Production, Manufacturing, Transportation and Logistics

Railway timetabling with integrated passenger distribution

Johann Hartleb, Marie Schmidt

Abstract

Timetabling for railway services often aims at optimizing travel times for passengers. At the same time, restricting assumptions on passenger behavior and passenger modeling are made. While research has shown that discrete choice models are suitable to estimate the distribution of passengers on routes, this has not been considered in timetabling yet. We investigate how to integrate a passenger distribution into an optimization framework for timetabling and present two mixed integer linear programs for this problem. Both approaches design timetables and simultaneously find a corresponding passenger distribution on available routes. One model uses a linear distribution model to estimate passenger route choices. The other model uses an integrated simulation framework to approximate a passenger distribution according to the logit model, a commonly used route choice model. We compare both new approaches with three state-of-the-art timetabling methods and a heuristic approach on a set of artificial instances and a partial network of Netherlands Railways (NS). Our experiments provide insights into the impact of considering multiple routes instead of a single route, and of integrated route choice versus predetermined route assignment with respect to the solution quality.

© 2021 The Author(s). Published by Elsevier B.V.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

1. Introduction

Public transport is important to our society for various reasons, such as increased mobility for the general public or lower air pollution compared to individual transport. Especially the potential of public transport to reduce emissions is recently much discussed in the context of climate change. To be considered an alternative to individual transport, public transport has to be as attractive as possible to passengers. For decades both researchers and practitioners have been working on the improvement of public transport from different perspectives using various approaches. Most of them follow the same pattern and design public transport sequentially. First, long term planning decisions are taken, such as stop location planning and, in the case of railways, network design. Afterward, the line routes are designed and the corresponding frequencies of lines are fixed. On the tactical level, a timetable is determined, based on the results of the previous steps. Finally, vehicles and crews are scheduled.

Finding a good timetable is an integral step for providing high-quality public transport services to passengers. Next to driving times of vehicles, the timetable determines the transfer times and thereby the travel times of passengers. Since transfer and travel times have a significant effect on the chosen routes of passengers and also their satisfaction with public transport, timetabling is a relevant problem with high practical impact. Moreover, from an algorithmic perspective timetabling is an interesting task since finding a feasible periodic timetable is NP-complete. For this reason, research often focuses on efficient solution strategies. In recent years, many publications deal with the question of how passenger travel time can be used as an objective to guide the search for timetables of high quality.

At the design of public transport systems, a good trade-off between service quality and costs for operating a public transport service has to be found. Since costs are mainly determined by the line plan as well as the vehicle and crew schedule, many optimization approaches for timetabling only aim at providing the best quality to passengers. Even though the focus is on the quality for passengers, strong assumptions on passenger demand are made. Among them, two assumptions are commonly found. First, all passengers travel on their shortest available route. Second, a predetermined passenger assignment to routes is sufficient to estimate passenger loads in the public transport network. In this context, a passenger route defines when and on which lines passengers

https://doi.org/10.1016/j.ejor.2021.06.025
0377-2217/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)
travel. As summarized in Table 1.1, the impact of each of these two assumptions has been studied individually. Improvements could be achieved by considering a passenger distribution on multiple routes and by integrating a shortest-route search into optimization, respectively.

Motivated by these improvements, we relax both assumptions at the same time. We study the problem of finding a travel-time minimal timetable under the assumption that passengers route choice can be modeled using a discrete choice model. To our best knowledge, this is the first time that a choice model is used to derive a passenger distribution within a timetable optimization model.

Depending on the quality of all available routes, discrete choice models estimate the probability that a route is chosen by passengers. This route choice corresponds to a passenger distribution in the network. We use the logit model, a commonly used passenger route choice model in transport applications, to estimate passenger distributions on available routes and incorporate it in an optimization framework for timetabling. Due to the non-linear structure of the logit model, the mathematical program for this problem requires reformulation to be solved exactly. We present two ways to integrate a passenger distribution on multiple routes into a timetabling model as a linear formulation. Our first model uses a novel linear distribution model. This distribution model is designed to have the same characteristics as the logit model. Due to its linear formulation, it can easily be incorporated into an optimization model. The second model relies on a simulation of the logit distribution of passengers. By considering multiple scenarios, the distribution of passengers according to a logit model can be approximated within an optimization model that is linear in all variables.

We aim at maximizing the quality of timetables for passengers. Researchers and practitioners developed a variety of ways to evaluate timetables from the passengers’ perspective. Due to their design, not all of these evaluations are suitable as an objective function in an optimization program. Moreover, Hartleb, Schmidt, Friedrich, & Huisman (2019) showed in an empirical comparison that different evaluation methods do not necessarily yield a consistent valuation of timetables. To best reflect the quality of the found solutions, we evaluate all timetables in our experiments with multiple evaluation functions. As an objective function, the first model uses the absolute travel time to minimize the time spent in the public transport system, which matches common practice in timetabling literature. In the second model, simulated travel times are minimized to incorporate passengers’ preferences that are not captured by absolute travel times only. These preferences can include any kind of non-modeled factors of influence, from differently perceived transfer times through to a popular ice cream shop at a certain transfer station. We discuss the theoretical properties of the chosen objective functions of the two models and analyze their influence on the resulting timetable in the experiments. This discussion suggests that the absolute travel time, although commonly used in literature, might not be suitable for evaluating timetables when considering multiple alternative routes for passengers.

We compare our models for timetabling with integrated passenger distribution with four timetabling methods motivated by approaches in the literature. Two of these methods assume that a passenger assignment to routes is fixed before optimizing the timetable, using either a single route for all passengers traveling between the same stations or a distribution on multiple routes. Another method finds optimal timetables based on the assumption that passengers use the shortest available routes. A fourth approach solves the problem of timetabling with integrated passenger distribution heuristically by iterating between assigning passengers to routes according to the logit model and finding optimal timetables.

We use multiple evaluation functions to thoroughly compare our timetabling methods with the four methods described in the previous paragraph. The comparison shows that the timetables found with our methods performed better concerning some evaluation functions while being of comparable quality concerning other evaluation functions. These improvements come at the expense of increased complexity of the models. From this, we conclude that the integration of a passenger distribution model has the potential to find better timetables for passengers, but more efficient solution strategies have to be developed.

We want to highlight two contributions of this paper. First, we present a novel timetabling model with an integrated choice model to derive a passenger distribution on multiple routes. We provide and discuss alternative representations of the passenger distribution and develop two linear timetabling programs. Second, we show on multiple artificial instances and a partial real-world network the advantages and disadvantages of the novel approaches when compared to state-of-the-art methods. In particular, our experiments provide insight into (1) how considering multiple routes for passengers instead of a single route, and (2) how integrating route choice instead of a predetermined route assignment affects solution quality.

The remainder is structured as follows. We summarize the literature on passenger distribution models, on optimization approaches for timetabling and on the evaluation of timetables in Section 2. In Section 3, the basic models relevant for this paper are introduced and the problem is defined. In Section 4, we develop and discuss two linear timetabling models with an integrated passenger distribution model. Section 5 describes the experimental setup, such as considered instances, benchmark methods, and used evaluation functions. We report and discuss our results of the experiments in Section 6 and conclude in Section 7.

2. Related literature

2.1. Passenger route choice

Passengers may have different preferences that lead to different route choices. State-of-the-art discrete choice models can estimate which routes are chosen by passengers and can predict a passenger distribution in a public transport network. Ben-Akiva & Lerman (1985) give a comprehensive introduction into the theory of choice models and de Dios Ortuzar & Willumsen (2011) give an overview

<table>
<thead>
<tr>
<th>Table 1.1</th>
<th>Selection of timetabling papers, categorized by (1) whether a predetermined route choice is assumed or a route choice model is integrated and (2) whether it is assumed that passengers use a single route only or distribute on multiple routes. The mentioned papers are discussed together with other related literature in Section 2.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predetermined route choice</td>
<td>Integrated route choice</td>
</tr>
<tr>
<td>Single route</td>
<td></td>
</tr>
<tr>
<td>Distribution</td>
<td></td>
</tr>
<tr>
<td>Parbo et al. (2014)</td>
<td>Sels (2015)</td>
</tr>
<tr>
<td>this paper</td>
<td></td>
</tr>
</tbody>
</table>

1 Papers with predetermined route choice mostly assume a passenger weight to be given without explicitly mentioning which distribution was used to obtain the weights. The authors know from conference presentations and personal conversations that almost always a shortest path routing is used. If a passenger distribution is derived from a choice model or historic data, this is usually reported. Therefore, we assume that papers with a predetermined route choice applied single route assignment, unless explicitly stated otherwise. This matches with reports of other authors, see for example Siebert & Goerigk (2013), Schmidt & Schöbel (2015) or Schiewe & Schöbel (2020).
of how choice models are used to estimate passenger decisions in public transport modeling. For estimating passenger distributions in public transport applications, the logit model is most commonly applied.

Choice models are based on the concept of assigning a utility to each route: The percentage of passengers who take a certain route depends on the utility of that route and the utility of alternative routes. The utility of a public transport route is mainly determined by the travel time, number of transfers and time of departure.

