective function for each passenger which is composed of travel time, transfer penalties, and a share of the overall cost of the solution. This way, the line planning problem can be interpreted as a game in which the passengers are the players who aim at minimizing their objective functions.

In this paper, we study this so-defined *line planning routing game*. To find equilibria we propose a best-response algorithm. We investigate, under which conditions on the line planning model a passenger's best-response can be calculated efficiently and which properties are needed to guarantee convergence of the best-response algorithm. Furthermore, we determine the price of anarchy which bounds the objective value of an equilibrium with respect to a system-optimal solution of the line planning problem. We demonstrate our results on some numerical examples.

The remainder of this paper is structured as follows. In Section 2 we review literature on line planning before we introduce a general line planning problem (LPQC) in Sections 3.1 and 3.2. We briefly introduce some concepts from game theory in Section 3.3. In Section 3.4 we define the line planning routing game LPRG defined by LPQC and discuss the relations between the two problems in Section 3.5.

In Section 4 we investigate properties of the line planning routing game. In Section 4.1 we sketch the best-response algorithm used to find equilibria to LPRG. In Section 4.2 we investigate, under which conditions on the line planning model a passenger's best-response can be calculated efficiently. The existence of equilibria and the convergence of the best-response algorithm is investigated in Section 4.3. Section 4.4 evaluates the solutions found by the best-response algorithm with respect to a system-optimal solution. Finally, in Section 5 we illustrate and compare the different models on some small line planning instances.

2 Related literature

Line planning is an important step in the public transportation planning process. There are many line planning models, which differ with respect to the decisions covered by the term *line planning*, the level of detail with which real-world constraints are included in the model, and the way of measuring the quality of a line plan. In this paper, we only give a brief overview on the line planning models and solution methods which are most relevant for this paper. See, e.g., [Sch11, Sch14] for more extensive overviews on line planning.

Line planning aims at finding a line concept (consisting of line routes and frequencies) which is good from an operational point of view and offers good quality for the passengers. Cost-oriented line planning models focus on minimizing the operational costs subject to the constraint that passenger demand has to be satisfied (see, e.g., [CvDZ98, Bus98, GvHK06, BHK⁺13]).

Many passenger-oriented line planning models aim at minimizing the overall travel time while keeping the costs below a predefined threshold [SS06a, Sch14]. There are also some passenger-oriented models which measure quality by the number of direct travelers [Die78, Bus98, BKZ97]. Some models combine quality and cost into one objective [BGP08, GYW06, PB06].

In most literature on line planning, line planning problems are modeled and solved as integer programs. Solution approaches for cost-oriented models often assign the demand to the network edges in a preprocessing step and formulate covering or packing models. Solution techniques include branch-and-bound [Bus98, CvDZ98], branch-and-cut [GvHK04], and variable fixing heuristics [BLL04].

Passenger-oriented line planning assumes that passengers choose the "best" route with respect to the chosen line concept (where "best" is often understood as travel-time minimal). For this purpose, passengers' routes cannot be determined in a preprocessing step but have to be determined together with the line concept. [SS06a] model passengers as flows in a *change-and-go network*, which allows to include transfer times in the travel time, and solve the LP-relaxation using Dantzig-Wolfe decomposition. However, this leads to very large IP models and relatively long solution times. [BGP07, BN10, BK12] use column generation to generate passengers' routes. In [BN10] it is shown that this can lead to a significant speed-up with respect to flow formulations in change-and-go networks. However, transfer time are not taken into account [BGP07] or only approximated [BN10, BK12] in these models.

A drawback of the described passenger-oriented models is that they can only find system-optimal solution, that is solutions with a minimal *overall* travel time. I.e., a solution to these models, consisting of line concept and passenger routes could actually require that some passengers take long detours to avoid capacity violations and allow a minimal overall travel time. [Sch14] introduces a model where only line concepts which allow *all* passengers to travel on shortest paths (with respect to the line concept) are considered feasible and proposes an IP formulation. However, since the formulation is based on a flow formulation in a change-and-go network, fast solution times for real-world instances are not to be expected.

Solution approaches to line planning which are not IP-based, often concentrate on the line routes only and postpone frequency setting to a later step. They use greedy strategies [CW91, PRR95, Qua03] to construct lines or successively remove lines from a big line pool [Pat25, Son79]. Furthermore, metaheuristics like genetic algorithms [FGP02, FM06a, SW11], neighborhood search [SW11], and simulated annealing [FM06b] are used. [Man80, JBT10, Sch14] describe iterative approaches, where line planning/frequency setting and route assignment steps are iterated.

There are also game-theoretic approaches to line planning which model line operators as players who compete for a good utilization of the lines they offer [SS06b, Sch09, BKZ09, BKZ11, SS13, Neu14]. In [LMP10], the problem of finding a line concept which is robust against link failures is modeled as a game between the network provider and an adversary. However, to the extent of our knowledge, so far no attempt has been made to model line planning as a game with *passengers* as players.

In the field of transit assignment, models from game theory are used to model passenger flows on networks (see, e.g. [SF89, She85, DCF93, NP88, SFS⁺11, SSJ11, CF95]). These models take into account different modeling requirements from practice, like e.g., limited seat capacity or uncertain information about the next arriving trains. Equilibria are often found by mathematical programming.

Routing games on networks are also studied from a more theoretical perspective in the area of algorithmic game theory. A good overview of this line of research, both for atomic and non-atomic flow, is given, e.g., in [Rou07]. Questions of interest cover the existence and quality of equilibria and algorithmic approaches to identify equilibria (see, e.g. [Ros73, AAE05, Rou05, ADK⁺04, Rou07, TW07])

3 The line planning routing game

3.1 Line planning

Line planning aims at determining routes and frequencies of trains. As a basis, we consider the underlying *public transportation network (PTN)* $G = (V, E)$. The nodes V of this network represent stations. Two stations are connected by an edge e if there is a direct track connection between the corresponding stations. In this paper, we assume that we are given a line pool \mathcal{L} of possible lines. The main task of line planning is to find a *line concept*, i.e., to assign a frequency $f_l \in \mathbb{N}_0$ to every line l in the line pool \mathcal{L} .

Passenger demand is given in form of *origin-destination (OD)*-pairs (u_q, v_q) , specifying a passengers' origin u_q and his destination v_q . Apart from the line concept, passenger-oriented line

planning models determine a set of *passenger routes* $\mathcal{R} := \{R_q : q \in \mathcal{Q}\}$. A route R_q for passenger q specifies a path $P'_q = (e_1, \dots, e_n)$ from u_q to v_q and for every $e_i \in P'_q$ a line l_i which is used while traveling on e_i . I.e., R_q can be written as a sequence $R_q = ((e_1, l_1), (e_2, l_2), \dots, (e_n, l_n))$.

For a given set of routes \mathcal{R} we denote the number of passengers who use line $l \in \mathcal{L}$ on edge $e \in l$ by

$$x_{(e,l)}(\mathcal{R}) := |\{q \in \mathcal{Q} : (e, l) \in R_q\}|.$$

A pair of frequencies f and passenger route set \mathcal{R} is *feasible* if for every l and every $e \in l$ the number of passengers traveling on the line does not exceed the capacity of a single train multiplied with the frequency of the line, that is, if

$$x_{(e,l)}(\mathcal{R}) \leq f_l B.$$

The objective of line planning is to minimize the *costs* of the implemented line concept as well as its *quality*, estimated, e.g., by the travel time and the number of transfers.

I.e., the line planning problem can be summarized as follows: given a PTN G , a line pool \mathcal{L} , a capacity bound B , and a set of passengers \mathcal{Q} : find a pair of frequencies f and routes \mathcal{R} such that

$$x_{(e,l)}(\mathcal{R}) \leq f_l B$$

and a weighted sum of quality and cost is minimized, as detailed in the following section.

3.2 The line planning objective function

There are many approaches on how to measure cost and quality, as described in the literature overview in Section 2. The models used here are described in more detail in the following:

In our model the costs cost_l of a line l are composed of a constant cost k_1^l and a per-frequency cost k_2^l , i.e.,

$$\text{cost}_l(f) = \begin{cases} \gamma_1 \cdot k_1^l + \gamma_2 \cdot k_2^l \cdot f_l & \text{if } f_l > 0 \\ 0 & \text{otherwise} \end{cases}$$

where γ_1 and γ_2 are non-negative constants. The cost of a line concept represented by frequencies f is hence given as

$$\text{cost}(f) := \sum_{l: f_l > 0} \text{cost}_l(f). \quad (1)$$

Quality, as perceived by a passenger, can be measured by driving time and transfer time of the passenger's journey. We denote $\text{travel}_q(f, R_q) := \alpha_1 c_q(R_q) + \alpha_2 \cdot \tau_q(f, R_q)$, where α_1 and α_2 denote non-negative weighting factors.

The driving time on route R_q depends only on the chosen route in the PTN. It is given as

$$c_q(R_q) := \sum_{(e,l) \in R_q} c(e).$$

The transfer time $\tau_q(f, R_q)$ is estimated based on the frequencies of the lines involved in the transfers on the route. In this paper, for a transfer from line l to line l' we assume a transfer time of $\frac{1}{f_l + f_{l'}}$. Then, the overall transfer time of passenger q on route R_q is

$$\tau_q(f, R_q) := \sum_{i=1}^{n-1} \frac{1}{f_{l_i} + f_{l_{i+1}}}, \quad (2)$$

where (l_1, l_2, \dots, l_n) is the sequence of lines used on R_q .

Additionally, in $\text{transfer}_q(R_q)$ a penalty β can be added for every transfer on R_q , modeling the inconvenience arising for the passenger when transferring.

We obtain the line planning objective

$$G(\mathcal{R}, f) := \sum_{q \in \mathcal{Q}} \underbrace{(\alpha_1 c_q(R_q) + \alpha_2 \tau_q(f, R_q))}_{\text{travel}_q(f, R_q)} + \sum_{q \in \mathcal{Q}} \text{transfer}_q(R_q) + \underbrace{\sum_{l \in \mathcal{L}: f_l > 0} (\gamma_1 \cdot k_1^l + \gamma_2 \cdot k_2^l f_l)}_{\text{cost}(f)}. \quad (3)$$

We call the problem described in Sections 3.1 and 3.2 *line planning with travel quality and cost objective (LPQC)*.

The following observation from [Sch14] will be useful in the remainder of this paper:

Observation 3.1. *Given a route set \mathcal{R} we can easily determine a corresponding line concept $f(\mathcal{R})$ by setting $f_l(\mathcal{R}) := \max_{e \in l} \left\lfloor \frac{x_{(e,l)}(\mathcal{R})}{B} \right\rfloor$.*

The pair $(\mathcal{R}, f(\mathcal{R}))$ is feasible by definition. In [Sch14] it was shown that it has minimal cost assuming the cost model $\text{cost}(\mathcal{R}, f) := \sum_{l \in \mathcal{L}} \gamma_2 \cdot k_2^l \cdot f_l$. However, it is easy to see that given \mathcal{R} this line concept is cost-optimal for any *reasonable* cost model, in particular for the ones we will consider later.

3.3 Basics from game theory

In this section we describe some basic concepts from game theory which are used in the remainder of this paper. See, e.g., [NRTV07] for details.

Game theory studies the dynamics of situations where players try to minimize individual, conflicting objective functions. In a game $(\mathcal{Q}, \mathcal{S}, g)$, each player $q \in \mathcal{Q}$ has a set of *strategies* \mathcal{S}^q among which he can choose. The individual objective function $g_q(S_q, S^{-q})$ of player q depends on his chosen strategy S_q , but also on the strategies $S^{-q} = (S_1, S_2, \dots, S_{q-1}, S_{q+1}, \dots, S_{|\mathcal{Q}|})$ chosen by the other players.

A central concept of game theory is the concept of equilibria. A set of strategies $(S_1, \dots, S_{|\mathcal{Q}|})$ is called (*Nash*) *equilibrium* if none of the players can improve his individual objective function by changing his strategy (given that all other players do *not* change their strategies). I.e., $\hat{S} = (\hat{S}_1, \dots, \hat{S}_{|\mathcal{Q}|})$ is an equilibrium if for all $q \in \mathcal{Q}$ it holds that

$$g_q(\hat{S}_q, \hat{S}^{-q}) \leq g_q(S_q, \hat{S}^{-q}) \quad \forall S_q \in \mathcal{S}^q.$$

Not all games have equilibria, and even if equilibria exist, they can be hard to find and they do not need to be unique.