Since recently, choice models in general and the logit model in particular are also applied in optimization approaches for public transport applications. Canca, De-Los-Santos, Laporte, & Mesa (2019) used it to estimate a passenger distribution and mode choice in the context of transit network planning. They solve the resulting non-linear program with a neighborhood search based matheuristic. Due to the non-linear structure, exact solution approaches rely mostly on a linearization of the logit model. De-Los-Santos, Laporte, Mesa, & Perea (2017) developed a linear approximation by using that one alternative with fixed utility is available. An overview of common linearizations of the logit model is given by Haase & Müller (2014).

An alternative approach to linear approximation of the choice models is to simulate them (Train, 2009). The utility of each alternative is then represented by the sum of a deterministic part and an error term. The deterministic part quantifies the known part of the utility that can be derived from the characteristics such as travel time or number of transfers. The error term is randomly drawn and models different sources of uncertainty and imperfect knowledge of analyses, such as unobserved route attributes, unobserved passenger preferences or measurement errors (Ben-Akiva & Lerman, 1985). Facho, Sharif Azadeh, & Bierlaire (2016) described such a simulation framework to compute optimal pricing strategies for different parking options while considering passenger behavior.

2.2. Timetabling

2.2.1. Timetabling models and methods

There are different models for public transport timetabling (Cacchiani & Toth, 2012; Galli & Stiller, 2018). For periodic timetabling, the most common model is the periodic event scheduling problem (PESP) as introduced by Serafini & Ukovich (1989). For the PESP model, there are two commonly used integer programming formulations. The first formulation is a direct translation of the modulo constraints in the PESP model and usually called PESP as well. Another formulation is the cyclic periodicity formulation (CPF), which is a further development of the PESP model by Nachtigall (1994). While the PESP has one variable for each event modeling points in time, the CPF uses one variable for each activity expressing a time duration. Compared to the PESP formulation, the CPF features a lower number of integer variables and tighter constraints, resulting in stronger linear programming relaxations and making the CPF easier to solve (Cacchiani & Toth, 2012). Peeters (2003) discussed and compared different cycle bases to improve the computational performance when solving the CPF. Liebschen & Möhring (2002) used the CPF for designing several timetables for the Berlin Underground, which lead to the first optimized railway timetable in practice (Liebschen, 2008).

For solving the PESP in timetabling, multiple solution methods have been developed. Serafini & Ukovich (1989) showed that the problem of finding a periodic timetable is NP-complete. Many publications focus on finding efficient ways to solve periodic timetabling. Schrijver & Steenbeek (1994) developed a constraint propagation algorithm which later on served as a basis for one of the first successful implementations of a timetable found with methods of Operations Research (Kroon et al., 2009). A powerful heuristic to solve the PESP model is the modulo network simplex algorithm developed by Nachtigall & Opitz (2008). The algorithm is inspired by the simplex algorithm for solving linear programs where a feasible solution is improved in each iteration by exchanging a basis and a non-basis variable. After translation of the PESP to propositional logic, Kümmeling, Großmann, Nachtigall, Opitz, & Weiß (2015) could apply efficient SAT solvers to solve the PESP. Pätzold & Schöbel (2016) proposed a promising matching-based heuristic that could find timetables in short computation times. Their algorithm was designed for a reduced PESP model with fixed drive and dwell times for vehicles. Liebschen (2018) described how to exploit the specific structure of a PESP instance to derive effective preprocessing techniques that reduce the complexity of the timetabling problem.

An overview of models and solution methods for railway timetabling is given in Cacchiani & Toth (2012) and Galli & Stiller (2018) and an empirical comparison of multiple solution approaches for the periodic timetable problem is provided in Liebschen, Proksch, & Wagner (2008).

2.2.2. Objective functions in timetabling

Initially introduced as a feasibility program, the PESP model was quickly extended by objective functions to guide the optimization. Recent publications often aim at designing timetables with minimal passenger travel time or with the lowest energy consumption during operation. We refer to Scheepmaker, Goverde, & Kroon (2017) for a summary of energy-efficient timetabling approaches and focus on passenger travel time in this overview. To model the objective of passenger travel time, two restrictive assumptions are often made about passenger behavior. These assumptions have been shown to distort the search for an optimal solution.

First, passengers are in some models assigned to routes in the transport network before the timetable optimization. With this passenger assignment to routes, the arcs in the network are assigned weights to consider passenger routes during optimization in a heuristic way. Many publications have challenged this assumption and shown that the routes passengers use depend on the timetable (Schmidt, 2014) and, therefore, cannot be reliably determined beforehand. To take passengers’ reactions on the designed public transport into consideration, Nachtigall (1998) and Siebert & Goerigk (2013) experimented with iterative approaches. They alternately assigned passengers to shortest routes and optimized the timetable given the updated passenger routes. Schmidt & Schöbel (2015) integrated a shortest-route search for passengers into the timetabling optimization model and further improved the quality of timetables found. They used the fact that the exact route of passengers does not need to be known in the periodic case since start and end events contain sufficient information for travel time computation. With this trick, the resulting timetabling model with integrated passenger assignment to shortest routes could be solved efficiently. Borndörfer, Hoppmann, & Karbstein (2017) developed a general timetable optimization model that allows the implementation of different passenger routing models. They discussed theoretical bounds for four passenger routing models: a lower-bound routing model where passenger routes are found before knowing the timetable; a shortest path routing model where passengers use the shortest path depending on the timetable; a capacitated multi-path routing model where passengers distribute on several paths to avoid violation of capacity constraints; and a capacitated unsplittable path routing model, where all passengers between the same origin and destination travel on the same path while respecting capacity constraints. Their results include the finding that, for different objectives, the travel time on a timetable optimized with predetermined passenger routes can be arbitrarily higher than the travel time on a timetable optimized with integrated passenger routing. Next to theoretical gaps,
Borndörfer et al. (2017) also compared the lower-bound routing model with the shortest path routing model in experiments and found significantly improved transfer waiting times for passengers by integration of the passenger routing model. A different solution approach to periodic timetabling with integrated shortest-route search was described in Gattermann, Großmann, Nachtigall, & Schöbel (2016). They used time slices to model departure time preferences and defined a translation of the integrated model to a satisfiability problem. Schiewe & Schöbel (2020) provided a heuristic approach for the timetabling problem with an integrated shortest-route search that considers only a small share of the OD pairs for timetable-dependent routes. Depending on whether the remaining OD pairs are assigned to fixed routes or not, upper or lower bounds for the exact solution can be found. Together with a preprocessing procedure that reduces the problem size by eliminating unnecessary routing variables, they are able to find improved solutions for close-to-real-world instances. Recently, Löbel, Lindner, & Borndörfer (2019) proposed an adjustment of the modulo simplex algorithm to incorporate a shortest-route search during optimization. Assuming that passengers always take the next available train in a high-frequency network, Polinder, Schmidt, & Huisman (2019) and Polinder, Schmidt, & Huisman (2020) integrated a route selection of passengers in a PESP model. A different approach was followed in an integrated timetabling and stop planning model by Hao, Meng, Corman, Long, & Jiang (2019). While in this paper passenger assignment to routes is not predetermined, the assigned routes do not necessarily correspond to shortest routes or to a distribution (aligned with a choice model).

Second, for the design of a majority of timetable objective functions, it is assumed that passengers only travel on the shortest route. van der Hurk, Kroon, Maröti, & Verest (2014) concluded from their study based on smart card travel data that this is one of the common misassumptions on passenger behavior. Many publications challenged this assumption and proposed enhanced models to develop better timetables for passengers. As input to their timetabling model, Sels (2015) described a passenger assignment to routes that are at most 20% longer than the potentially shortest route. Parbo, Nielsen, & Prato (2014) used estimates for utilities of available connections for deriving passenger distributions and updated the passenger distribution after each timetable computation. A similar approach was used by Robenek, Maknoon, Azadeh, Chen, & Bierlaire (2016) who used estimates for connection utilities as defined for choice models together with time-dependent demand structures to estimate the distribution of passengers. In an extension in Robenek, Azadeh, Maknoon, de Lapper, & Bierlaire (2018), the passenger route choice was estimated with a logit model within the timetabling model. To solve the corresponding timetabling problem with integrated passenger distribution, a simulated annealing heuristic was described that iteratively adjusts the timetable or passenger fares and the passenger distribution. In case of overcrowded trains, their algorithm had the flexibility to selectively remove alternative routes from the passenger choice set with the goal to maximize operator profit.

As mentioned in the literature review on passenger choice models in Section 2.1, some choice models were already integrated into optimization approaches of other public transport applications. For the case of timetabling, Cordone & Redaelli (2011) included a logit model to estimate the passenger mode choice between the modes car, bus, and train. The aim is to find a train timetable that maximizes the total number of passengers. Although the number of passengers is variable, they predetermined a shortest route between each station pair for the passengers. To the best of our knowledge, for passenger route choice no other choice models than a shortest-route search were integrated into an optimization framework for timetabling, which we do in this paper.

2.3. Timetable evaluation

As discussed in Section 2.2, the majority of publications in Operations Research use the absolute travel time of passengers on predetermined routes as objective. This evaluation function is suitable for optimization because of its simple structure. In other research areas, timetables are usually evaluated differently. For evaluation purposes in Transport Engineering, the perceived travel time is often used. That is a weighted travel time equivalent that incorporates more factors of influence, such as penalties for transfers, fares, or adaption time (de Dios Ortuzar & Willumsen, 2011). In contrast to that, commonly applied choice models use an evaluated utility to measure the quality of a timetable for passengers.

An evaluated utility is usually a non-linear function of a weighted travel time equivalent, such as the perceived travel time. Recently, evaluated utilities are often proposed as a replacement for established evaluation functions. For example, de Jong, Daly, Pieters, & van der Hoorn (2007) summarized the literature on ‘logsums’, an evaluated utility, and showcased the advantages of this evaluation in a case study on high-speed trains in the Netherlands. Indeed, Hartleb et al. (2019) showed that timetable evaluation functions do not yield consistent evaluation results, although they are all designed to evaluate the quality of timetables for passengers. This suggests that timetables should be evaluated from different perspectives.

3. Problem definition

In this section, we define the problem of timetabling with an integrated passenger distribution on multiple routes. To this end, we give a basic formulation for both problems: timetabling assuming that a passenger assignment to routes is given, and route choice modeling assuming that a timetable is known.

All formulations are based on an event network $\mathcal{N} = (\mathcal{E}, A)$ with a set of events $\mathcal{E}$ and a set of activities $A$. In this context, an event $i \in \mathcal{E}$ denotes an arrival or a departure of a vehicle at a station, and an activity $ij \in A$ represents a drive or a wait activity of a vehicle between two events $i \in \mathcal{E}$ and $j \in \mathcal{E}$. Activities can be used to model more than vehicle actions, for example, transfer activities of passengers and headway or synchronization constraints between vehicles (Liebchen & Möhring, 2007).

3.1. Passenger distribution

Discrete choice models can be used to describe passengers’ behavior concerning route choice when a timetable is known. We use the logit model to estimate the distribution of passengers on their routes. The passenger routes in the public transport network are represented by paths in the event activity network. A path $p = (i_1, \ldots, i_m)$ is a sequence of events $i$ in the event activity network such that two consecutive events are connected by a drive, wait, or transfer activity. We denote the perceived travel time of path $p$ by $\bar{t}_p$, which is a weighted linear combination of the influencing factors such as travel time and the number of transfers. The perceived travel time is often interpreted as a negative utility of the path $p$ and assumed to be given in the context of choice modeling.

Let a set $P$ of alternative paths with perceived travel times $\bar{t}_p$ for all paths $p \in P$ be given. Then, the logit model can be interpreted as a probability function $w^\text{ln}_p((\bar{t}_p)_{q \in P})$ that assigns a probability to alternative $p$, based on the utility of all considered alternatives;

$$w^\text{ln}_p((\bar{t}_p)_{q \in P}) = \frac{e^{\bar{t}_p}}{\sum_{q \in P} e^{\bar{t}_q}}.$$  

(3.1)
where \((\bar{t}_p)_{p\in P}\) is a vector containing the utilities of all paths in the set \(P\). With the scalar \(\beta \in \mathbb{R}\), the logit model can be adjusted to suit the specific instance.

### 3.2. Timetabling

In the literature, an instance \(T = (\mathcal{N}, l, u, OD)\) for a timetabling problem usually consists of an event activity network \(\mathcal{N}\) with lower and upper bounds \(l\) and \(u\) on the activity durations, and a demand matrix \(OD\) that indicates how many passengers wish to travel from each origin to destination. It remains to find arrival and departure times for each vehicle on each line. As mentioned in Section 2.2.1, the two most commonly used formulations for timetabling are the periodic event scheduling problem (PESP) formulation and the cyclic periodicity formulation (CPF). The main difference between these formulations is that the PESP has one variable per node in the event activity network, denoting the point in time of an event, and the CPF has one variable per arc in the network, denoting the duration of an activity. Since our goal is to compute the lengths of passenger paths in the event activity network as sum of activity durations, we focus on the cyclic periodicity formulation for periodic timetabling problems. This integer linear formulation is based on a set of constraints ensuring that the duration \(\Delta_{ij} \in \mathbb{Z}_+\) of each activity \(ij \in A\) is between a given lower \(l_{ij} \leq \delta_{ij} \leq u_{ij}\) and upper bound \(u_{ij} \in \mathbb{Z}_+,\) i.e.,

\[l_{ij} \leq \delta_{ij} \leq u_{ij} \quad \forall j \in A. \tag{3.2}\]

We assume that the timetable has an accuracy of one time unit and the duration between events is integer valued. A solution to this formulation specifies how much time elapses between two events, but not when the events occur. Since the events are to be repeated periodically, the activity durations cannot be determined independently. To ensure that the durations can be transformed into a feasible periodic timetable that assigns a point in time to each event, cycle constraints need to be added to the CPF (Nachtigall, 1994). Each cycle constraint assures that the sum of durations on a cycle in the event activity network is a multiple of the period \(T\), which implies the periodicity of events. It is sufficient to include cycle constraints for each cycle \(c\) in an integral cycle basis \(\mathcal{C}\) of the event activity network (Liebchen & Peeters, 2009). We add

\[\Gamma_c \delta = T \cdot \mu_c \quad \forall c \in \mathcal{C}. \tag{3.3}\]

to the constraints, using an integer cycle variable \(\mu_c \in \mathbb{Z}\). The vector \(\Gamma_c\) indicates the forward and backward edges in cycle \(c\). The objective of most timetabling formulations is to minimize the total travel time of passengers. Mostly, this is achieved with the help of passenger weights \(\bar{x}_{ij}\) on each activity \(ij\) and by minimizing

\[\sum_{i,j,A} \bar{x}_{ij} \cdot \delta_{ij}. \tag{3.4}\]

Note that the passenger weights \(\bar{x}_{ij}\) are predetermined by assigning passengers to routes before optimization. The cyclic periodicity formulation for timetabling with predetermined passenger routes uses Constraints (3.2) and (3.3) and is given by

\[
\min \sum_{i,j,A} \bar{x}_{ij} \cdot \delta_{ij} \quad \text{s.t.} \quad \begin{align*}
\delta_{ij} &\geq l_{ij} \quad \forall i,j \in A \\
\delta_{ij} &\leq u_{ij} \quad \forall i,j \in A \\
\Gamma_c \delta &= T \cdot \mu_c \quad \forall c \in \mathcal{C} \\
\delta_{ij} &\in \mathbb{Z}_+ \quad \forall i \in \mathcal{I} \\
\mu_c &\in \mathbb{Z}_+ \quad \forall c \in \mathcal{C}
\end{align*}
\]

### 3.3. Integration of passenger distribution and timetabling

Section 3.1 defines the logit model to estimate passengers’ route choice for a given timetable, and Section 3.2 provides a standard model to optimize a timetable for a predetermined passenger route choice. Since the result of one model is the input for the other and vice versa, we aim at developing a model integrating both aspects.

We assume that for each OD pair \(k\) a finite choice set \(P_k\) of \(n_k\) possible paths is given. Each path \(p \in P_k\) is a sequence of events in the event activity network that could be taken by the passengers of OD pair \(k\). Since these paths are defined in the event activity network, two paths for one OD pair can be different although they might use the same stations and tracks. In fact, such two paths do not need to have any event in common.

The logit model in Section 3.1 and the timetabling model in Section 3.2 assumed the received travel time \(t_p\) on paths and the passenger weights \(x_{ij}\) on activities to be predetermined, respectively. Both are variables in the integrated model, which we indicate by omitting the bars above the identifiers. The aim is to derive the passenger weight \(x_{ij}\) on each activity \(ij\) from the distribution on the paths. To this end, we compute the respective length

\[t_p = \sum_{i,j,p} \delta_{ij} \quad \forall p \in P_c, \quad \forall k \in OD \tag{3.5}\]

of each path for all OD pairs as the sum of durations of the activities. Note that the definition of \(t_p\) can easily be extended by additional external influencing factors such as a fare for taking path \(p\) or a penalty for each transfer included in path \(p\). Since the path choice sets for OD pairs are assumed to be given, fares or transfer penalties can be determined in a preprocessing step for each path and are constant in the model formulation. As these constants added to \(t_p\) do not affect the structure of the model, they are omitted in the problem formulation for ease of exposition. Given the path lengths \(t_p\), we can use the logit distribution \(w_p^m\) to compute a share of each OD pair using the path \(p\). Multiplied by the number of passengers \(o_{ij}\) of OD pair \(k\), this yields the number of passengers on each activity \(ij\) using path \(p\), which we denote by

\[x_{ij}^p = w_p^m \cdot (t_{ij} x_{ij}) \cdot q_k \quad \forall i,j \in p, \quad \forall p \in P_k, \quad \forall k \in OD. \tag{3.6}\]