A special class of games with good properties is the class of potential games. We call a function $\Phi : \mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \times \dots \times \mathcal{S}_{|\mathcal{Q}|} \rightarrow \mathbb{R}$ *potential function*, if it satisfies the relation

$$\Phi(S) - \Phi(S') = g_q(S_q, S^{-q}) - g_q(S'_q, S^{-q}) \quad (4)$$

for all solutions $S = (S_1, \dots, S_{|\mathcal{Q}|}) \in \mathcal{S}$, all players $q \in \mathcal{Q}$ and all solutions $S' = (S_1, \dots, S_{q-1}, S'_q, S_{q+1}, S_{|\mathcal{Q}|})$ which can be obtained from S by exchanging the strategy of player q . A game with potential function is called *potential game*. The existence of a potential function allows us to interpret the problem of finding an equilibrium to $(\mathcal{Q}, \mathcal{S}, g)$ as an optimization problem. As we can easily verify in (4), an optimal solution to Φ is an equilibrium for the considered game (although there may be equilibria which are not optimal for Φ).

Furthermore, the relation (4) implies that every time a player changes his strategy to improve his personal objective (while the other players' strategies remain unchanged), the solution becomes better with respect to Φ and, in this sense, closer to an equilibrium. This motivates the approach of using *best-response algorithms* to find equilibria: in every step, one of the players changes his strategy to the *best response* with respect to the other players' strategies, i.e., he picks a solution of the optimization problem $\min_{S_q \in \mathcal{S}^q} g_q(S_q, S^{-q})$ as a new strategy. If there is only a finite number of strategies, this procedure converges to an optimum of Φ , and hence to an equilibrium of the game in a finite number of steps.

A centralized way to evaluate a solution $S = (S_1, \dots, S_{|\mathcal{Q}|})$ is to sum up the individual objective functions to a centralized objective function $G(S) = \sum_{q \in \mathcal{Q}} g_q(S_q, S^{-q})$. We call $S \in \mathcal{S}$ *system-optimal* if it minimizes G .

There exist different concepts to measure the inefficiency of equilibria with respect to the centralized objective. The *price of anarchy* is defined as

$$\max_{S^* \text{ is an equilibrium}} \frac{G(S^*)}{\min_S G(S)}.$$

Assuming that over time, selfish behavior will converge to equilibrium solutions, the price of anarchy gives a worst-case bound on the quality of such a convergence process.

The *price of stability*,

$$\min_{S^* \text{ is an equilibrium}} \frac{G(S^*)}{\min_S G(S)},$$

in contrast, quantifies how far the best equilibrium (i.e.: the best solution that would be accepted by the players) is away from system optimality.

3.4 The line planning routing game

Our approach to finding good line planning solutions is to determine passengers' routes \mathcal{R} in \mathcal{G} in such a way that $(\mathcal{R}, f(\mathcal{R}))$ is a good solution to the line planning problem.

To this end, we model the line planning problem as a routing game. The passengers \mathcal{Q} are the players. The strategies of a passenger q are the routes \mathcal{R}_q from $(0, u_q)$ to $(v_q, 0)$. Each passenger has an individual objective function $g_q(R_q, \mathcal{R}^{-q})$ which depends on his chosen route R_q and the routes chosen by the other passengers \mathcal{R}^{-q} . We call this game *line planning routing game (LPRG)*. Equilibria \mathcal{R}^* of this game will be interpreted as solutions $(\mathcal{R}^*, f(\mathcal{R}^*))$ of the line planning problem. The choice of the individual objective functions g_q is of course crucial for the quality of the obtained solutions. Ideally, the individual objective functions should

- account for individual travel quality as well as costs in order to find a solution which is balanced between the two partly contradicting objectives of minimizing costs while maximizing quality
- model passengers' behavior as realistically as possible.

We propose the following general model. The passengers objective is composed of the quality of the solution and a share of the overall costs as defined in (1):

$$g_q(R_q, \mathcal{R}^{-q}) := \underbrace{\text{travel}_q(R_q, \mathcal{R}^{-q}) + \text{transfer}_q(R_q)}_{=-\text{quality}_q(f(\mathcal{R}), R_q)} + \underbrace{\text{cost}_q(R_q, \mathcal{R}^{-q})}_{\text{cost share}}. \quad (5)$$

To share the costs among the passengers, we equally divide the cost of all lines among all users

$$\text{cost}_q(R_q, \mathcal{R}^{-q}) := \sum_{l \in R_q} \frac{\text{cost}_l(f(\mathcal{R}))}{|\{q' \in \mathcal{Q} : l \in R_{q'}\}|}$$

(this is referred to as *line-based* cost model in the following) or assign the line costs to the edges $e \in l$ as *edge costs* $\text{cost}_{(e,l)}$ and set

$$\text{cost}_q(R_q, \mathcal{R}^{-q}) := \sum_{(e,l) \in R_q} \frac{\text{cost}_{(e,l)}(f(\mathcal{R}))}{x_{(e,l)}(R_q, \mathcal{R}^{-q})}$$

(called *edge-based* model in the following). If $\gamma_2 = 0$, $\text{cost}_{(e,l)}(f(\mathcal{R}))$ is constant and we write $\text{cost}_{(e,l)} := \text{cost}_{(e,l)}(f(\mathcal{R}))$.

Note that in the definition of the quality functions in Section 3.1 and the individual objective functions in the section, we implicitly assumed that the passenger all have the same perception of quality of a travel route since we assume the weighting factors α_i , β , and γ_i $i = 1, 2$ to be the same for each passenger. It would be possible to replace these common weighting factors by a set of individual weighting factors for each passenger. However, for the sake of simplicity, in this paper we only consider the case of common weighting factors for all passengers.

3.5 Relation between LPQC and LPRG

In the line-based cost model and it holds that

$$\sum_{q \in \mathcal{Q}} \text{cost}_q(R_q, \mathcal{R}^{-q}) = \text{cost}(f(\mathcal{R})).$$

This is also true in the edge-based cost model, as long as it is ensured that a line does not cross an edge completely empty.

Furthermore, by definition

$$\sum_{q \in \mathcal{Q}} \text{travel}_q(R_q, \mathcal{R}^{-q}) = \text{travel}(f(\mathcal{R}), \mathcal{R}) \quad \text{and} \quad \sum_{q \in \mathcal{Q}} \text{transfer}_q(R_q) = \text{transfer}(\mathcal{R}).$$

We obtain

$$\sum_{q \in \mathcal{Q}} g_q(R_q, \mathcal{R}^{-q}) = G(\mathcal{R}, f(\mathcal{R})),$$

which implies that a system-optimal routing corresponds to an optimal solution of the line planning problem described in Sections 3.1 and 3.2.

Hence, on the one hand, if the price of anarchy in the LPRG is small, an equilibrium \mathcal{R}^* of the game provides us with a good approximation $(f(\mathcal{R}^*), \mathcal{R}^*)$ for LPQC.

Lemma 3.2. *Denote by I an instance of the LPQC. Assume that the price of anarchy for the corresponding instance I_{RG} of LPRG is bounded by ξ . Then any equilibrium \mathcal{R}^* of I_{RG} is a ξ -approximation $(f(\mathcal{R}^*), \mathcal{R}^*)$ for I .*

So, on the one hand, finding an equilibrium to LPRG may be regarded as a new, decentralized, way of solving LPQC.

On the other hand, one may argue that in some cases, optimal solutions to LPQC are not desirable in practice. Indeed, it may happen that the routing \mathcal{R} in a solution (f, \mathcal{R}) to LPQC allots very long routes to some passengers for the "greater good" of a solution which is optimal with respect to objective (3).

As an example, consider the situation shown in Figure 1. There are two lines (depicted by gray arrows) from station v_1 to station v_7 . One is a fast line which stops only at one intermediate station, the other one is a regional line which serves a geographically different route and visits many small stations in between. Assume that the transportation capacity of each line is $B = 100$. There are 100 passengers who want to travel from v_1 to v_7 , 50 who want to travel from v_2 to v_7 , and some smaller amounts of passengers traveling to and from the regional stations. Hence, both lines have to be established. Now, if the cost factors γ_1 and γ_2 in the objective function (3) are comparatively big, both lines will be established with frequency 1 in an optimal solution (f', \mathcal{R}') to LPQC. This means that 50 of the 100 passengers from v_1 to v_2 will be sent via the regional train route in an optimal solution.

However, if this solution was implemented in real life, at station v_1 , when the passengers from v_1 and v_7 have to make a decision which train to board, the fast train is still empty. To implement the solution (f', \mathcal{R}') into practice, somebody would have to convince these 50 passengers to use a slower connection to reserve the seats in the fast train for the passengers from v_2 to v_7 boarding later. It is not hard to imagine, that the passengers from v_1 to v_7 would board the train anyway so that the starting in v_2 could not board or the train would be overcrowded.

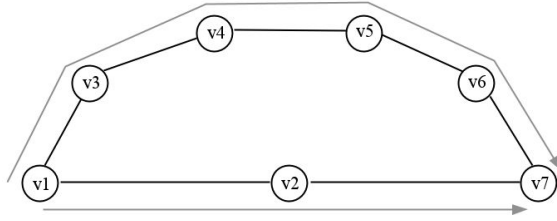


Figure 1: Example instance where LPQC finds undesirable solution.

This would not happen in the solution $(f(\mathcal{R}^*), \mathcal{R}^*)$ provided by an equilibrium \mathcal{R}^* of the corresponding routing game LPRG. In this solution, all passengers from v_1 to v_7 would choose the fast train and the planner would be *forced* to provide enough frequency here to avoid overcrowding - unless taking the slow line along would be cheap enough to be a favorable option for the passengers. Hence, if we assume that $\text{cost}_q(R_q)$ is an estimate of the real costs that a passenger would pay on a route R_q , the solution $(f(\mathcal{R}^*), \mathcal{R}^*)$ defined by an equilibrium \mathcal{R}^* of LPRG models passenger behavior in a better way, provides better estimates of actual solution quality and helps to avoid overcrowding and is hence, from this perspective, preferable to the solution (f', \mathcal{R}') found by LPQC.

4 Finding good line planning solutions using best-response algorithms

The approach of this paper is to find line planning solutions by finding equilibria of LPRG. We justified this approach in the preceding section. In this section, we concentrate on the process of finding such equilibria.

4.1 The best-response algorithm

A popular method for finding equilibria are *best-response algorithms* (compare Section 3.3). The scheme of a best-response algorithm for LPRG is given in Algorithm 1.

Algorithm 1 Best response algorithm

Require: PTN, line pool, set of passengers \mathcal{Q} , individual cost functions g_q , maximal number of iterations in $\mathbb{N} \cup \infty$

Ensure: A routing $\hat{\mathcal{R}}$

Start with an empty routing (or with an arbitrary non-empty routing).

while there are improvements for the passengers and maximal number of iterations is not reached **do**

for Passenger $q \in \mathcal{Q}$ **do**

 Calculate optimal passenger route R_q based on the route choices of the other passenger \mathcal{R}^{-q} (*routing problem*)

end for

end while

In the remainder of this paper we are going to discuss the following questions:

1. How can we solve the *routing problem*?
2. Does the best-response algorithm converge?
3. How good is the solution $(f(\hat{\mathcal{R}}), \hat{\mathcal{R}})$ to LPQC defined by the solution $\hat{\mathcal{R}}$ from Algorithm 1?

4.2 The routing problem

In every step of the (LPRG) we have to solve the following *routing problem* (RP_q) for passenger q :

Definition 4.1. *Given the PTN G , the line pool \mathcal{L} , origin u_q , destination v_q and individual objective function g_q for passenger q , and routes $R_{q'}$ for all passengers $q' \in \mathcal{Q} \setminus \{q\}$, the routing problem (RP_q) consists of finding a route R_q from u_q to v_q such that $g_q(R_q, \mathcal{R}^{-q})$ is minimized.*

Unfortunately, in general the routing problem which has to be solved in each iteration of the best-response algorithm is NP-hard. We see in Section 4.2.1 that there are two components which make the problem hard:

1. line-based costs (see Theorem 4.2),
2. and frequency-based transfer times (see Theorem 4.3).

On the contrary, if costs are assumed to be edge-based with $\gamma_2 = 0$ and transfer times are neglected, the problem becomes much better tractable, as we are going to discuss in Section 4.2.2. Section 4.5 discusses how to incorporate frequency-based transfer times in a heuristic way.