This is an expected value and not necessarily integral. By aggregating these numbers over all paths \(p\) for each OD pair, we obtain the number of passengers on each activity

\[x_{ij} = \sum_{k=0}^{OD} x_{ij}^p \quad \forall i,j \in A. \tag{3.7}\]

As in the timetabling formulation from Section 3.2, this number is used in the objective function to find a travel time minimal timetable. We formulate a general optimization problem for timetabling assuming that a passenger distribution can be modeled with a logit model:

\[
\min \sum_{i,j\in A} x_{ij} \cdot \delta_{ij} \quad \text{s.t.} \quad \begin{align*}
\delta_{ij} &\geq l_{ij} \quad \forall i,j \in A \\
\delta_{ij} &\leq u_{ij} \quad \forall i,j \in A \\
\Gamma_c \delta &= T \cdot \mu_c \quad \forall c \in \mathcal{C} \\
\delta_{ij} &\in \mathbb{Z}_+ \quad \forall i \in \mathcal{I} \\
\mu_c &\in \mathbb{Z}_+ \quad \forall c \in \mathcal{C}
\end{align*}
\]

Note that the variables \(\delta\) and \(t\) can be relaxed to be continuous since the lower \(l\) and upper bounds \(u\) are integer and \(\mathcal{C}\) is an integral cycle basis. No matter which domain is chosen, this formulation cannot be solved efficiently due to the passenger distribution function \(w_p^m\). Furthermore, the objective is non-linear in the variables since the passenger loads \(x\) are modeled to be dependent on the durations \(\delta\).
4. Models

Already Parbo et al. (2014) argued that the problem from Section 3.3 is “extremely difficult to solve mathematically, since the timetable optimisation is a non-linear non-convex mixed integer problem, with passenger flows defined by the route choice model, where the route choice model is a non-linear non-continuous mapping of the timetable.” In this section, we describe two different representations of the route choice model. Using these, we introduce two linear formulations for the problem of finding travel-time minimal routes under the assumption that passengers’ routes choice can be modeled using a logit model.

4.1. Model 1 - timetabling with linear distribution model

In a first model, we use a novel linear passenger distribution model that is inspired by characteristics of the logit model. Furthermore, the quadratic objective is linearized. We address these two details in the following two sections and provide a linear formulation for timetabling with an integrated passenger distribution model.

4.1.1. Linear distribution model

The literature provides multiple linearizations of the logit model for applications in Operation Research. To our best knowledge, these linearizations can be classified into two cases. Either, just the utility of a single alternative is variable while the utilities of all remaining alternatives are fixed. Or, the utilities of all alternatives for customers are fixed and the decision is whether to offer alternatives. Since in our case all alternative paths are always available and their utility depends on the timetable, these linearizations are not appropriate.

Therefore, we develop a linear distribution model to approximate the logit model. Our model allows all utilities to be flexible in their domain, i.e., $t_p \in [\underline{m}_k, \overline{m}_k] \forall p \in P_k$, and satisfies the probability characteristics. For each OD pair $k \in OD$, we require the following five characteristics

**Distribution characteristics**

$$w_p((t_q)_{q \in R_k}) \in [0, 1] \text{ and } \sum_{p \in P_k} w_p((t_q)_{q \in R_k}) = 1 \quad (4.1)$$

**Monotonicity** Let $[R_k] > 1$, let $\varepsilon > 0$ and let $e_p$ be the unit vector with a 1 at the position of path $p$. Then

$$w_p((t_q)_{q \in R_k} + \varepsilon \cdot e_p) < w_p((t_q)_{q \in R_k}) \quad (4.2)$$

**Uniform distribution on equivalent alternatives**

$$w_p((t_q)_{q \in R_k}) = \frac{1}{|R_k|} \quad (4.3)$$

if all paths have the same length, that is, $t_p = t_q \forall q \in P_k$.

**Independence of order** Let $\pi : P_k \rightarrow P_k$ be any permutation on a set of paths $P_k$ that keeps path $p$ constant, i.e., $\pi(p) = p$. Then

$$w_p((t_q)_{q \in R_k}) = w_p((t_{\pi(q)})_{q \in R_k}) \quad (4.4)$$

**Logit characteristic: absolute differences in utility determine probability**

$$w_p((t_q + \delta)_{q \in R_k}) = w_p((t_q)_{q \in R_k}) \cdot \alpha \quad (4.5)$$

where $\delta \in \mathbb{R}^n$ is a constant.

This yields a family of linear distribution functions.

**Lemma 4.1.** Let $n_k = |P_k|$ be the number of alternative paths and let $\underline{m}_k$ and $\overline{m}_k$ be the minimal and maximal possible length of any considered path in the event activity network for OD pair $k$, respectively. Then all linear distribution functions fulfilling the five characteristics (4.1) to (4.5) can be characterized according to the three following cases:

1. $n_k = 1$
   - If there is just one path $p$ for OD pair $k$ given, then $P_k = \{p\}$ and $w_p((t_p)) = 1$.
   - (4.6)

2. $n_k \neq 1$ and $\underline{m}_k = \overline{m}_k$
   - If $\underline{m}_k = \overline{m}_k$, all paths have the same fixed length, i.e., $t_q p \forall p \in P_k$. Then
   $$w_p((t_q)_{q \in R_k}) = w_p((t_q, \ldots, t_q)) = \frac{1}{n_k} \quad (4.7)$$

3. $n_k \neq 1$ and $\underline{m}_k \neq \overline{m}_k$
   - In the general case all linear functions with the required characteristics have the form
   $$w_p((t_q)_{q \in R_k}) = \frac{\alpha}{n_k (\overline{m}_k - \underline{m}_k)} \left( \frac{t_p - \frac{1}{n_k - 1} \sum_{q \neq p} t_q}{\frac{1}{n_k} - \frac{1}{n_k}} \right) + \frac{1}{n_k} \quad (4.8)$$
   with $\alpha \in (0, 1]$.

A constructive proof for Lemma 4.1 is given in Appendix B. We replace the logit model by the linear distribution functions (4.6)–(4.8) in their respective cases in the model from Section 3.3. This yields a linearly constrained feasible region of the optimization problem and further ensures that the five characteristics (4.1) to (4.5) hold.

The linear distribution function is defined in the range $[\underline{m}_k, \overline{m}_k]$ for the length $t_q$ of each path $q$. The slope of the function in that domain can be adjusted with the parameter $\alpha \in [0, 1]$. For example, for $\alpha \rightarrow 0$, we approximate the uniform distribution, independent of the path lengths. The higher $\alpha$, the more do passengers react on differences in path lengths. In experiments, we learned that the linear distribution model from Lemma 4.1 tends to distribute passengers more evenly on paths than a logit distribution. Therefore, we use a value of $\alpha = 1$ to scale the linear distribution model in all experiments.

Fig. 4.1 visualizes the probabilities that path $p$ is chosen according to a logit and a linear distribution model, given a second path $q$ with fixed length $t_q$. To better demonstrate the linear distribution model, three cases for the fixed path length $t_q$ are considered. For example, in Fig. 4.1a it is assumed that the length of the alternative path $q$ is as short as possible, i.e., $t_q = \underline{m}_k$. Then, the probability that path $p$ is chosen is at most 0.5 since it cannot be shorter than path $q$. The higher the length of path $q$, the higher the probability that path $p$ is chosen, see Fig. 4.1b and c.

This figure also illustrates how the probability of the logit distribution can be over- or underestimated by the linear distribution model. Knowing the length $t_q$ of the alternative path $q$, a better linear approximation of the logit model is possible. However, since the utilities of all alternatives depend on the timetable, a linear distribution model can only rely on the bounds $\underline{m}_k$ and $\overline{m}_k$.

4.1.2. Reformulation of the model

Using the linear distribution model from Lemma 4.1 instead of the logit distribution allows us to express the number of passengers on each activity $x_{ij}$ as a linear function of the durations $\delta_{ij}$. We obtain the quadratic integer program for timetabling with

**Integrated passenger distribution according to a LINEar distribution model (ID-LIN)name tabling with linearized logit:**

$$\min \delta^T A \delta + b^T \delta$$

s.t. $\delta_{ij} \geq l_{ij} \forall ij \in A$

$\delta_{ij} \leq u_{ij} \forall ij \in A$

$\Gamma_c^{ij} \delta = T \cdot c_{ij} \forall c \in C$

$\delta_{ij} \in \mathbb{Z}^+ \forall ij \in A$

$\mu_c \in \mathbb{Z}^+ \forall c \in C$
where $\delta^l$ and $b^l$ denote the transpose of the column vectors $\delta$ and $b$, respectively. The coefficients in the objective function are defined as

$$A_{ij,j'} := \sum_{k \in OD} \frac{\alpha \cdot q_k}{n_k (m_k - \bar{m}_k)} \left( \sum_{p \in P} \left[ \sum_{q \in OD} \left( \sum_{k \in OD} \sum_{r \in R} \frac{1}{n_k - 1} \right) \right) \right) + \delta_{ij,j'} \in A$$

and

$$b_{ij} := \sum_{k \in OD} \frac{\alpha_k}{n_k} \forall j \in A.$$ 

In Eq. (4.9), $OD^*$ denotes the set of OD pairs $k$ with $n_k > 1$ and $m_k \neq \bar{m}_k$. This means, that only OD pairs with multiple paths contribute to the matrix $A$, and thus add to the quadratic part of the objective function. OD pairs with only one path, or with multiple paths of fixed length, only add to the linear part of the objective. The derivation of the coefficient matrix $A$ and vector $b$ can be found in Appendix C.