4.2.1 NP-hardness of the routing problem

For determining the complexity of our problems we will use reductions from the set covering decision problem (SCP), where an instance is given by a set of elements $\mathcal{M} = \{m_1, \dots, m_n\}$, a set of subsets \mathcal{C} with $C \subseteq \mathcal{M}$ and an integer $K \in \mathbb{N}$. The problem is to find a subset $\mathcal{C}' \subseteq \mathcal{C}$ such that $\bigcup_{C \in \mathcal{C}'} C \supseteq \mathcal{M}$ and $|\mathcal{C}'| \leq K$.

We first show that the assumption of line-based costs leads to an NP-hard routing problem.

Theorem 4.2. *The routing problem (4.1) with line-based cost model is NP-hard, even if transfer times and transfer penalties are not taken into account, i.e. if $\beta = 0$ and $\gamma_2 = 0$.*

Proof. We show that (SCP) given by $(\mathcal{M}, \mathcal{C}, K)$ can be reduced to the routing decision problem with line-based costs. Given an instance $(\mathcal{M}, \mathcal{C}, K)$ of (SCP) we construct an instance of the routing problem as follows.

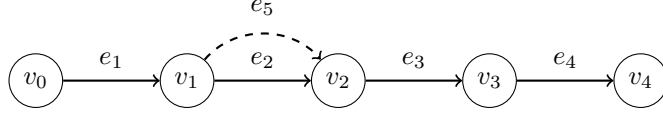
We create a node v_0 and for each $m_i \in \mathcal{M}$, $i = 1, \dots, n$ a node v_i and an edge $e_i = (v_{i-1}, v_i)$. For all $C \in \mathcal{C}$ we create a line $l_C \in \mathcal{L}$ containing all edges $\{e_i : m_i \in C\}$ and additional edges to ensure that the lines are connected paths in the PTN. We set edge lengths to $c(e) := 0$ for all edges related to $m \in \mathcal{M}$ and to $c(e) := K + 1$ for all additional edges. Furthermore we set the costs to $\text{cost}_l = 1$ for all lines. We consider a passenger q who wants to travel from v_0 to v_n . The parameters of the objective function are $\alpha_1 = \gamma_1 = 1$ and $\alpha_2 = \beta = \gamma_2 = 0$. (Note that the same construction is possible if $\gamma_1 = 0$ and $\gamma_2 = 0$.) Now there is a solution to the routing problem with objective value lower or equal to K if and only if there is a solution to (SCP) with objective value lower or equal to K :

Let \mathcal{C}' be a solution to (SCP). Then the set of lines $\mathcal{L}' := \{l_C : C \in \mathcal{C}'\}$ has costs $\leq K$ and allows q to travel from origin to destination with 0 travel time. On the other hand, in every solution to the constructed instance of the routing problem with travel time $\leq K$, q uses the edge sequence (e_1, \dots, e_n) , because otherwise his travel time would be $> K$. Hence, $\mathcal{C}' = \{C \in \mathcal{C} : q \text{ uses } l_C\}$ is a solution to (SCP). \square

The following example illustrates this construction. Consider the instance of (SCP) given by

$$\begin{aligned} \mathcal{M} &= \{1, 2, 3, 4\}, \\ \mathcal{C} &= \{C_1 = \{1, 2\}, C_2 = \{1, 3\}, C_3 = \{3, 4\}\}, \\ K &= 2. \end{aligned}$$

This leads to the PTN



where e_1, \dots, e_4 correspond to \mathcal{M} and e_5 is a helping edge for C_2 with

$$c(e_i) = \begin{cases} 0 & , i = 1, \dots, 4 \\ 3 & , \text{otherwise} \end{cases}$$

and

$$\mathcal{L}_0 = \{l_1 = (e_1, e_2), l_2 = (e_1, e_5, e_3), l_3 = (e_3, e_4)\}.$$

Due to the result of Theorem 4.2, in the remainder of this paper we restrict to edge-based cost functions. However, even without considering costs, a similar hardness result can be proven for the routing problem with frequency-based transfer times:

Theorem 4.3. *The routing problem (4.1) is NP-hard, even if transfer penalties and operational costs are not taken into account, i.e., $\beta = 0$ and $\gamma_1 = \gamma_2 = 0$.*

Proof. Similarly to the proof of Theorem 4.2 we prove this theorem by reduction from (SCP). Let $(\mathcal{M}, \mathcal{C}, K)$ denote an instance of (SCP) with $n = |\mathcal{M}|$. We create a linear PTN consisting of two parts. The first part is used to ensure that at most K sets are chosen from \mathcal{C} . The second part is similar to the construction in the proof of Theorem 4.2 and is used to determine whether the chosen sets cover \mathcal{M} .

The first part of the PTN consists of vertices v_i for $i = 1, \dots, 2K + 1$ and edges $e_i = (v_i, v_{i+1})$, $i = 1, \dots, 2K$ with $c(e_i) = 0$. Additionally there are edges $\bar{e}_{2i-1} = (v_{2i-1}, v_{2i})$ with costs $K' + 1$, where $K' := \frac{2K+2n}{3}$. For every edge e_{2i-1} with an odd index we introduce a line \bar{l}_{2i-1} which consists of this edge only.

The second part of the PTN consists of vertices w_i for $i = 1, \dots, 2n + 1$ and edges $a_i = (w_i, w_{i+1})$ for $i = 1, \dots, 2n$ with $c(a_i) = 0$. Furthermore, we add edges \bar{a}_{ij} which connect all pairs of vertices w_i and w_j with $i < j$ and whose length is $c(\bar{a}_{ij}) := K' + 1$. For each $i = 1, \dots, n$ we introduce a line \tilde{l}_{2i-1} which covers the edge a_{2i-1} . We connect both parts of the PTN by a transition edge $t = (v_{2K+1}, w_1)$.

For every $C \in \mathcal{C}$ we create a line $l_C \in \mathcal{L}$ containing all edges $\{e_{2i} : m_i \in C\}$ from the first part of the PTN, the transition edge t , and the edges a_{2i} with $m_i \in C$ from the second part of the PTN. We add additional edges with lengths $K' + 1$ whenever needed to ensure that the lines are connected paths in the PTN.

In contrast to the proof of Theorem 4.2, in this proof we have $|\mathcal{C}| + 1$ passengers. Each passenger q_C with $C \in \mathcal{C}$ has origin v_1 and destination v_{2K+1} and his route \mathcal{R}_{q_C} is identical to line l_C from v_1 to v_{2K+1} . The passenger q for which we have to solve the routing problem has origin v_1 and destination w_{2n+1} . We set the capacity in each train to $B := 1$. For the objective function we use the parameters $\alpha_1 = \alpha_2 = 1$ and $\beta = \gamma_1 = \gamma_2 = 0$. Note that line costs can be set to arbitrary values, since $\gamma_1 = \gamma_2 = 0$.

We now show that there is a solution to the considered instance of (SCP) if and only if there is a solution R_q to the routing problem (RP_q) with individual objective value $g_q(R_q, \mathcal{R}^{-q}) \leq K'$.

First note that any such route in the first part of the PTN will use the lines \bar{l}_i on edges with an odd index and some lines l_i on the ones with an even index. Furthermore, for all of these paths the contribution from the first part of the PTN to the individual objective function is $\frac{2K}{3}$, since the length of every used edge is 0 and on each such path there is a transfer at each station between a line \bar{l} that is used by one passenger and a line l that is used by two passengers. In the second part of the PTN, only edges a_i can be used in a route R_q with $g_q(R_q, \mathcal{R}^{-q}) \leq K'$. Now consider the transfer costs of such a path in the second part of the PTN. At each node in the second part of the PTN a transfer has to take place. Thereby, transfer time is $\frac{1}{2}$ if passenger q did not use line l_j in the first part of the PTN, $\frac{1}{3}$ if he used it. Since there are $2n$ such transfers, any path with

individual objective value $\leq K'$ uses on edge a_{2i} a line that was already used in the first part of the PTN.

Due to the construction of the lines l_C , this means that if there is a route R_q with $g_q(R_q, \mathcal{R}^{-q}) \leq K'$, for each element $m_i \in \mathcal{M}$ at least one line l_C with $C \ni m_i$ is used in the first part of the PTN. Since not more than K such lines can be used in R_q , there must be a solution to the considered instance of (SCP).

On the other hand, if there is a solution $\mathcal{C}' = \{C_1, \dots, C_k\}$ with $k \leq K$ to the considered instance of (SCP), using line l_{C_i} on edge e_{2i} for $i = 1, \dots, k$ (and arbitrary lines on e_{2i} for $i = k+1, \dots, K$) allows the passenger to choose a path with transfer time $\frac{n}{3}$ in the second part of the PTN and thus yields an individual objective value $\leq K'$. \square

4.2.2 Cases with polynomially solvable routing problem

A convenient way to represent line planning problems is the change-and-go network (CGN) $\mathcal{G} = (\mathcal{V}, \mathcal{A})$, which was first introduced in [SS06a]. The set of nodes consists of *station nodes* $\mathcal{V}_{stat} := \{(v, \text{board}) : v \in V\} \cup \{(v, \text{alight}) : v \in V\}$ and *travel nodes* $\mathcal{V}_{trav} := \{(v, l) : l \in \mathcal{L}, v \in l\}$. The set of arcs is $\mathcal{A} := \mathcal{A}_{OD} \cup \mathcal{A}_{trans} \cup \mathcal{A}_{line}$ with

- line arcs $\mathcal{A}_{line} := \{(e, l) : l \in \mathcal{L}, e \in l\}$ for each edge e covered by a line l ,
- transfer arcs $\mathcal{A}_{trans} := \{((v, l_1), (v, l_2)) : v \in V, l_1 \ni v, l_2 \ni v\}$,
- and arcs for boarding and alighting

$$\mathcal{A}_{OD} := \{((v, \text{board}), (v, l)) : l \in \mathcal{L}, v \in l\} \cup \{((v, l), (v, \text{alight})) : l \in \mathcal{L}, v \in l\}.$$

Now every route R_q for a passenger q can be uniquely represented in \mathcal{G} as a path P_q from (u_q, board) to (v_q, alight) in \mathcal{G} .

For $a \in \mathcal{A}$ we denote by $x_a(\mathcal{R})$ the number of passengers, using arc a of the CGN, i.e., $x_a(\mathcal{R}) := \{q \in \mathcal{Q} : P_q \ni a\}$ where P_q is the path in the CGN corresponding to R_q . To abbreviate, we sometimes omit the routing and use the notation $x_a := x_a(\mathcal{R})$.

Let us now assume that (given \mathcal{R}^{-q}) the objective value of a route R_q is the sum of edge weights over all edges contained in the corresponding path P_q , i.e., that there are arc weights $w_a^q(\mathcal{R}^{-q}) \geq 0 \forall a \in \mathcal{A}$ such that

$$g_q(R_q, \mathcal{R}^{-q}) = \sum_{a \in P_q} w_a^q(\mathcal{R}^{-q}). \quad (6)$$

This is the case if costs are edge-based with $\gamma_2 = 0$ and $\alpha_2 = 0$. Indeed, it is easy to check that the weights

$$w_a^q(\mathcal{R}^{-q}) := \begin{cases} \alpha_1 c(e) + \frac{\text{cost}_{(e,l)}}{x_{(e,l)}(\mathcal{R}^{-q}) + 1} & \text{if } a = (e, l) \in \mathcal{A}_{drive} \\ \beta & \text{if } a \in \mathcal{A}_{transfer} \end{cases}$$

fulfill (6). In Section 4.5, different approaches to define arc weights are studied.

If edge weights of the form (6) can be found, we obtain the following lemma:

Lemma 4.4. *Consider an instance of the routing problem defined by a network \mathcal{G} , a set of passengers \mathcal{Q} , origin u_q and destination v_q of a passenger q with individual objective function g_q and set of routes \mathcal{R}^{-q} for all $q' \in \mathcal{Q} \setminus \{q\}$. If there are arc weights $w_a^q(\mathcal{R}^{-q})$ as defined in (6), (RP_q) can be solved in polynomial time.*

Proof. In this case, any shortest path from (u_q, board) to (v_q, alight) with respect to the edge weights $w_a^q(\mathcal{R}^{-q})$ is an optimal solution to I . Hence, we can find a solution using, e.g., Dijkstra's algorithm. \square

Hence, in this case, we can use the best-response algorithm 1 with, e.g., Dijkstra's algorithm in the routing step to search for an equilibrium of the LPRG.

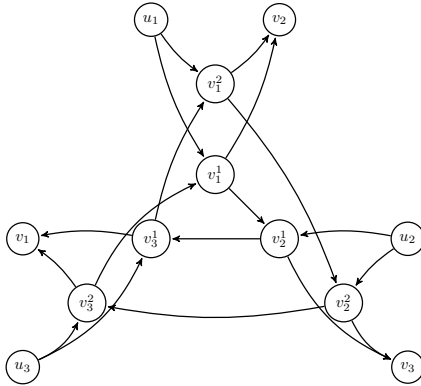


Figure 2: PTN for an example instance where there are no equilibria for LPRG and Algorithm 1 does not converge.

4.3 Existence of equilibria and convergence of the best-response algorithm

In this section we study under which assumptions equilibria to the LPRG exist and can be found by the best-response algorithm. We start with an example which shows that in the general case the existence of an equilibrium is not guaranteed.

4.3.1 Non-existence of equilibria

The following example shows that for some instances of LPRG there are no equilibria. This also implies that the best-response algorithm does not converge in this case.

We consider the PTN from Figure 2 with 12 nodes and 18 edges. Every edge is served by one directed line (which contains only the edge). The capacity of a train is $B = 1$. There are three *main* passengers q_1 from u_1 to v_1 , q_2 from u_2 to v_2 , and q_3 from u_3 to v_3 and six sets of *auxiliary* passengers: \mathcal{Q}_i^j for $i = 1, \dots, 3$ and $j = 1, 2$ contains M passengers from v_i^j to v_i (where M is a sufficiently large number, e.g., $M > 12$). We denote by \mathcal{Q}' the union of the auxiliary passengers. We consider the objective functions $g_q(\mathcal{R}) := \tau_q(R_q, \mathcal{R}^{-q})$ which, for the sake of simplicity, takes only the transfer time into account, i.e., $\alpha_1 = \beta = \gamma_1 = \gamma_2 = 0$.

Note that the auxiliary passengers only have one strategy, let \mathcal{R}' denote the set of these strategies. Each of the main passengers q_i has two different strategies: to take the route R_i^1 starting with edge (u_i, v_i^1) or to take the route R_i^2 starting with edge (u_i, v_i^2) .

We now show that there does not exist an equilibrium in the described situation. Assume that \mathcal{R} is an equilibrium of the described line planning routing game. Denote by $R_i^{j_i}$ the strategy chosen by q_i . Without loss of generality, assume that $j_1 = 1$. Then

$$g_2(R_1^1, R_2^1, R_3^{j_3}, \mathcal{R}') = \begin{cases} \frac{1}{1+2} + \frac{1}{2+2} + \frac{1}{2+M+1} = \frac{7}{12} + \frac{1}{M+3} & \text{if } j_3 = 1 \\ \frac{1}{1+2} + \frac{1}{2+1} + \frac{1}{1+M+1} = \frac{8}{12} + \frac{1}{M+2} & \text{if } j_3 = 2 \end{cases}$$

and

$$g_3(R_1^1, R_2^2, R_3^{j_3}, \mathcal{R}') = \begin{cases} \frac{1}{1+1} + \frac{1}{1+1} + \frac{1}{1+M+1} = \frac{12}{12} + \frac{1}{M+2} & \text{if } j_3 = 1 \\ \frac{1}{1+1} + \frac{1}{1+2} + \frac{1}{2+M+1} = \frac{10}{12} + \frac{1}{M+2} & \text{if } j_3 = 2 \end{cases}$$

Since \mathcal{R} is an equilibrium, we conclude that $j_2 = 1$, i.e., $R_2^{j_2} = R_2^1$.

Now

$$g_3(R_1^1, R_2^1, R_3^1, \mathcal{R}') = \frac{1}{1+2} + \frac{1}{2+1} + \frac{1}{1+M+1} = \frac{4}{6} + \frac{1}{M+2}$$

and

$$g_3(R_1^1, R_2^1, R_3^2, \mathcal{R}') = \frac{1}{1+1} + \frac{1}{1+2} + \frac{1}{2+M+1} = \frac{5}{6} + \frac{1}{M+3}.$$

Since \mathcal{R} is an equilibrium, we conclude that $j_3 = 1$, i.e., $R_3^{j_3} = R_3^1$.
Now we have a look at the strategies for q_1 :

$$g_1(R_1^1, R_2^1, R_3^1, \mathcal{R}') = \frac{1}{1+1} + \frac{1}{1+2} + \frac{1}{2+M+1} = \frac{5}{6} + \frac{1}{M+3}$$

and

$$g_1(R_1^2, R_2^1, R_3^1, \mathcal{R}') = \frac{1}{1+2} + \frac{1}{2+1} + \frac{1}{1+M+1} = \frac{4}{6} + \frac{1}{M+2}.$$

Thus, $g_1(R_1^1, R_2^1, R_3^1, \mathcal{R}') > g_1(R_1^2, R_2^1, R_3^1, \mathcal{R}')$. This is a contradiction to R_1^1 being part of an equilibrium.

Due to the symmetry of the construction of the instance, the assumption that R_1^2 is part of an equilibrium leads to a contradiction in the same way.

4.3.2 Line planning routing games with potential functions

In contrast to the example from Section 4.3.1 as we see in Lemma 4.5, existence of equilibria and convergence can be guaranteed if for every $a \in \mathcal{A}$ there is an *arc weight function* $\bar{w}_a : \mathbb{N} \rightarrow \mathbb{R}$ such that

$$g_q(R_q, \mathcal{R}^{-q}) = \sum_{a \in P_q} \bar{w}_a(x_a) \quad (7)$$

for every route R_q from u_q to v_q and its corresponding path P_q in the CGN.

In case of edge-based costs with $\gamma_2 = 0$ and $\alpha_2 = 0$, such arc weight functions are given by

$$\bar{w}_a(x) := \begin{cases} \alpha_1 c(e) + \frac{\text{cost}_{(e,l)}}{x} & \text{if } a = (e, l) \in \mathcal{A}_{drive} \\ \beta & \text{if } a \in \mathcal{A}_{transfer} \end{cases} \quad (8)$$

Lemma 4.5. *Let $I := (G, \mathcal{L}, \mathcal{Q}, \{g_q : q \in \mathcal{Q}\})$ be an instance of the LPRG such that arc weight functions as specified in (7) exist. Then*

1. $\Phi(\mathcal{R}) := \sum_{a \in \mathcal{A}} \sum_{i=1}^{x_a(\mathcal{R})} \bar{w}_a(i)$ is a potential function for I ,
2. there exists an equilibrium to I ,
3. the best-response algorithm converges to an equilibrium in a finite number of steps,
4. each of the steps can be executed in polynomial time.

Proof. This proof follows standard arguments for convergence of atomic routing games, compare, e.g., [Rou07].

1. Let \mathcal{R} and \mathcal{R}' be two routings. We denote with P_q and P'_q the corresponding paths for passenger q in the CGN and with $x_a := x_a(\mathcal{R})$ and $x'_a := x_a(\mathcal{R}')$ the corresponding flows on edge a of the CGN. We first observe that

$$\begin{aligned} \Phi(R_q, \mathcal{R}^{-q}) - \Phi(R'_q, \mathcal{R}^{-q}) &= \sum_{a \in P_q \setminus P'_q} \bar{w}_a(x_a) - \sum_{a \in P'_q \setminus P_q} \bar{w}_a(x'_a) \\ &= g_q(R_q, \mathcal{R}^{-q}) - g_q(R'_q, \mathcal{R}^{-q}), \end{aligned}$$

hence Φ indeed is a potential function by (4).

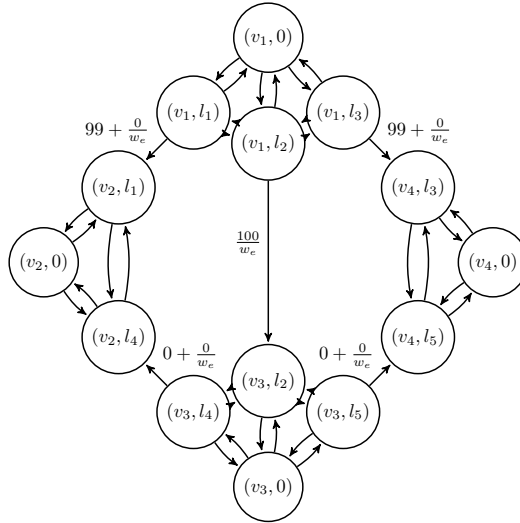


Figure 3: Two equilibria with different objective values

2. Hence, every optimum of Φ is an equilibrium of the game. Since the number of solutions is finite, there exists at least one optimum of Φ /equilibrium of I .
3. Since in each step of the best-response algorithm there is a non-zero improvement in the individual objective function and thus also in the potential function, and the number of solutions is bounded, the best-response algorithm converges to an optimum of Φ which is an equilibrium.
4. We set $w_a^q(\mathcal{R}^{-q}) := \bar{w}_a(x_a(\mathcal{R}^{-q}) + 1)$. Then

$$\begin{aligned}
 g_q(R_q, \mathcal{R}^{-q}) &= \sum_{a \in P_q} \bar{w}_a(x_a(\mathcal{R})) \\
 &= \sum_{a \in P_q} \bar{w}_a(x_a(\mathcal{R}^{-q}) + 1) \\
 &= \sum_{a \in P_q} w_a^q(\mathcal{R}^{-q}).
 \end{aligned}$$

The proposition follows from Lemma 4.4. □

Thus, if objective functions fulfill the conditions of Lemma 4.5 an equilibrium will be obtained after a finite number of steps, this holds in particular for all line planning routing games with $\alpha_2 = \gamma_2 = 0$ and edge-based costs.

4.4 Equilibria and system optima

In this section we investigate the quality of the equilibria of LPRG by comparing them to system-optimal solution of the line planning routing game. First, we illustrate that the equilibria which are found by the best-response algorithm are not necessarily optimal with respect to the line planning problem defined in Section 3.1, even if convergence is guaranteed by Lemma 4.5.

To this end we consider the PTN G and line pool \mathcal{L} associated to the CGN shown in Figure 3 with $c((v_1, v_2)) = c((v_1, v_3)) = 99$, $c(e) = 0$ and $\forall e \in E \setminus \{(v_1, v_2), (v_1, v_3)\}$. We set the edge-based line costs of edge (v_1, v_3) in line l_2 to $\text{cost}_{((v_1, l_2), (v_3, l_2)), l_2} := 100$, all other line costs are 0. We

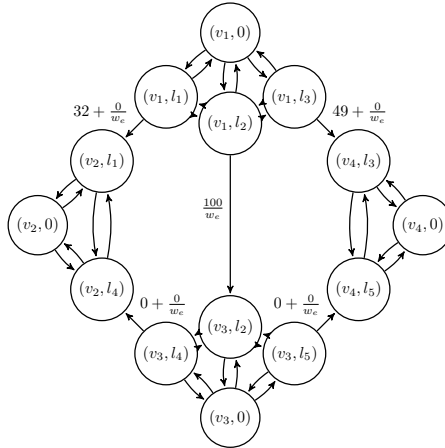


Figure 4: The system optimum is not necessarily an equilibrium.

consider two passengers: q_1 wants to travel from v_1 to v_2 and q_2 wants to travel from v_1 to v_4 , with individual objective functions $g_q(\mathcal{R}) = c(\mathcal{R}^q) + \sum_{(e,l) \in R_q} \frac{\text{cost}_{(e,l)}(f(\mathcal{R}))}{x_{(e,l)}}$, i.e., $\alpha_1 = \gamma_1 = 1$, $\beta = \alpha_2 = \gamma_2 = 0$. (The same construction is possible if we switch the role of γ_1 and γ_2 and assign high enough train capacity).

There are two equilibria:

1. \mathcal{R}' : q_1 uses line 1 and q_2 uses line 3. Both of them have costs of 99, the social costs are 198.
2. \mathcal{R}^* : q_1 uses line 2 and 4, q_2 uses line 2 and 5. Both of them have costs of 50, the social costs are 100.

Clearly, the second equilibrium is preferable to the first one, since for both passengers the individual objective function are almost double in the first one. However, e.g., when starting with an empty solution, the best-response algorithm will find the first equilibrium.

In the following example we see that the optimal solution to the line planning problem is not necessarily an equilibrium of the corresponding routing game at all: Consider the PTN G and line pool \mathcal{L} associated to the CGN shown in Figure 4. This PTN is the same as the one used in the previous example, except that edge lengths $c((v_1, v_2)) = 32$ and $c((v_1, v_3)) = 49$ are changed. Like in the previous example we have $c(e) = 0 \forall e \in E \setminus \{(v_1, v_2), (v_1, v_3)\}$ and the edge-based line cost of l_2 on edge (v_1, v_3) is 100, all other line costs are 0. For the objective function we choose the parameters $\alpha_1 = \gamma_1 = 1$, $\beta = \alpha_2 = \gamma_2 = 0$ (again, the same construction is possible for $\gamma_1 = 0$ and $\gamma_2 = 1$ if capacity is high enough). Figure 4 shows the resulting CGN where all nonzero arc weight functions are shown. Let there be three passengers. q_1 wants to travel from v_1 to v_2 , q_2 wants to travel from v_1 to v_4 , q_3 wants to travel from v_1 to v_3 .