(ID-LIN) is a minimization program and the coefficient matrix $A$ can be proven to be negative semi-definite, see Appendix D. That means the objective function is concave and standard methods for quadratic programs are not expected to be efficient. Therefore apply a linearization to the objective function. To this end, we express the integer variables $\delta_{ij}$ as a sum of binary variables,

$$\delta_{ij} = l_{ij} + \sum_{m=0}^{[\log(l_{ij}-l_{ij})]} 2^m \sigma_{ij}^m$$

and linearize the products of binaries $\sigma_{ij}^m \cdot \sigma_{ij}^{m'}$. The corresponding linearization of the optimization program (ID-LIN) as used for the experiments can be found in Appendix E.

### 4.2. Model 2 - simulation of logit model

In a second model, we integrate a simulated passenger distribution into the timetabling function. The simulation is based on an alternative way to compute the logit probabilities. According to Train (2009), it holds that

$$W_p(t_q | q) = \frac{e^\beta_q}{\sum_{q \in OD} e^\beta_q}$$

$$= Prob\left( t_p + \varepsilon_p \leq \min_{q \in OD} \left( t_q + \varepsilon_q \right) \right).$$

where the $\varepsilon_p$ are independent and identically Gumbel distributed. That means the logit probability that alternative $p$ is chosen equals the probability that the length of path $p$, deferred by some random value $\varepsilon_p$, is shorter than the length of any alternative path $q$, deferred by some random value $\varepsilon_q$. Following similar steps as Pacheco et al. (2016), we use the representation in Eq. (4.10) to simulate the logit model by drawing random values for $\varepsilon$. That means we consider several scenarios $r \in R$, draw a random value $\varepsilon_{pr}$ for each path $r$ in each scenario, and add these to the path lengths. This yields a different, randomized path length in each scenario, which we denote by

$$t_{pr} = \sum_{ij \in OD} \delta_{ij} + \varepsilon_{pr} \forall k \in OD, \forall p \in P_k, \forall r \in R.$$ 

Note that similar to the path length computation in Eq. (3.5), this modeling can easily be extended by additional factors like fares or a penalty for each transfer as well. Then, we choose the shortest path in each scenario for each OD pair and denote the travel time for OD pair $k$ in scenario $r$ by

$$t_{kr} = \min_{\{p \in P_k\}} t_{pr} \forall k \in OD, \forall r \in R.$$ 

This discrete choice of the shortest path in each scenario $r$ yields a distribution of the passengers of OD pair $k$ over the available paths in the path choice set $P_k$. Since we choose the random terms $\varepsilon_{pr}$ to be independent and identically Gumbel distributed, this distribution converges towards a logit distribution for an increasing number of scenarios, see Eq. (4.10).

Using a binary choice variable $z_{pr}$ that is set to one if and only if path $p$ is the shortest in scenario $r$, constraint (4.11) can be linearized to

$$t_{kr} \leq t_{pr} \forall k \in OD, \forall p \in P_k, \forall r \in R$$

$$t_{kr} \geq t_{pr} - (1 - z_{pr}) M_{kr} \forall k \in OD, \forall p \in P_k, \forall r \in R$$

$$\sum_{p \in P_k} z_{pr} = 1 \forall k \in OD, \forall r \in R$$

where

$$M_{kr} = \max_{p \in P_k} \left( \sum_{q \in OD} \left( t_{qr} + \varepsilon_q \right) \right) - \min_{p \in P_k} \left( \sum_{q \in OD} \left( t_{qr} + \varepsilon_q \right) \right)$$

is sufficiently large.

Note that in the case of equality of best randomized path lengths in a scenario, this modeling will do a random assignment of the passenger choice. We obtain the model for timetabling with an Integrated passenger Distribution by SIMulation of the logit
model (ID-SIM) name timetabling with simulated logit:

\[
\begin{align*}
\min & \sum_{k \in \text{OD}} \frac{1}{|R|} \sum_{r \in R} t_{yr} \\
\text{s.t.} & \delta_{ij} \geq t_{ij} \quad \forall i \in A \\
& \delta_{ij} \leq u_{ij} \quad \forall i \in A \\
& T_{ij} \mu_c = \sum_{k \in \text{OD}} \delta_{ij} + \epsilon_{pr} \\
& \sum_{pr} \epsilon_{pr} \sum_{ij} \epsilon_{ij} \geq 1 \\
& t_{yr} \geq \delta_{ij} + z_{pr} (1 - \epsilon_{pr}) |
\end{align*}
\]

The constraints and the objective function of this formulation are linear in the variables.

Obviously, there is a trade-off between the solvability of the MILP model (ID-SIM) and the accuracy of the simulation. Considering only a few scenarios results in a small model which, yields a random solution because a path could be privileged or disadvan-
taged by chance. With an increasing number of scenarios, we ex-
pect the passenger distribution on paths to converge and the solu-
tion to stabilize, but also the model size and hence solution time to increase. To choose a setting that balances solvability and accuracy, we ran preliminary experiments with varying numbers of scenarios. Based on this, we choose to use a low number of \(|R| = 10\) sce-
narios and pick the best solution of 10 repetitions instead of using a large number of scenarios. In our experiments, this has shown to yield a good trade-off between computation time and a high probability to find a solution of high quality. Another advantage of solving each instance multiple times with a small number of sce-
narios over considering large scenario sets is that the repetitions are independent, and therefore easily parallelizable.

4.3. Illustration of model differences

In this section, we compare the two models (ID-LIN) and (ID-
SIM) concerning their objective functions. The objective function of (ID-LIN) is the sum of the absolute travel times of all passen-
gers on their respective paths, which are chosen based on the lin-
ear distribution model introduced in Lemma 4.1. This distribution assumption implies that not everyone travels on a shortest path, but passengers make use of paths with slightly longer travel times than the shortest as well. The following minimal example shows that with the use of (ID-LIN) seemingly unintuitive results can be obtained: a timetable that provides better or equal travel time on all routes, is not necessarily evaluated as better by the objective function of (ID-LIN).

Example 1. Consider a network consisting of two stations \(A\) and \(B\) and one OD pair \(k\) that wants to travel from \(A\) to \(B\). Assume, there are two available paths, \(p\) and \(q\), with respective bounds \([10, 22]\) and \([11, 21]\). The example network is illustrated in Fig. 4.2.

We compare three different timetables \(t^1\), \(t^2\) and \(t^3\). The first timetable offers two equally good paths, these are \(t_{yq}^3 = 11 = t_{yq}^2\). The second and third timetable have one short path and one longer alternative, respectively. These are \(t_{yq}^2 = 10\), \(t_{yq}^3 = 13\) and \(t_{yq}^3 = 10\), \(t_{yq}^3 = 21\), respectively.

We evaluate the three timetables with (1) the travel time on the passengers’ respective shortest path, (2) the travel time when assuming that passengers distribute according to a logit distri-
bution, and (3) the travel time when assuming that pas-
sengers distribute according to the linear distribution model from Lemma 4.1 which is the objective function of (ID-LIN). For the logit and linear distribution models, we use the parameters \(\beta = -0.22\) and \(\alpha = 1.0\), respectively. The objective values of one passenger of OD pair \(k\) can be found in Table 4.1.

We find, as expected, that the travel time on the shortest path is best in timetables \(t^2\) or \(t^3\), regardless of the length of alternative \(q\). Regarding travel time according to a linear or logit distri-
bution, timetable \(t^2\) is worse than timetable \(t^1\). This result is open for discussion as none of the two timetables is obviously better than the other. However, according to the linear or logit distribution, timetable \(t^1\) performs better than timetable \(t^2\). This is an interesting observation since \(t^2\) is intuitively perceived as the better timetable. In \(t^2\), path \(p\) has the same length as in \(t^3\), and path \(q\) is shorter than in \(t^2\). Thus, this evaluation result might be unexpected at first glance, but it has a simple explanation. The worse the travel time \(t_{yq}^3\) of alternative \(q\), the more probable it is that passengers choose to travel via path \(p\), which yields a lower total travel time.

This example shows that combining the distribution of passen-
gers on multiple paths with the evaluation function of total travel time can lead to evaluation results that contradict our intuition. Timetables that are better or equal in every aspect are evaluated as the worse timetables. This implies that using the total travel time as an objective function and distributing passengers on multi-
ple paths as done in model (ID-LIN) can incite to design timetables with longer alternatives than necessary.