There is only one equilibrium: \mathcal{R}' : q_1 uses line 1, q_2 uses line 3, q_3 uses line 2. $G(\mathcal{R}', f(\mathcal{R}')) = 181$. The solution which is optimal for the line planning problem is \mathcal{R}^* : q_1 uses line 2 and 4, q_2 uses line 2 and 5, q_3 uses line 2, with overall objective value $G(\mathcal{R}^*, f(\mathcal{R}^*)) = 100$.

By extending the example given in Figure 4 in a straight-forward way, we see that for instances with an unbounded number of passengers, the price of stability is not bounded for the considered games: for n passengers we can construct an instance with price of stability $H_n = \sum_{i=1}^n \frac{1}{i}$.

However, we can bound the price of anarchy by the number of passengers if the arc weight functions (7) fulfill an additional property.

Lemma 4.6. *If there exist non-increasing arc weight functions \bar{w}_a with $\bar{w}_a(1) \leq x \cdot \bar{w}_a(x)$ for all $x \in \mathbb{N}$, the price of anarchy in the LPRG is at most the number of passengers.*

Proof. Let $\mathcal{X} := \{X_1, \dots, X_n\}$ be a socially optimal solution and let $\mathcal{R} := \{R_1, \dots, R_n\}$ be an equilibrium. Assume that $G(\mathcal{R}, f(\mathcal{R})) > |\mathcal{Q}|G(\mathcal{X}, f(\mathcal{X}))$. Then there is at least one passenger q with $g_q(\mathcal{R}) > |\mathcal{Q}|g_q(\mathcal{X})$. For this passenger q it follows that

$$g_q(X_q, \mathcal{R}^{-q}) = \sum_{a \in X_q} \bar{w}_a(x'_a) \leq \sum_{a \in X_q} \bar{w}_a(1) \leq \sum_{a \in X_q} \hat{x}_a \bar{w}_a(\hat{x}_a) \leq \sum_{a \in X_q} |\mathcal{Q}| \bar{w}_a(\hat{x}_a) < g_q(R_q, \mathcal{R}^{-q}),$$

where $\hat{x}_a := x_a(X_q, \mathcal{R}^{-q})$ denotes the number of passengers on arc a in (X_q, \mathcal{R}^{-q}) and $x'_a := x_a(\mathcal{R})$ the number of passengers on arc a in \mathcal{R} . This is a contradiction to the assumption that \mathcal{R} is an equilibrium. \square

Corollary 4.7. *If edge-based cost functions with $\gamma_2 = 0$ are considered and $\alpha_2 = 0$, the price of anarchy is bounded by the number of passengers.*

Proof. The functions given in (8) are non-increasing. Furthermore we have for $x > 0$

$$x\bar{w}_a(x) = \begin{cases} x\alpha_1 c(a) + \text{cost}(a) & \text{if } a \in \mathcal{A}_{drive} \\ x\beta & \text{if } a \in \mathcal{A}_{transfer} \end{cases} \geq \begin{cases} \alpha_1 c(a) + \text{cost}(a) & \text{if } a \in \mathcal{A}_{drive} \\ \beta & \text{if } a \in \mathcal{A}_{transfer} \end{cases} = \bar{w}_a(1).$$

\square

Hence, if we use the best-response algorithm as a heuristic for the line planning problem, we obtain a solution with approximation ratio $\frac{1}{|\mathcal{Q}|}$ where \mathcal{Q} is the number of passengers.

To see that there are indeed instances I with a price of anarchy that equals $|\mathcal{Q}|$ consider the example given in Figure 3. If we set the travel costs on (v_1, v_2) and (v_1, v_4) to 100, \mathcal{R}' and \mathcal{R} are still both equilibria and the price of anarchy is 2. We can easily extend this construction to an arbitrary number of passengers.

However, convergence to an equilibrium may be slow. To find a polynomial-time approximation algorithm for the initial line planning problem, the next lemma is more helpful.

Lemma 4.8. *If there exist non-increasing arc weight functions \bar{w}_a with $\bar{w}_a(1) \leq x \cdot \bar{w}_a(x)$ for all $x \in \mathbb{N}$, given an empty state of the game, calculating the best response once for every passenger leads to a routing \mathcal{R} with*

$$\frac{G(\mathcal{R}, f(\mathcal{R}))}{G(\mathcal{X}, f(\mathcal{X}))} \leq |\mathcal{Q}|$$

where \mathcal{X} is a system-optimal solution.

Proof. Let $\mathcal{Q} = \{1, \dots, n\}$ be the set of passengers and \mathcal{S}^q for $q = 1, \dots, n$ the route combination after choosing the best response R_q for passenger q , i.e., $\mathcal{S}^q = (R_1, R_2, \dots, R_q, \emptyset, \emptyset, \dots, \emptyset)$. Furthermore let \mathcal{X} be the socially optimal solution, where the passengers choose the route X_q , $q = 1, \dots, n$ with corresponding paths Y_q in the CGN.

Since arc weight functions are non-increasing, it holds that $g_q(\mathcal{S}^n) \leq g_q(\mathcal{S}^q)$. Since R_q is a best response to $(R_1, R_2, \dots, R_{q-1}, \emptyset, \emptyset, \dots, \emptyset)$ we have

$$g_q(\mathcal{S}^q) = \sum_{a \in P_q} \bar{w}_a(x_a(\mathcal{S}^q)) = \sum_{a \in P_q} \bar{w}_a(x_a(\mathcal{S}^{q-1}) + 1) \leq \sum_{a \in Y_q} \bar{w}_a(x_a(\mathcal{S}^{q-1}) + 1) \quad (9)$$

where P_q denotes the path in the CGN corresponding to R_q . With this, it holds

$$\begin{aligned}
G(\mathcal{S}^n, f(\mathcal{S}^n)) &= \sum_{q \in \mathcal{Q}} g_q(\mathcal{S}^n) \leq \sum_{q \in \mathcal{Q}} g_q(\mathcal{S}^q) \\
&\leq \sum_{q \in \mathcal{Q}} \sum_{a \in Y_q} \bar{w}_a(x_a(\mathcal{S}^{q-1}) + 1) && \text{due to (9)} \\
&\leq \sum_{q \in \mathcal{Q}} \sum_{a \in Y_q} \bar{w}_a(1) && \text{since } \bar{w}_a \text{ non-increasing} \\
&\leq \sum_{q \in \mathcal{Q}} \sum_{a \in Y_q} x_a(\mathcal{X}) \cdot \bar{w}_a(x_a(\mathcal{X})) && \text{since } \bar{w}_a(1) \leq x \bar{w}_a(x) \\
&\leq |\mathcal{Q}| \sum_{q \in \mathcal{Q}} \sum_{a \in Y_q} \bar{w}_a(x_a(\mathcal{X})) \\
&= |\mathcal{Q}| \sum_{q \in \mathcal{Q}} g_q(\mathcal{X}^n) = |\mathcal{Q}| \cdot G(\mathcal{X}^n, f(\mathcal{X}^n)).
\end{aligned}$$

□

Corollary 4.9. *If there exist non-increasing arc weight functions \bar{w}_a with $\bar{w}_a(1) \leq x \cdot \bar{w}_a(x)$ for all $x \in \mathbb{N}$, a solution (\mathcal{R}, f) to the line planning problem with $\frac{G(\mathcal{R}, f)}{G(\mathcal{R}^*, f)} \leq |\mathcal{Q}|$ can be found in polynomial time.*

As described above, we can show that this bound is tight, i.e., there are instances where the best-response algorithm can get stuck in an equilibrium whose objective value is $|\mathcal{Q}|$ -times the optimal solution value.

4.5 Heuristic approaches to the routing problem

In the preceding Sections 4.2-4.4 we have seen that in order to achieve polynomial running time of the best-response algorithm, to be able to prove convergence to an equilibrium, and give bounds on the quality of an equilibrium, strong restrictions on the parameters of the objective function have to be imposed.

In this section we investigate heuristic approaches to the routing problem for the general case, i.e., for individual objective functions

$$g_q(R_q, \mathcal{R}^{-q}) = \text{travel}_q(R_q, \mathcal{R}^{-q}) + \text{transfer}_q(R_q) + \text{cost}_q(R_q, \mathcal{R}^{-q})$$

with edge-based costs

$$\text{cost}_q(R_q, \mathcal{R}^{-q}) := \sum_{(e,l) \in R_q} \frac{\text{cost}_{(e,l)}(f(R_q, \mathcal{R}^{-q}))}{x_{(e,l)}(R_q, \mathcal{R}^{-q})}.$$

In this general case, the routing problem is NP-hard (Theorem 4.3) and the best-response algorithm does not necessarily converge (see Section 4.3.1). To overcome these difficulties, in this section we simplify the transfer time function τ_q and the edge-based cost function cost_q .

4.5.1 Auxiliary frequencies

The first approach is to replace the frequencies $f(\mathcal{R})$ by *auxiliary frequencies* $\tilde{f}(\mathcal{R}^{-q})$ when determining a route for passenger q . This small trick allows us to define arc weights in accordance to Lemma 4.4 and hence, to solve the routing problem using Dijkstra's algorithm in the CGN.

Let \mathcal{Q} be a set of passengers and let $\mathcal{R} = \{R_q : q \in \mathcal{Q}\}$ be a set of strategies represented by paths in the CGN. We call an *edge* $(e, l) \in \mathcal{A}$ *critical for* \mathcal{R} if one additional passenger on the edge would increase the frequency, i.e., if $x_{(e,l)}(\mathcal{R}) \equiv 0 \pmod{B}$. A *line* $l \in \mathcal{L}$ is *critical for* \mathcal{R} if it contains an

edge which is critical for \mathcal{R} . In order to find a route, represented by a path P_q in the CGN, given the routes for all other passengers \mathcal{R}^{-q} , we define the auxiliary frequencies

$$\tilde{f}_l(\mathcal{R}^{-q}) := \begin{cases} f_l(\mathcal{R}^{-q}) + 1 & \text{if } l \text{ is critical for } \mathcal{R}^{-q} \\ f_l(\mathcal{R}^{-q}) & \text{otherwise.} \end{cases}$$

We observe that for every line l and every passenger $q \in \mathcal{Q}$,

$$\tilde{f}_l(\mathcal{R}^{-q}) \geq f_l(\mathcal{R}) \geq f_l(\mathcal{R}^{-q}). \quad (10)$$

For all non-critical lines we even have equality. Plugging in the auxiliary frequencies into τ_q we obtain an *auxiliary* transfer time function

$$\tilde{\tau}_q^{\text{lb}}(\mathcal{R}) := \sum_{i=1}^{n-1} \frac{1}{\tilde{f}_{l_i}(\mathcal{R}^{-q}) + \tilde{f}_{l_{i+1}}(\mathcal{R})}$$

(where l_1, \dots, l_n are the lines used in R_q) which *underestimates* the transfer times $\tau(\mathcal{R})$ in a routing \mathcal{R} . To find an *overestimating* heuristic measure for transfer times, we can consider

$$\tilde{\tau}_q^{\text{ub}}(\mathcal{R}) := \tau_q^{\text{ub}}(\mathcal{R}^{-q}) = \sum_{i=1}^{n-1} \frac{1}{f_{l_i}(\mathcal{R}^{-q}) + f_{l_{i+1}}(\mathcal{R}^{-q})}.$$

Using the same approach, we can define overestimating auxiliary edge-based cost functions as

$$\tilde{\text{cost}}_q^{\text{ub}}(\mathcal{R}) := \sum_{(e,l) \in R_q} \frac{\text{cost}_{(e,l)}(\tilde{f}(\mathcal{R}^{-q}))}{x_{(e,l)}(R_q, \mathcal{R}^{-q})} \geq \text{cost}_q(\mathcal{R})$$

and underestimating auxiliary edge-based cost functions

$$\tilde{\text{cost}}_q^{\text{lb}}(\mathcal{R}) := \text{cost}_q(\mathcal{R}^{-q}) = \sum_{(e,l) \in R_q} \frac{\text{cost}_{(e,l)}(f(\mathcal{R}^{-q}))}{x_{(e,l)}(R_q, \mathcal{R}^{-q})} \leq \text{cost}_q(\mathcal{R}_q).$$

We define over- and underestimated versions of the individual objective functions

$$\begin{aligned} \tilde{g}_q^{\text{ub}}(R_q, \mathcal{R}^{-q}) &:= \alpha_1 c(\mathcal{R}) + \alpha_2 \tilde{\tau}^{\text{ub}}(\mathcal{R}^{-q}) + \text{transfer}_q(R_q) + \tilde{\text{cost}}_q^{\text{ub}}(R_q, \mathcal{R}^{-q}), \\ \tilde{g}_q^{\text{lb}}(R_q, \mathcal{R}^{-q}) &:= \alpha_1 c(\mathcal{R}) + \alpha_2 \tilde{\tau}^{\text{lb}}(\mathcal{R}^{-q}) + \text{transfer}_q(R_q) + \tilde{\text{cost}}_q^{\text{lb}}(R_q, \mathcal{R}^{-q}) \end{aligned}$$

and obtain

$$\tilde{g}_q^{\text{lb}}(R_q, \mathcal{R}^{-q}) \leq g_q(R_q, \mathcal{R}^{-q}) \leq \tilde{g}_q^{\text{ub}}(R_q, \mathcal{R}^{-q}).$$

Given a passenger q and a set of strategies \mathcal{R}^{-q} for the remaining passengers, the auxiliary frequencies allow us to define weights for the arcs in the CGN which depend only on the strategy choices of the remaining passengers \mathcal{R}^{-q} . This observation is summarized in the following lemma.

Lemma 4.10. *For arc weights*

$$\begin{aligned} \tilde{w}_a^{\text{ub}}(\mathcal{R}^{-q}) &:= \begin{cases} \alpha_1 c(e) + \frac{\text{cost}_{(e,l)}(\tilde{f}(\mathcal{R}^{-q}))}{x_a(\mathcal{R}^{-q})+1} & \forall a = (e, l) \in \mathcal{A}_{\text{drive}} \\ \frac{1}{f_l(\mathcal{R}^{-q}) + f_{l'}(\mathcal{R}^{-q})} + \beta & \forall a = ((v, l), (v, l')) \end{cases} \\ \text{or } \tilde{w}_a^{\text{lb}}(\mathcal{R}^{-q}) &:= \begin{cases} \alpha_1 c(e) + \frac{\text{cost}_{(e,l)}(f(\mathcal{R}^{-q}))}{x_a(\mathcal{R}^{-q})+1} & \forall a = (e, l) \in \mathcal{A}_{\text{drive}} \\ \frac{1}{f_l(\mathcal{R}^{-q}) + f_{l'}(\mathcal{R}^{-q})} + \beta & \forall a = ((v, l), (v, l')) \end{cases} \end{aligned}$$

we have

$$\tilde{g}_q^{\text{ub}}(R_q, \mathcal{R}^{-q}) = \sum_{a \in P_q} \tilde{w}_a^{\text{ub}}(\mathcal{R}^{-q}) \text{ and } \tilde{g}_q^{\text{lb}}(R_q, \mathcal{R}^{-q}) = \sum_{a \in P_q} \tilde{w}_a^{\text{lb}}(\mathcal{R}^{-q})$$

(where P_q denotes the path in the CGN corresponding to R_q) and the routing problem can be solved in polynomial time.

Here, the last statement follows from Lemma 4.4.

However, the use of the auxiliary objective functions \tilde{g}_q does not guarantee the existence of an equilibrium: In fact, in the counter example shown in Section 4.3.1 we have $f_l(\mathcal{R}) = \tilde{f}_l(\mathcal{R}^{-q})$ for all choices of q and R_q . Hence, this example also proves the non-existence of an equilibrium for objective functions \tilde{g}_q .

4.5.2 Auxiliary arc weights

Since the heuristic from section 4.5.1 does not always lead to an equilibrium, we consider a further simplification which guarantees the existence of an equilibrium and the convergence of the best-response-algorithm.

Consider a set of passenger routes $\mathcal{R} = \{R_q : q \in \mathcal{Q}\}$ in the CGN and a transfer edge $a = ((v, l), (v, l'))$. Then the frequency of l and l' respectively is at least $\left\lceil \frac{x_a(\mathcal{R})}{B} \right\rceil$, since at least all passengers changing from l to l' have to use l and l' respectively. Additionally all frequencies are at most $\left\lceil \frac{|\mathcal{Q}|}{B} \right\rceil$ since no more than all passengers can use any given line. This leads to the following approximate arc weight functions:

$$\bar{w}_a^{\text{lb}}(x) := \begin{cases} \alpha_1 c(e) + \frac{\gamma_1 \cdot k_1^l + \gamma_2 \cdot k_2^l \cdot \left\lceil \frac{x}{B} \right\rceil}{x} & \text{if } a = (e, l) \in \mathcal{A}_{\text{drive}} \\ \frac{\alpha_2}{2 \cdot \left\lceil \frac{|\mathcal{Q}|}{B} \right\rceil} + \beta & \text{if } a \in \mathcal{A}_{\text{transfer}} \end{cases} \quad (11)$$

and

$$\bar{w}_a^{\text{ub}}(x) := \begin{cases} \alpha_1 c(e) + \frac{\gamma_1 \cdot k_1^l + \gamma_2 \cdot k_2^l \cdot \left\lceil \frac{|\mathcal{Q}|}{B} \right\rceil}{x} & \text{if } a = (e, l) \in \mathcal{A}_{\text{drive}} \\ \frac{\alpha_2}{2 \cdot \left\lceil \frac{x}{B} \right\rceil} + \beta & \text{if } a \in \mathcal{A}_{\text{transfer}} \end{cases}. \quad (12)$$

With $\bar{g}_q^{\text{lb}}(R_q, \mathcal{R}^{-q}) := \sum_{a \in P_q} \bar{w}_a^{\text{lb}}(x_a)$ and $\bar{g}_q^{\text{ub}}(P_q, \mathcal{R}^{-q}) := \sum_{a \in P_q} \bar{w}_a^{\text{ub}}(x_a)$, (where P_q is the path in the CGN corresponding to R_q) we obtain:

Lemma 4.11. *For every passenger $p \in \mathcal{Q}$ with route R_q and $\mathcal{R} = (R_q, \mathcal{R}^{-q})$ we have*

$$\bar{g}_q^{\text{lb}}(R_q, \mathcal{R}^{-q}) = \sum_{a \in P_q} \bar{w}_a^{\text{lb}}(x_a) \leq g_q(R_q) \leq \sum_{a \in P_q} \bar{w}_a^{\text{ub}}(x_a) = \bar{g}_q^{\text{ub}}(R_q, \mathcal{R}^{-q}).$$

From Lemma 4.4, Lemma 4.5, Lemma 4.6 and Lemma 4.8 we conclude:

Lemma 4.12. *For individual objective functions \bar{g}_q^{lb} and \bar{g}_q^{ub} , the routing step of the best-response algorithm can be executed in polynomial time using arc weights $\bar{w}_a^{\text{lb}}(x_a)$ or $\bar{w}_a^{\text{ub}}(x)$, respectively, in the CGN.*

With respect to these objective functions equilibria exist, the best-response algorithm converges towards an equilibrium, and the price of anarchy is at most $|\mathcal{Q}|$. Furthermore, given an empty state of the so-defined game, calculating the best response once for every passenger leads to a state with costs of at most number of passengers times the costs of the social optimal solution.

5 Experiments

5.1 Test instances

In order to find out how our best-response algorithm behaves and to gain first insights on its performance, we tested our best-response algorithm it on the three small test instances illustrated in Figures 5-7. The figure on the left shows the PTN with travel times and the lines, on the right

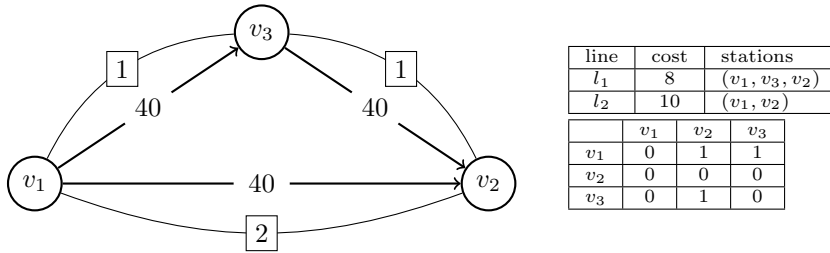


Figure 5: Description of instance I1

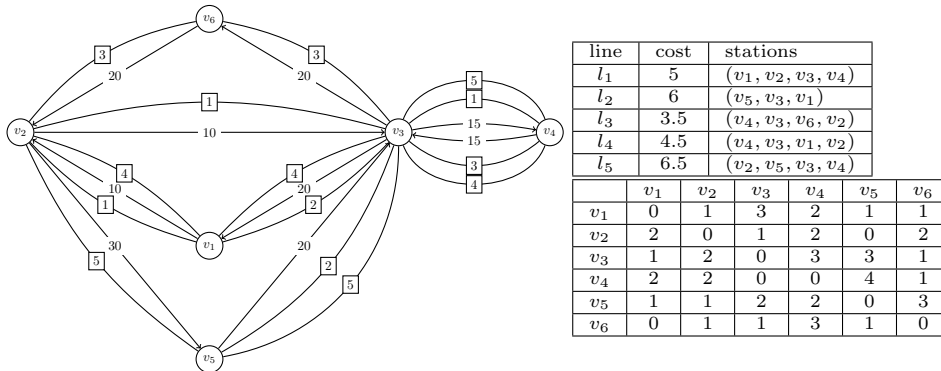


Figure 6: Description of instance I2

line costs and OD-matrix are given. I1, shown in Figure 5 is our smallest example consisting of only three stations, two lines, and three passengers. I2 and I3 are bigger, the former consists of six stations, five lines, and 49 passengers, the latter contains eight stations, eight lines, and 2620 passengers.

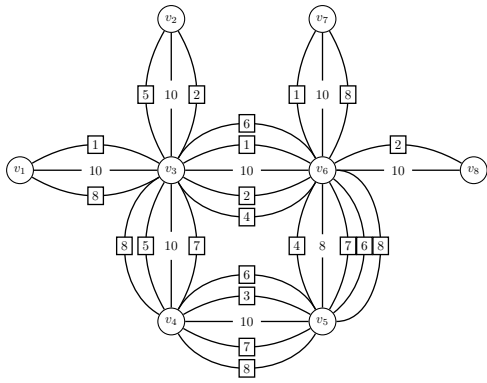
Additionally we tested the algorithm on two more realistic instances GOE1 and GOE2 which are taken from the LinTim toolbox [Lin14]. These instances are both based on the PTN shown in Figure 8 which is based on the bus-network in Göttingen, Germany. It consists of 257 stations and 548 edges. A line pool consisting of 44 lines was generated for these experiments. The instances GOE1 and GOE2 differ in the number of passengers. For our experiments we modified the original OD-data from LinTim toolbox which consisted of 460146 passengers on 66048 OD-pairs. For GOE2, the number of passengers are scaled down by the factor $\frac{1}{30}$ and rounded, which leads to 6321 passengers on 6114 OD-pairs. For GOE1, the same calculation is done with the factor $\frac{1}{50}$, leading to OD-data of 1943 passengers and as many OD-pairs

5.2 Running time

Table 1 gives a short overview of some computational aspects for using auxiliary arc-weight \bar{w}_a^{lb} described in Section 4.5.2 with parameter set $(\alpha_1/\alpha_2, \beta, \gamma_1/\gamma_2) = (0.5/0.5, 0.5, 0.1/0.1)$ for I1, I2 and I3 and $(\alpha_1/\alpha_2, \beta, \gamma_1/\gamma_2) = (0.5/0.01, 0.1, 0.01/0.01)$ for GOE1 and GOE2. The number of iterations for different parameter choices and for the other heuristics described in Sections 4.5.1

parameters:	I1	I2	I3	GOE1	GOE2
# iterations	2	3	3	3	3
computation time per iteration (in seconds)	0.01	0.045	2.26	45.78	143.35
% pass. without transfer	100	55.1	82.44	3.45	6.3

Table 1: Characteristica of found solutions for different instances for individual cost function (11).



line	cost	stations
l_1	4	(v_1, v_3, v_6, v_7)
l_2	4	(v_2, v_3, v_6, v_8)
l_3	1.8	(v_4, v_5)
l_4	2.8	(v_3, v_6, v_5)
l_5	3	(v_2, v_3, v_4)
l_6	4	(v_3, v_6, v_5, v_4)
l_7	3.8	(v_3, v_4, v_5, v_6)
l_8	5.8	$(v_1, v_3, v_4, v_5, v_6, v_7)$

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
v_1	0	10	20	10	60	100	0	0
v_2	10	0	40	20	70	100	0	10
v_3	20	40	0	20	100	50	20	10
v_4	10	20	20	0	100	150	0	0
v_5	60	70	100	100	0	120	100	90
v_6	100	100	50	150	120	0	60	50
v_7	0	0	20	0	100	60	0	0
v_8	0	10	10	0	90	50	0	0

Figure 7: Description of instance I3



Figure 8: Underlying public transportation network for instances GOE1 and GOE2

	auxiliary freq ub	auxiliary freq lb	aux. arc-weight ub	aux. arc-weight lb
# iterations	4	5	4	3
<u>travel time</u>				
drive	16.85	16.79	16.85	16.5
transfer real(est.)	1.87(1.87)	1.92(1.9)	0.54(2.06)	0.72(0.1)
total real(est.)	18.72(18.72)	18.72(18.69)	17.4(18.92)	17.22(16.6)
# transfers	1210	1530	340	460
% pass. without transfer	54.58	47.33	87.02	82.44
<u>lines</u>				
# lines	7	9	8	11
\sum fixed line costs	1884	2352	2016	2664
\sum freq line costs	4840	5800	3800	4352
\sum line costs	6724	8152	5816	7016
objective value	25559.7	25732.82	23542.43	23437.75

Table 2: Comparison of different heuristics for instance I3, using parameter set (0.5/0.5, 0.5, 0.1/0.1).

and 4.5.2 is comparable.