This issue is approached with the objective function of model (ID-SIM). There, the objective is to minimize the weighted sum of randomized shortest path lengths \(t_{yw}\) instead of the ab-
solute travel time as used in the first model (ID-LIN). In the model (ID-SIM), a path only enters the objective function if it is perceived better than any alternative in at least one scenario. Hence, no considered path in the path choice set can deteriorate, but only improve the objective value. That means the objective function of the program (ID-SIM) does not have a bias towards de-
signing longer paths than necessary.

5. Experimental setup

5.1. Instances

To test and compare our approaches, we run experiments on a number of instances. Each instance \(I\) consists of an event activity

<table>
<thead>
<tr>
<th>((t_{yr}, \mu_c))</th>
<th>Shortest path</th>
<th>Logit distribution</th>
<th>Linear distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>((11, 11))</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>((10, 13))</td>
<td>10</td>
<td>11.02</td>
<td>11.13</td>
</tr>
<tr>
<td>((10, 21))</td>
<td>10</td>
<td>10.90</td>
<td>10.46</td>
</tr>
</tbody>
</table>

ARTICLE IN PRESS

network \(N\) with lower and upper bounds \(l\) and \(u\) and a de-
mind situation. The event activity network is derived from infor-
mation about the public transport network, i.e., stations and tracks, as well as a line plan. Both models (ID-LIN) and (ID-SIM) assume a choice set of paths \(P_k\) for each OD pair \(k\) to be given. The paths are defined in the event activity network and can be interpreted as a sequence of line trips. Depending on the line plan, there can be multiple paths on the same geographical route or just a single path although origin and destination are connected by a number of geographical routes. Hence, a track network can give an indica-
tion but is not determining for the number of passenger paths for an OD pair. How we preprocess the instances and derive a path choice set is described in Appendix F.

5.1.1. Instances on grid network

We consider 32 instances defined on a $3 \times 3$ grid network, which is depicted in Fig. 5.1a. On this network, we consider four different demand situations and for each of them several line plans with corresponding event activity networks. The number of events and activities in the corresponding timetabling instances range from 120 to 208 and 326 to 760, respectively. The instances are partial instances of a bigger grid network introduced by Friedrich, Hartl, Schiewe, & Schöbel (2017) and made available in an online repository.2 The grid infrastructure has several geographically different routes of comparable length for passengers. Depending on the line plan, this provides good conditions to find multiple passenger paths in the event activity network. On the 32 instances, there are on average 1.7 paths for each OD pair with a maximum of 8 paths for one OD pair across all instances. On average, 29.9 OD pairs have more than one path.

5.1.2. Instance on Dutch railway network

To test our approaches on a real-world instance, we consider a part of the Dutch railway network operated by Netherlands Railways (NS). The partial network includes the stations Amsterdam, Den Haag, Den Haag HS, Haarlem, Gouda, Leiden, Rotterdam, Rotterdam Alexander, and Utrecht in the Randstad, a metropolitan region in the Netherlands. The track network is depicted in Fig. 5.1b. We consider eight intercity lines operating between the stations, yielding 128 events and 357 activities for the timetabling model. Based on this, 1 to 7 paths are available per OD pair, with an average of 2.4 paths. On the Dutch railway network, 40 OD pairs have multiple available paths. Both the number of OD pairs with multiple paths and the average number of paths per OD pair are higher than in the grid network although this network contains less cycles. This indicates that the optimization problem for the Dutch instance is larger and thus potentially harder to solve.

5.2. Timetabling approaches

We compare the timetabling models with integrated passenger distribution (ID-LIN) and (ID-SIM) with three state-of-the-art methods for timetabling: two methods (PS) and (PD) assume a predetermined passenger assignment to routes, and one method (IS) has an integrated passenger routing on the shortest paths. Besides the timetabling models (ID-SIM) and (ID-LIN) that integrate the passenger distribution, we also test and compare a heuristic solution approach (ID-ITR) for timetabling with passenger distribution. These approaches are described in more detail below.

(P) First, a timetabling model with a Predetermined passenger assignment on a Single path is considered. In this model, the passengers’ routes are fixed before the optimization step. We assign passengers to the shortest route using the average bounds $\frac{1}{2} (l_i + u_i)$ on edges in the event activity network. This basic version of the timetabling model is the subject of many publications since the development of the PESP model, see for example Nachtrig & Opitz (2008) or Liebchen (2018). An integer programming formulation is given by Eqs. (3.2) to (3.4), as described in Section 3.2.

(PD) Second, we implement another model with Predetermined passenger routes. In contrast to the model (PS), passengers are Distributed on multiple paths according to a logit model with the parameter $\beta = -0.22$ and using average bounds on edges. In consultation with traffic engineers, the value of $\beta$ is chosen similar to values that are typically found when fitting the logit model on instances with similar travel distances. We are not aware of a published timetabling approach that explicitly states a predetermined passenger distribution according to a logit model. Still, this strategy can be compared to those made in Parbo et al. (2014) or Robenek et al. (2016), where passenger distributions were derived from utilities of alternative routes. The underlying integer programming model is all the same as the one in (PS), only the passenger weights are predetermined in a different way.

(IS) Third, we consider a timetabling model with an Integrated Shortest path search. The timetable is optimized with the objective of minimizing passenger travel times for passengers that choose the shortest path based on the timetable. This approach resembles the idea of the integrated shortest path models described in Siebert & Goerigk (2013), Gattermann et al. (2016) and Bornsdörfer et al. (2017), for example. An integer programming formulation of this model is attached in Appendix G.1.

(ID-ITR) Fourth, we consider a heuristic approach for timetabling with an Integrated passenger Distribution that iterates between timetable design and passenger distribution. To compute the passenger distribution based on a fixed timetable, we use the logit model with the parameter $\beta = -0.22$. The initial passenger loads are determined by using the average bounds as edge lengths. In all following iterations, the realized edge lengths of the timetable are used. This yields fixed passenger loads on each edge in the event activity network in each iteration and a standard timetabling model assuming a predetermined passenger distribution can be solved with the given loads. We iterate until the solution value does not change significantly between two iterations or a maximum number of iterations is reached. Similar iterative approaches for

\[ \text{https://github.com/FO2083/PublicTransportNetworks/tree/master/Grid_5x5} \]
Table 5.1  
Summary indicating which solution approach (1) assumes a predetermined route choice or has an integrated route choice and (2) assumes that passengers use a single route only or distribute on multiple routes.

<table>
<thead>
<tr>
<th>Predetermined route choice</th>
<th>Integrated route choice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single route (PS)</td>
<td>(IS)</td>
</tr>
<tr>
<td>Distribution (PD)</td>
<td>(ID-ITR), (ID-LIN), (ID-SIM)</td>
</tr>
</tbody>
</table>

timetabling and passenger route choice are described in Sels, Dewilde, Cattrysse, & Vansteenwegen (2011) or Parbo et al. (2014), for example. The pseudocode used for this method can be found in Appendix G.2.

We refer to these benchmark models by (PS), (PD), (IS), and (ID-ITR), respectively. Table 5.1 indicates whether the route choice is integrated into the methods as well as which kind of route choice model is assumed.

By comparing the models (ID-LIN) and (ID-SIM) with the heuristic approach (ID-ITR) and the three benchmark models (PS), (PD) and (IS), we can identify the benefits of integrating (1) passenger route search and (2) simultaneous modeling of a passenger distribution.

5.3. Implementation

To reduce the size of the search space, the domain of the variables $\mu_c$ is constrained in all models with the following inequalities:

$$\left[ \frac{1}{t} \sum_{ij \in c} l_{ij} - \frac{1}{t} \sum_{ij \in c} u_{ij} \right] \leq \mu_c \leq \left[ \frac{1}{t} \sum_{ij \in c} u_{ij} - \frac{1}{t} \sum_{ij \in c} l_{ij} \right] \forall c \in C.$$

Here, $c^+$ and $c^-$ denote the set of edges in cycle $c$ in forward and backward direction, respectively, and $l_{ij}$ and $u_{ij}$ are the lower and upper bounds of activity $ij$. These well-established inequalities were first described in Odkij (1996).

All mixed integer linear programs are solved with the general-purpose solver Fico Xpress 8.5 on a laptop with 32 GB RAM and an Intel® Core™ i7-6700HQ. A time limit of one hour is used for the grid instances and no time limit for the Dutch railway instance. For all experiments, we use a start solution to warm start the optimization. This start solution consists of an initial timetable for the instance and, if applicable, a corresponding passenger routing according to the passenger distribution model of the solution approach used.