We observe that the time needed for one iteration increases with increasing network size and in particular with increasing number of passengers. This is to be expected, since each iteration consists of finding one shortest path per passenger and a subsequent adaption of arc lengths in the network. Note that in instances GOE1 and GOE2, the solution time *per passenger* is roughly the same in both instances. Furthermore, it is interesting to see that the number of iterations did not increase with the problem size. This indicates that with respect to running time, the approach could also be applicable for bigger instances, in particular in conjunction with speed-up techniques for shortest paths.

The fourth row in Table 1 shows the percentage of passengers who can travel without a transfer from origin to destination. This number reveals the different levels of closeness to reality of the five instances. While in the smaller instances all or nearly all passengers can travel without transfers, this number drops dramatically for instance GOE1 and GOE2. Comparing GOE1 and GOE2, it is interesting to see that less passengers have to transfer in GOE2. This can be explained by the fact that in case of higher demand, more lines can be established and hence there are more choices for the passengers.

5.3 Comparison of the different heuristic approaches to include frequencies

In this section we compare the different heuristic approaches from Section 4.5 on the instances I3 and GOE2, using the parameter sets $(\alpha_1/\alpha_2, \beta, \gamma_1/\gamma_2) = (0.5/0.5, 0.5, 0.1/0.1)$ and $(\alpha_1/\alpha_2, \beta, \gamma_1/\gamma_2) = (0.5/0.01, 0.1, 0.01/0.01)$ respectively. The results are shown in Table 2 and Table 3, where *auxiliary freq ub* and *auxiliary freq lb* use the individual objective functions \tilde{g}_q^{ub} and \tilde{g}_q^{lb} described in Section 4.5.1, respectively, and *aux. arc-weight ub* and *aux. arc-weight lb* denote the heuristics based on auxiliary arc weights \bar{w}_a^{ub} and \bar{w}_a^{lb} as described in Section 4.5.2, respectively.

We compare the average travel time per passenger, split up into driving time and transfer time, the total number of transfers, the percentage of passengers without a transfer, the number of lines, the costs of the line concept, split up into fixed costs and costs depending on the frequency and the objective value (according to the given parameter set) as sum of the objective values of all passengers.

For the transfer time and travel time, we show both the exact value (calculated using g) and the value estimated by the heuristic individual objective functions (in brackets).

We observe that, as proven in Section 4.5, transfer times are overestimated in the heuristics *auxiliary freq ub* and *aux. arc-weight ub*, while they are underestimated in *auxiliary freq lb* and *aux. arc-weight lb*. As was to be expected, the estimated transfer time in the heuristics using auxiliary frequencies is much closer to the exact transfer time than in the heuristics using auxiliary arcs weights.

Comparing heuristics based on auxiliary frequencies and heuristics based on auxiliary arc weights

	auxiliary freq ub	auxiliary freq lb	edge-based ub	edge-based lb
# iterations	5	4	6	3
<u>travel time</u>				
drive	49.55	50	42.23	41.91
transfer real(est.)	9.9(9.93)	9.9(9.89)	9.36(27.48)	9.51(0.88)
total real(est.)	59.45(59.48)	59.89(59.89)	51.59(59.71)	51.42(42.79)
# transfers	24927	25612	15838	15688
% pass. without transfer	2.77	2.09	6.83	6.3
<u>lines</u>				
# lines	37	38	39	39
\sum fixed line costs	2398710	2402736	2542626	2542626
\sum freq line costs	3476608	3649197	2671705	2672386
\sum line costs	5875318	6051933	5214331	5215012
objective value	204291.96	206111.89	185927.84	185237.34

Table 3: Comparison of different heuristics for the frequencies for instance GOE2, using parameter set (0.5/0.01, 0.1, 0.01/0.01).

	parameter changes wrt. (0.5/0.5, 0.5, 0.1/0.1)			
	no changes	$\alpha_1 = \alpha_2 = 17.5$	$\alpha_2 = 15$	$\alpha_2 = 0$
# iterations	3	2	3	3
<u>travel time</u>				
drive	16.5	16.5	16.6	16.5
transfer real(est.)	0.72(0.1)	0(0)	0(0)	0.85(0.1)
total real(est.)	17.22(16.6)	16.5(16.5)	16.6(16.6)	17.35(16.6)
# transfers	460	0	0	460
% pass. without transfer	82.44	100	100	82.44
<u>lines</u>				
# lines	11	11	10	11
\sum fixed line costs	2664	2664	2424	2664
\sum freq line costs	4352	4352	4096	4352
\sum line costs	7016	7016	6520	7016
objective value for (0.5/0.5, 0.5, 0.1/0.1)	23437.75	22264.91	22377.71	23437.75

Table 4: Comparison of different parameter sets in the individual objective functions using auxiliary arc weights \bar{w}^{lb} on instance I3.

among each other, we observe that the heuristics which *overestimate* transfer time and line costs, find solutions with less transfers and less lines. However, travel time is slightly higher in these cases.

A comparison of the objective values of the results provided by the different heuristics surprisingly shows that the simple heuristics using auxiliary arc weights provide better results than the heuristics using auxiliary frequencies for both the small instance I3 and the more realistic instance GOE2 - although the latter provide much better estimates of travel times. In particular, the heuristics based on auxiliary arc weights seem to find paths with a smaller overall number of transfers and a bigger percentage of direct travelers. This effect is a bit more pronounced for the heuristic *aux. arc-weight ub* which *overestimates* transfer times, but also clearly visible for *aux. arc-weight lb*.

5.4 Influence of the parameters

In order to investigate how the choice of parameters in the objective function influences the outcome of the best-response algorithm, in Table 4 and 5 we vary the parameters in the heuristic individual objective functions used in *aux. arc-weight lb* on instance 3.

We observe that the equilibrium the algorithm found for the given parameter set is not the best one for the objective function for this parameter set. Surprisingly, increasing the weighting parameter for travel-time (α_1), transfer-time (α_2) or transfers (β) each lead to an equilibrium that is better socially and reduces the number of transfers to 0.

Increasing the influence of the costs - regardless of considering fixed costs, frequency-based costs or both - leads to an socially considerable worse equilibrium which is on top of that more expensive than any other considered here. This undesirable property might be a general disadvantage of the heuristic used here.

parameter changes wrt. (0.5/0.5, 0.5, 0.1/0.1)				
	no changes	$\beta = 20$	$\gamma_1 = \gamma_2 = 0.75/\gamma_1 = 150/\gamma_2 = 150$	$\beta = \gamma_1 = \gamma_2 = 0$
# iterations	3	4	2	
travel time				
drive	16.5	16.6	22.29	16.5
transfer real(est.)	0.72(0.1)	0(0)	2.43(0,6)	0.85(0.1)
total real(est.)	17.22(17.6)	16.6(16.6)	24.72(22.89)	17.35(17.6)
# transfers	460	0	2780	460
% pass. without transfer	82.44	100	35.11	82.44
lines				
# lines	11	10	8	12
\sum fixed line costs	2664	2424	1800	2772
\sum freq line costs	4352	4096	7080	3580
\sum line costs	7016	6520	8880	6352
objective value for (0.5/0.5, 0.5, 0.1/0.1)	23437.75	22377.71	34339.9	23549.8

Table 5: Comparison of different parameter set in the individual objective functions using auxiliary arc weights $\bar{w}lb$ on instance I3.

The different results are also shown in Figures 9(a)-9(d), where the thickness of the line arcs corresponds to the number of passengers using it or the frequencies of the lines respectively. The direct edges correspond to edges in the PTN, the bent edges to edges in the CGN.

The first image in Figure 9(a) is the solution for the original parameter set. For Figure 9(b), the transfer time is taken into account more which leads to a higher usage of the long line 8, leading to fewer transfers in this instance. A similar picture can be derived by increasing β , i.e. the transfer penalty. In Figure 9(c) less lines are used, due to the increased cost influence in the solutions. A corresponding effect is the higher load on the remaining lines. Additionally one edge is not used anymore in this solution. Note that the solution is exactly the same regardless of which cost parameter is increased. For the last solution 9(d), only travel and transfer times are taken into account. This results in a very level picture regarding the usage of the different lines, probably because taking new lines into the solution does not cost more and lines with similar frequencies are nice for the transfer times.

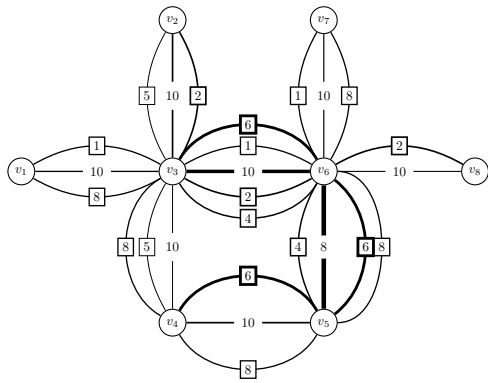
6 Conclusions and further research

We presented a new approach to solve line planning by solving a routing game where the passengers are the players who aim at minimizing a weighted sum of their travel time, transfer penalties, and a cost share. Under strong assumptions on the objective function (transfer time is not taken into account and line costs can be assigned to edges and are independent of frequencies) equilibria of this game can be found using the described best-response algorithm. In case that the objective function does not fulfill these properties, applicability and convergence of the best-response approach can be achieved by a slight modification of the individual objective functions.

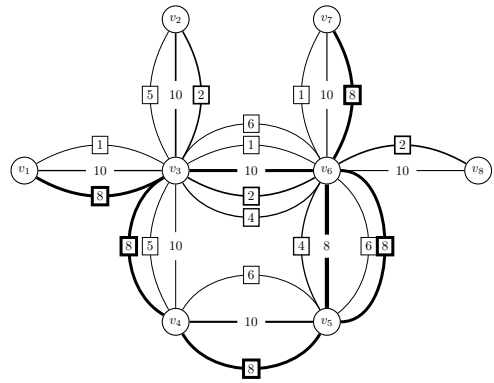
A logical next step will be to evaluate whether the line planning routing game, besides being an interesting object of study in itself, does indeed lead to a good heuristic for line planning.

First, more experiments of the type presented in Section 5 on instances of realistic size (in particular also with respect to passenger numbers) may lead to more insights on the performance of the different approaches presented in Section 4.5. A positive effect of increasing passenger numbers is that the approximate frequencies $f(\mathcal{R}^{-q})$ and $\tilde{f}(\mathcal{R}^{-q})$ become better estimates of actual frequencies $f(\mathcal{R})$. However, in the current version of the best-response strategy, in each iteration a shortest path for each passenger has to be found, hence running time increases with increasing number of passengers. For large passenger numbers it may thus make sense to use flow equilibration techniques in the inner loop instead of shortest path computations for each individual passenger.