5.4. Evaluation of timetables

Different research areas apply different measures to evaluate timetables from the passengers’ perspective. We could see in Example 1 that different evaluation functions can yield different results on small networks. This small example suggests two features. First, although travel time is commonly used to evaluate timetables, it might not be suitable when considering a passenger distribution on multiple routes. Second, different evaluation measures may consider different timetables to be better, although the functions are commonly accepted to serve for the evaluation of timetables. Hartleb et al. (2019) compared multiple timetable evaluation functions for passengers on different instances and indeed found that these functions are often not consistent in their evaluation. We learn that there is no default objective function to be used when optimizing timetables with an integrated passenger distribution. To avoid misinterpretation of the results due to a simplistic or biased evaluation, we evaluate all resulting timetables with four different evaluation functions. As before, we denote the total passenger load of OD pair $k$ with $o_k$ and the length of path $p$ with $l_p$, as defined in Eq. (3.5). Let $\beta_p$ be a set of available paths for OD pair $k$. The used evaluation functions are:

$${\it t}_{tp}$$ The total travel time of all passengers on their shortest path:

$$\sum_{k \in OD} \sum_{p \in \beta_p} w_{kp}^p \cdot t_p,$$

where $w_{kp}^p$ is the probability that passengers choose path $p$ assuming that all passengers use their shortest paths only.

$${\it t}_{mp}$$ The total travel time of all passengers when distributed on multiple paths according to the logit model:

$$\sum_{k \in OD} \sum_{p \in \beta_p} w_{kp}^m \cdot t_p,$$

where $w_{kp}^m$ is the probability that passengers choose path $p$ assuming that all passengers distribute on their paths according to a logit distribution.

$${\it u}_{sum}$$ The evaluated total utility for all passengers, defined as the weighted sum of all logit denominators:

$$\sum_{k \in OD} \sum_{p \in \beta_p} e^{\beta_k},$$

with $\beta = -0.22$. Derived from the logit model, this measure indicates utility of a public transport service for passengers. The utility of a path $p$ is weighted with the passenger load $o_k$.

$${\it u}_{log}$$ The logsums, a utility based evaluation function, defined as the weighted sum of the logarithm of all logit denominators:

$$\sum_{k \in OD} \sum_{p \in \beta_p} \ln \left( \sum_{p \in \beta_p} e^{\beta_k} \right),$$

with $\beta = -0.22$. Similar to the evaluated total utility, the logsums are a measure of utility for passengers. Due to the logarithm in this evaluation function, the effect of changing travel time $t_p$ on the evaluation $u_{log}$ depends not only on the passenger load, but also on the number of paths of that OD pair. For example, changing the travel time on one path of an OD pair with many good paths might affect the value of $u_{log}$ less than changing the travel time of one path of an OD pair with just few and bad paths, even if the passenger load in the first case is higher than in the second case. That means, OD pairs have different weights relative to each other, as contrasted with the evaluated total utility.

All four functions evaluate the quality of timetables from the passengers’ perspective. Note that these functions are commonly used for evaluation, but due to their structure, not all are suitable as objective functions in an optimization program. The first two evaluation functions are travel time based and thus to be minimized while the latter two evaluation functions are utility based and hence to be maximized. Considering all four evaluation functions allows a thorough investigation and comparison of the timetables and, in this way, of the proposed timetabling methods.

For better comparability, we present the relative solution values when compared to an ideal solution. In an ideal solution, it is assumed that the travel time on each path for each OD pair is equal to the length of the path using the lower bounds on all edges. That means, in an ideal solution we set $\delta_{ij} := l_{ij} \quad \forall ij \in A$

without considering the cycle constraints (3.3). This is also called lower-bound routing of passengers, see Borndörfer et al. (2017). For most instances, such an ideal solution does not exist, but it is a common measure to see how close solutions are to perfect conditions. More details about ideal solutions and about how they are used in practice can be found in Caimi, Kroon, & Liebchen (2017).
6. Results

In the experiments, we showcase the benefits and drawbacks of the timetabling models with integrated passenger distribution (ID-LIN) and (ID-SIM) when compared to existing timetabling approaches.

6.1. Experiments on 32 instances on the grid network

We conduct experiments on 32 instances on the grid network described in Section 5.1.1. On seven instances, all six methods find an ideal solution, and on another four instances, the model (ID-LIN) could not find an optimal solution or could not prove optimality in tests with a time limit of ten hours. Therefore, we exclude these 11 instances from the discussion. In Fig. 6.1, we present the evaluation values of the solutions found by the different approaches averaged over the remaining 21 instances on the grid network. This figure shows the average performance of the six methods on the four considered evaluation functions introduced in Section 5.4. All values are given in percent, relative to the evaluation value of an ideal solution.

The relative evaluation values can be read as follows. For example, a relative value of 1.77 for $t_{sp}$ in Fig. 6.1a of the model (PS) means that the travel time on the shortest connection in the solution of (PS) is, on average, 1.77 percent longer than the travel time on the shortest connection in an ideal solution. Comparing this to the relative travel time on a shortest connection of the model (IS), 0.57, shows that (IS) performs, on average, better than (PS) regarding the travel time on the shortest path. In general, the relative evaluation values show to what extent a solution is worse than an ideal solution, according to the used evaluation function. We discuss the results per evaluation function.

Figure 6.1a When evaluating timetables with travel time on the shortest path $t_{sp}$, on average, the methods (IS) and (ID-SIM) provide the best solutions. This is expected for the method (IS) since its objective is to minimize the total travel time of passengers on their shortest paths. To simulate a logit distribution in the model (ID-SIM), in each scenario the shortest path is chosen, as modeled in Eq. (4.11). It seems that in many scenarios the same path is chosen, which in turn gets assigned high weights in the objective function. The model (ID-LIN) finds solutions with travel times on the shortest route that are, on average, higher than those of methods (IS) and (ID-SIM) and only slightly lower than those of methods (PD) and (ID-ITR). As discussed in Section 4.1.1, the linear distribution model in (ID-LIN) tends to distribute passengers more evenly on paths than the logit model. Thus, the weights assigned to the shortest paths are lower compared to those in the models (IS) and (ID-SIM). This could explain the worse performance of (ID-LIN) regarding travel time on the shortest path. The remaining three methods, (PS), (PD), and (ID-ITR) perform worse according to travel time on the shortest path. Compared to the best found solutions, their respective travel times are up to three times as far away from an ideal solution.

Figure 6.1b In the case of evaluating travel time using a logit distribution $t_{mp}$, the method (ID-SIM) performs best, which is presumably due to the simulated logit distribution of passengers. The model (ID-LIN) performs, on average, worse than (ID-SIM) and finds solutions that are as good as those found by (PD) and (ID-ITR). This indicates that the passenger distribution of the linear distribution model used in (ID-LIN) is different from the distribution according to a logit model, which is used for evaluation. Furthermore, we can observe that the method (IS) finds better solutions than (PD) and (ID-ITR) averaged over all 21 instances. This is surprising since the methods (PD) and (ID-ITR) consider a passenger distribution according to a logit model, whereas (IS) does not consider any alternatives to the shortest route. We identify the combination of $t_{mp}$ as an evaluation function and a passenger distribution on multiple routes as the reason for this observation. In the model (IS), alternative routes might get assigned high travel times, which implies a low utilization of these routes in a subsequent distribution of passengers according to the logit model. As shown in Example 1 with the comparison of timetables $t^2$ and $t^3$, this can result in lower total travel times for passengers than providing low travel times on all alternative routes. Indeed, with all six methods, we find solutions on certain instances with negative relative evaluation values for $t_{mp}$, implying that the found solutions are ‘better’ than an ideal solution. As in Example 1, this finding appears unexpected at first glance and is undesired for evaluation. This questions whether the total (or average) travel time of passengers, while assuming that passengers distribute over
multiple routes in the network, is a valid evaluation function for public transport timetables.

**Figure 6.1c** The evaluation with the evaluated total utility $u_{\text{sum}}$ shows a different pattern. The methods (PD), (ID-ITR), (ID-LIN) and (ID-SIM) clearly outperform the methods (PS) and (IS). The gap to the evaluation value of an ideal solution is more than halved. On average, the method (ID-LIN) finds the best solutions, almost halving the gap to the ideal solution once more compared to the model (ID-SIM). This is contrary to the observations made with the travel time based evaluation functions $t_{\text{sp}}$ and $t_{\text{mp}}$ where (ID-SIM) performs better than (ID-LIN), see Fig. 6.1a and b. A similar observation can be made for the model (IS). While it performs very good on the travel time based evaluation functions, (IS) yields solutions that are among the worst according to the evaluated total utility.

**Figure 6.1d** We make similar observations with the total logs $u_{\text{log}}$ as the evaluation function. Also here, the methods (PD), (ID-ITR), (ID-LIN) and (ID-SIM) find clearly better solutions than the methods (PS) and (IS). However, when evaluating the found timetables with the total logs, the gaps to an ideal solution are far larger. Furthermore, the solutions of (IS) are, on average, rated better than those of (PS), which is not visible with the other utility based evaluation function $u_{\text{sum}}$ in Fig. 6.1c.

**Cross-figure discussion** As indicated in Table 5.1, we consider four different categories of modeling passengers in optimization approaches for timetabling. They result from a combination of (1) whether a predetermined route choice is assumed or a route choice model is integrated into optimization and (2) whether passengers are assumed to use a single route only or to distribute on multiple routes.