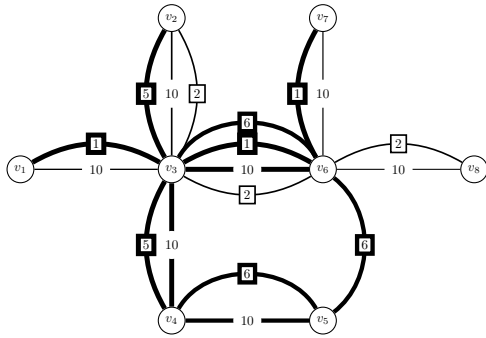
Second, line planning solutions obtained with the routing game approach should be compared to state-of-the-art exact and heuristic solution methods for line planning with respect to objective value, running time, and practicability of the found solution (in the sense of Section 3.5).



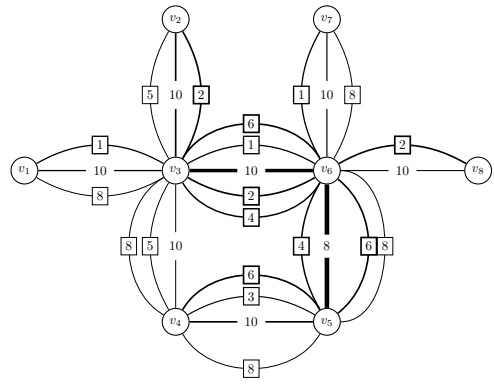
(a) Parameter set $(0.5/0.5, 0.5, 0.1/0.1)$.



(b) Parameter set $(0.5/7.5, 0.5, 0.1/0.1)$.



(c) Parameter set $(0.5/0.5, 0.5, 7.5/7.5) / (0.5/0.5, 0.5, 15/0.1) / (0.5/0.5, 0.5, 0.1/15)$.



(d) Parameter set $(0.5/0.5, 0, 0/0)$.

Figure 9: Dependence of equilibrium solution from choice of parameters using the auxiliary arc weights \bar{w}/b on instance I3.

While the terms for travel time and transfers are quite intuitive, many different choices are possible for the cost-sharing among passengers. It remains an interesting question how to divide operational costs among passengers such that, on the one hand, the algorithmic approach is still viable, and on the other hand, cost shares are comparable to real-world travel costs. Furthermore, it would be interesting to investigate whether the routing game approach can also be applied to other planning problems which can be considered integrated network design and routing problem like, e.g., timetabling or delay management with integrated routing.

References

- [AAE05] Baruch Awerbuch, Yossi Azar, and Amir Epstein. The price of routing unsplittable flow. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 57–66. ACM, 2005.
- [ADK⁺04] E. Anshelevich, A. Dasgupta, J. Kleinberg, E. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. pages 295 – 304, 2004.
- [BGP07] R. Borndörfer, M. Grötschel, and M. E. Pfetsch. A column-generation approach to line planning in public transport. *Transportation Science*, 41(1):123–132, 2007.
- [BGP08] R. Borndörfer, M. Grötschel, and M. E. Pfetsch. Models for line planning in public transport. In M. Hickman, P. Mirchandani, and S. Voß, editors, *Computer-aided Systems in Public Transport*, volume 600 of *Lecture Notes in Economics and Mathematical Systems*, pages 363–378. Springer Berlin Heidelberg, 2008.
- [BHK⁺13] Ralf Borndörfer, Heide Hoppmann, Marika Karbstein, et al. A configuration model for the line planning problem. In *ATMOS-13th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems-2013*, volume 33, pages 68–79, 2013.
- [BK12] R. Borndörfer and M. Karbstein. A Direct Connection Approach to Integrated Line Planning and Passenger Routing. In D. Delling and L. Liberti, editors, *12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*, volume 25 of *OpenAccess Series in Informatics (OASICS)*, pages 47–57, Dagstuhl, Germany, 2012. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- [BKZ97] M. R. Bussieck, P. Kreuzer, and U. T. Zimmermann. Optimal lines for railway systems. *European Journal of Operational Research*, 96(1):54–63, 1997.
- [BKZ09] Apostolos Bessas, Spyros Kontogiannis, and Christos Zaroliagis. Incentive-compatible robust line planning. In *Robust and Online Large-Scale Optimization*, pages 85–118. Springer, 2009.
- [BKZ11] Apostolos Bessas, Spyros Kontogiannis, and Christos Zaroliagis. Robust line planning in case of multiple pools and disruptions. In *Theory and Practice of Algorithms in (Computer) Systems*, pages 33–44. Springer, 2011.
- [BLL04] M. R. Bussieck, T. Lindner, and M. E. Lübbecke. A fast algorithm for near cost optimal line plans. *Mathematical Methods of Operations Research*, 59:205–220, 2004.
- [BN10] R. Borndörfer and M. Neumann. Models for line planning with transfers. Technical Report 10-11, ZIB, Takustr.7, 14195 Berlin, 2010.
- [Bus98] M. Bussieck. *Optimal Lines in Public Rail Transport*. PhD thesis, Technische Universität Braunschweig, 1998.

- [CF95] I. Constantin and M. Florian. Optimizing frequencies in a transit network: a nonlinear bi-level programming approach. *International Transactions in Operational Research*, 2(2):149, 1995.
- [CvDZ98] M. T. Claessens, N. M. van Dijk, and P. J. Zwaneveld. Cost optimal allocation of rail passenger lines. *European Journal of Operational Research*, 110:474–489, 1998.
- [CW91] A. Ceder and N. H. M. Wilson. Bus network design. *Transportation Research Part B*, 20(4):331–344, 1991.
- [DCF93] Joaquin De Cea and Enrique Fernández. Transit assignment for congested public transport systems: an equilibrium model. *Transportation science*, 27(2):133–147, 1993.
- [Die78] H. Dienst. *Linienplanung im spurgeführten Personenverkehr mit Hilfe eines heuristischen Verfahrens*. PhD thesis, Technische Universität Braunschweig, 1978. (in German).
- [FGP02] Gaetano Fusco, Stefano Gori, and Marco Petrelli. A heuristic transit network design algorithm for medium size towns. *Proceedings of 9th Euro working group on transportation, Bari, Italy*, pages 652–656, 2002.
- [FM06a] Wei Fan and Randy B. Machemehl. Optimal transit route network design problem with variable transit demand: genetic algorithm approach. *Journal of transportation engineering*, 132(1):40–51, 2006.
- [FM06b] Wei Fan and Randy B Machemehl. Using a simulated annealing algorithm to solve the transit route network design problem. *Journal of transportation engineering*, 132(2):122–132, 2006.
- [GvHK04] J.-W. Goossens, S. van Hoesel, and L. Kroon. A branch-and-cut approach for solving railway line planning problems. *Transportation Science*, 38(3):379–393, 2004.
- [GvHK06] J.-W. Goossens, S. van Hoesel, and L. Kroon. On solving multi-type railway line planning problems. *European Journal of Operational Research*, 168(2):403–424, 2006. Feature Cluster on Mathematical Finance and Risk Management.
- [GYW06] J. F. Guan, Hai Yang, and S.C. Wirasinghe. Simultaneous optimization of transit line configuration and passenger line assignment. *Transportation Research Part B*, 40(10):885–902, 2006.
- [JBT10] L’ Jánošíková, M Blatoň, and D Teichmann. Design of urban public transport lines as a multiple criteria optimisation problem. In *Proc. of the 16th International Conference on Urban Transport and the Environment—Urban Transport XVI*, pages 137–146, 2010.
- [Lin14] LinTim - Integrated optimization in public transportation, 2014. [Online, <http://lintim.math.uni-goettingen.de>; accessed 25-November-2014].
- [LMP10] Gilbert Laporte, Juan A Mesa, and Federico Perea. A game theoretic framework for the robust railway transit network design problem. *Transportation Research Part B: Methodological*, 44(4):447–459, 2010.
- [Man80] C. E. Mandl. Evaluation and optimization of urban public transportation networks. *European Journal of Operational Research*, 5(6):396–404, 1980.
- [Neu14] Andreas Neumann. *A paratransit-inspired evolutionary process for public transit network design*. PhD thesis, 2014.
- [NP88] S. Nguyen and S. Pallottino. Equilibrium traffic assignment for large scale transit networks. *European Journal of Operational Research*, 37(2):176 – 186, 1988.

- [NRTV07] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic game theory*. Cambridge University Press, 2007.
- [Pat25] A Patz. Die richtige auswahl von verkehrslinien bei großen strassenbahnnetzen. *Verkehrstechnik*, 50:51, 1925.
- [PB06] M. Pfetsch and R. Borndörfer. Routing in line planning for public transport. In H.-D. Haasis, H. Kopfer, and J. Schönberger, editors, *Operations Research Proceedings 2005*, volume 2005 of *Operations Research Proceedings*, pages 405–410. Springer Berlin Heidelberg, 2006.
- [PRR95] Uwe Pape, Yean-Suk Reinecke, and Erwin Reinecke. Line network planning. In JoachimR. Daduna, Isabel Branco, and Jos M. Pinto Paixo, editors, *Computer-Aided Transit Scheduling*, volume 430 of *Lecture Notes in Economics and Mathematical Systems*, pages 1–7. Springer Berlin Heidelberg, 1995.
- [Qua03] CB Quak. Bus line planning. *Master’s thesis, TU Delft*, 2003.
- [Ros73] R. W. Rosenthal. The network equilibrium problem in integers. *Networks*, 3(1):53–59, 1973.
- [Rou05] T. Roughgarden. *Selfish routing and the price of anarchy*. MIT press, 2005.
- [Rou07] Tim Roughgarden. Routing games. In N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, editors, *Algorithmic game theory*, volume 18. Cambridge University Press New York, 2007.
- [Sch09] Silvia Schwarze. *Path player games: analysis and applications*, volume 24. Springer, 2009.
- [Sch11] A. Schöbel. Line planning in public transportation: models and methods. *OR Spectrum*, pages 1–20, 2011. 10.1007/s00291-011-0251-6.
- [Sch14] Marie Schmidt. *Integrating Routing Decisions in Public Transportation Problems*. Optimization and its Applications. Springer, 2014.
- [SF89] Heinz Spiess and Michael Florian. Optimal strategies: a new assignment model for transit networks. *Transportation Research Part B: Methodological*, 23(2):83–102, 1989.
- [SFS⁺11] Jan-Dirk Schmcker, Achille Fonzone, Hiroshi Shimamoto, Fumitaka Kurauchi, and Michael G.H. Bell. Frequency-based transit assignment considering seat capacities. *Transportation Research Part B: Methodological*, 45(2):392 – 408, 2011.
- [She85] Y. Sheffi. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, 1985.
- [Son79] H. Sonntag. Ein heuristisches Verfahren zum Entwurf nachfrageorientierter Linienführung im öffentlichen Personennahverkehr. *ZOR - Zeitschrift für Operations-Research*, 23:B15, 1979. (in German).
- [SS06a] A. Schöbel and S. Scholl. Line planning with minimal transfers. In *5th Workshop on Algorithmic Methods and Models for Optimization of Railways*, number 06901 in Dagstuhl Seminar Proceedings, 2006.
- [SS06b] Anita Schöbel and Silvia Schwarze. A game-theoretic approach to line planning. In *Proceedings of ATMOS*, 2006.
- [SS13] A. Schöbel and S. Schwarze. Finding delay-resistant concepts using a game-theoretic approach. *Netnomics*, 14(3):95–117, 2013.

- [SSJ11] WY Szeto, Muthu Solayappan, and Yu Jiang. Reliability-based transit assignment for congested stochastic transit networks. *Computer-Aided Civil and Infrastructure Engineering*, 26(4):311–326, 2011.
- [SW11] W.Y. Szeto and Yongzhong Wu. A simultaneous bus route design and frequency setting problem for tin shui wai, hong kong. *European Journal of Operational Research*, 209(2):141 – 155, 2011.
- [TW07] E. Tardos and T. Wexler. Network formation games and the potential function method. In N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, editors, *Algorithmic Game Theory*. Cambridge University Prss, 2007.

ERIM Report Series <i>Research in Management</i>	
ERIM Report Series reference number	ERS-2014-017-LIS
Date of publication	2014-12-03
Version	03-12-2014
Number of pages	30
Persistent URL for paper	http://hdl.handle.net/1765/77431
Email address corresponding author	schmidt2@rsm.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus University Rotterdam PO Box 1738 3000 DR Rotterdam, The Netherlands Phone: +31104081182 Fax: +31104089640 Email: info@erim.eur.nl Internet: http://www.erim.eur.nl
Availability	The ERIM Report Series is distributed through the following platforms: RePub, the EUR institutional repository Social Science Research Network (SSRN) Research Papers in Economics (RePEc)
Classifications	The electronic versions of the papers in the ERIM Report Series contain bibliographic metadata from the following classification systems: Library of Congress Classification (LCC) Journal of Economic Literature (JEL) ACM Computing Classification System Inspec Classification Scheme (ICS)