With the utility based evaluation functions, $u_{\text{sum}}$ and $u_{\text{log}}$, our experiments show that the quality of timetables can be considerably improved by considering multiple routes instead of a single route for passengers. All four methods that consider a passenger distribution on multiple routes find solutions with a significantly lower gap to an ideal solution than the two models that assume passengers to use a single route only. In comparison, the integration of a passenger route choice model, as opposed to a predetermined route assignment, did not help to improve the quality of the found timetables according to the utility based evaluation functions. Only the solutions of (IS) are, on average, slightly better than those of (PS), but the others were not in comparison to (PD).

Regarding the travel time based evaluation functions, $t_{\text{sp}}$ and $t_{\text{mp}}$, the methods with an integrated route choice model find better timetables than the corresponding single or multiple route methods that assume a predetermined route choice. Especially the models (IS) and (ID-SIM) could find timetables with significantly better travel time on the shortest path and the latter also on a logit distribution. Considering multiple routes for passengers during optimization instead of only one route yields better solutions for $t_{\text{mp}}$, but not necessarily for $t_{\text{sp}}$ since there just the shortest path is considered for evaluation. Moreover, although the method (PD) finds, on average, better solutions than (PS), (PD) is outperformed by all other methods regarding travel time based evaluation functions. In our experiments, considering multiple routes for passengers is not sufficient to find timetables with best travel times.

We find that considering a passenger distribution on multiple routes mainly improves the utilities, and integrating a passenger route choice model mainly improves the travel times of the found timetables. Furthermore, by integrating a passenger distribution model, it is possible to find solutions with multiple good routes that yield both good travel times and high utilities for passengers on the considered instances. The model (ID-SIM) provided the best solutions regarding the travel time based evaluation functions and comparable solutions with respect to one utility based evaluation function. The model (ID-LIN) could not perform as well as one state-of-the-art approach according to the travel time based evaluation functions but provided the solutions with the best utilities. Thus, by integrating a passenger distribution model, it is possible to find better timetables than the benchmark methods regarding some evaluation functions while maintaining the quality regarding some other evaluation functions.

These improvements by the integration of a passenger distribution model come at the expense of significantly larger models. Table 6.1 shows the average solution times of the six different methods on the discussed 21 instances on the grid network. From the computation times, it is apparent that the two proposed models (ID-LIN) and (ID-SIM) need by far the most time for solving the instances. It took almost 20 minutes to solve the model (ID-SIM) and more than 30 minutes to solve the model (ID-LIN), on average. The other methods were solved within a few seconds.

The second column displays the number of instances that were solved within one hour. The model (ID-LIN) could only find optimal solutions for 17 of the 21 instances and (ID-SIM) provided optimal solutions for 19 instances. After one hour, the model (ID-LIN) had, on average, a gap of more than 5% to the best bound, whereas the simulation based model was close to an optimal solution with a remaining gap of a little more than 1%. The other four methods were always able to terminate within one hour.

**Table 6.1**

<table>
<thead>
<tr>
<th>Method</th>
<th>CPU time</th>
<th>No. instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>(PS)</td>
<td>0.4 s</td>
<td>21/21</td>
</tr>
<tr>
<td>(IS)</td>
<td>5.6 s</td>
<td>21/21</td>
</tr>
<tr>
<td>(PD)</td>
<td>0.8 s</td>
<td>21/21</td>
</tr>
<tr>
<td>(ID-ITR)</td>
<td>0.8 s</td>
<td>21/21</td>
</tr>
<tr>
<td>(ID-LIN)</td>
<td>1952.9 s</td>
<td>17/21 (5.68)</td>
</tr>
<tr>
<td>(ID-SIM)</td>
<td>11840.0 s</td>
<td>19/21 (1.17)</td>
</tr>
</tbody>
</table>

**Table 6.2**

<table>
<thead>
<tr>
<th>Method</th>
<th>CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td>(PS)</td>
<td>1.4 s</td>
</tr>
<tr>
<td>(IS)</td>
<td>9.8 s</td>
</tr>
<tr>
<td>(PD)</td>
<td>30.3 s</td>
</tr>
<tr>
<td>(ID-ITR)</td>
<td>35.7 s</td>
</tr>
<tr>
<td>(ID-LIN)</td>
<td>14007.6 s</td>
</tr>
<tr>
<td>(ID-SIM)</td>
<td>9150.2 s</td>
</tr>
</tbody>
</table>

**6.2. Experiments on Dutch railway network**

We also compare the six different methods on a part of the network of Netherlands Railways, which is described in Section 5.1.2. Table 6.2 shows that the solution times for the Dutch railway instance are generally higher compared to the solution times of the instances on the grid network. Model (ID-LIN) required almost four hours to be solved to optimality, and model (ID-SIM) took on average two and a half hours for solving, whereas the simulation based model was close to an optimal solution with a remaining gap of a little more than 1%. The other four methods were always able to terminate within one hour.
In Fig. 6.2, the evaluation values of all methods are given relative to those of an ideal solution. We observe in Fig. 6.2a and b that two models with an integrated passenger route choice model, (IS) and (ID-SIM), perform best. The gap to an ideal solution is significantly lower compared to the other methods. This is in line with the observation made in the evaluation by the travel time based evaluation functions on the grid instances and demonstrates once more the benefits of integrating a passenger route choice model into timetabling optimization. The model (ID-LIN) provides a solution with higher travel times, but it has notably shorter travel times than the remaining methods on the shortest path and comparable travel times assuming a passenger distribution.

The relative evaluation values of the utility based functions in Fig. 6.2c and d suggest that the method (ID-LIN) performs best, as it was observed on the grid instances. In contrast to the instances on the grid network, there seems to be no visible advantage of the methods that consider a passenger distribution on multiple routes over the methods that assume that passengers use only a single route. Instead, the method (IS) performs better than the two methods (PD) and (ID-ITR) with respect to the logsums.

We find that the solutions found by (IS) and (ID-SIM) dominate the solutions found by all other methods regarding the travel time based evaluation functions, while the consideration of multiple routes brings only a slight advantage to the model (ID-SIM). According to the utility based evaluation functions, the solution found by (ID-LIN) dominates all other solutions. Moreover, the results in Fig. 6.2 demonstrate the importance of a thorough evaluation with multiple evaluation functions. Together with the results on the grid network, these experiments illustrate that an evaluation with a single evaluation function is likely to falsify the interpretation.

7. Conclusion

In this paper, we study the problem of finding a travel time minimal timetable under the assumption that the distribution of passengers on available routes can be modeled using a discrete choice model. We use the logit model to estimate a passenger distribution and formulate this problem as a mixed integer program. Based on this, we develop two linear models proposing different ways to model the interaction of passenger route choice and timetable design. In the first model, we incorporate a novel multidimensional linear passenger distribution model that resembles the characteristics of the logit model. Our second model approximates a logit distribution of the passengers from an integrated simulation framework.

We compare the two timetabling models with integrated passenger distribution with three state-of-the-art methods and a heuristic approach that iterates between timetabling and passenger routing to find travel time optimal timetables for passengers. The experiments are conducted on a set of artificial instances and a part of the network of Netherlands Railways. We provide a thorough comparison of all solutions with four structurally different evaluation functions.

With the integration of a passenger distribution model into a timetabling framework, we were able to find better timetables for passengers than the considered state-of-the-art methods. The gap to an ideal solution for passengers could be significantly reduced for some evaluation functions while performing similarly according to other evaluation functions. In general, the experiments give insight into how two model decisions for passenger distribution on routes affect the solution quality. The first decision examined is whether to consider multiple routes or a single route for passengers, and the second is whether route choice is integrated or the assignment of passengers to routes is predetermined.

It is interesting to observe that the different evaluation functions yield different results for the considered methods. This supports the impression that a comprehensive evaluation with multiple functions is useful and necessary to make clear statements about the quality of methods. In particular, we address observations that a commonly used evaluation function for timetables, the total travel time of passengers, in combination with a passenger distribution model might yield an undesired assessment of the timetable. Our results and a simple example raise the question of whether this function is suitable for evaluation or as objective function when considering a distribution of passengers on multiple paths.

The integration of a passenger distribution model in both timetabling models comes at the expense of significantly higher solution times. Future research could deal with the development of solution approaches to be able to solve large instances. Furthermore, for solution approaches that are based on the simulation of the logit model, it should be investigated how many scenarios should be considered for a good trade-off between computation time and quality of the solution. In particular, it would be interesting to further examine whether the number of scenarios should
depend on characteristics of the problem, such as the number of alternative paths of the OD pairs.

**Funding**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Declaration of Competing Interest**

Authors declare that they have no conflict of interest.

**Acknowledgments**

We thank Dennis Huisman from the Erasmus School of Economics at the Erasmus University Rotterdam and Netherlands Railways as well as Markus Friedrich from the Institute for Road and Transport Science at the University of Stuttgart for their valuable comments and suggestions throughout this work. We also thank two anonymous reviewers for their helpful remarks which greatly contributed to improving this paper.

**Supplementary material**

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2021.06.025

**References**


