Gutenberg School of Management and Economics

& Research Unit “Interdisciplinary Public Policy”

Discussion Paper Series

Strategic Planning for Integrated Mobility-on-Demand and Urban Public Bus Networks

Konrad Steiner, Stefan Irnich

November 2, 2018

Discussion paper number 1819

Johannes Gutenberg University Mainz
Gutenberg School of Management and Economics
Jakob-Welder-Weg 9
55128 Mainz
Germany
https://wiwi.uni-mainz.de/

Contact details

Konrad Steiner
Chair of Logistics Management
University of Mainz
Jakob-Welder-Weg 9
55128 Mainz
Germany

steiner@uni-mainz.de

Stefan Irnich
Chair of Logistics Management
University of Mainz
Jakob-Welder-Weg 9
55128 Mainz
Germany

irnich@uni-mainz.de

Strategic Planning for Integrated Mobility-on-Demand and Urban Public Bus Networks

Konrad Steiner*,[a], Stefan Irnich[a]

[a] *Chair of Logistics Management, Gutenberg School of Management and Economics, Johannes Gutenberg University Mainz, Jakob-Welder-Weg 9, D-55128 Mainz, Germany.*

Abstract

Traditional planning for urban transportation usually assumes a given or calculated split between motorized individual transport and public transportation. However, app-based services and ride-sharing in the field of mobility on demand (MoD) create an intermediate mode of transport, whose long-term role in the urban mobility landscape and within public transport systems is not fully understood as of today. If the public transport industry wants to capture the opportunities these new services offer and mitigate the risks that come with them, planning tools for integrated intermodal networks are indispensable.

In this work, we develop a strategic network planning optimization model for bus lines that allows for intermodal trips with MoD as a first or last leg. For an existing public transport network, we decide simultaneously on the use of existing lines and segments in the future fixed-route network, on areas of the city where an integrated MoD service should be offered, on how MoD interacts with the fixed route network via transfer points, and on passenger routes.

The objective is the optimization of the financial performance of the resulting network, including the additional demand and revenues induced by enhanced network coverage. The model provides several options for modeling MoD costs, which allows for linking with operational models such as dial-a-ride. Moreover, our model considers a range of additional important strategic decisions like subsidies for a potential private sector operator of MoD services and the interplay between the demand for public transport and the role of MoD.

We develop a path-based formulation of the problem and a branch-and-price algorithm as well as an enhanced enumeration-based approach to solve real-world instances to proven optimality. The model is tested on instances generated with the help of real-world data from a medium-sized German city that currently operates around 20 bus lines.

Key words: mobility on demand, public transport, strategic network planning, intermodal networks, multimodality, dial-a-ride, branch-and-price

1. Introduction

Traditional planning for urban transportation usually assumes a given or calculated split between two primary modes of transportation: motorized individual transport and public transporta-

*Corresponding author.

Email addresses: ksteiner@uni-mainz.de (Konrad Steiner), irnich@uni-mainz.de (Stefan Irnich)

tion. However, new trends in mobility on demand (MoD), in particular ride-sharing and the app-based services offered by companies such as Uber and Lyft, induce a new mode of transport that sits between those existing modes. The full implications of this new mode, including its long-term role in the urban mobility landscape and within public transport systems, are not fully understood as of today.

On the one hand, there is huge potential for MoD service operators to decrease dependence on car ownership and to provide better access and egress to public transport, i.e., solving, or at least reducing, the last-mile problem. Within areas of lower demand and during off-peak times, MoD operators may provide less costly yet more attractive services than those currently offered by classical public transport companies. These characteristics have recently motivated the International Association of Public Transport (UITP) to create a new offering named "Digital Platforms" for players like Door2Door, Citymapper, and Uber, and thus to provide a formal framework for collaboration (UITP, 2017). On the other hand, for the public transport systems, there is a risk of revenue decline due to the new competition. Studies like Martínez (2015) even raise the question if urban transportation without classical public transport could be possible. The first showcase for the potential of MoD services to disrupt entire industries is the taxi business, which has faced significant revenue decline since the introduction of the new competition (Economist, 2015). Mainly as a consequence of the opposition from the taxi sector, providers of MoD are currently facing strong regulations that restrict their operations (e.g., Economist, 2017). However, we believe that bans and restrictive regulations will not be a long-term solution and that the overall market for on-demand mobility will grow.

Of course, it is not clear which players will offer MoD-based integrated public transport: It may be done by the likes of Uber and Lyft, or other technology platform providers who do not offer mobility solutions today, or car manufacturers, or rental companies that will establish the transition into this market. The alternative is that public authorities will start offering these services themselves. While it is not certain when exactly this will happen, it seems unlikely that a powerful technology-enabled product like on-demand mobility will be blocked from the mass-market in the long run. Furthermore, the potential of MoD could grow even stronger once it is merged with autonomous vehicles, which will lower operational costs and enable more efficient central planning of transportation systems.

In this paper, we look at the urban transport system from a holistic perspective, with a view to supporting cities and public transport operators in gaining a better understanding about potential future network structures. The key challenge for the public transport industry is to capture the opportunities these new services offer and to mitigate the risks that come with them. We hypothesize in line with recent studies like (masabi, 2018) that the most promising way to achieve this is integrating classical public transport and MoD into a combined system that offers high-quality services and flexibility to the customers based on their needs.

1.1. Scope and application context

In the following, we develop an optimization model that facilitates strategic planning of a public transport bus network. The model incorporates the possibility of intermodal trips using MoD as a first and/or last leg. The terms *intermodal* and *multimodal* are used in line with the most common convention (see, e.g., Willing *et al.*, 2017), where multimodality describes the general concept of travelers having access to different modes of transport, and intermodality the special case where at least two different modes are combined within a single trip.

We decide simultaneously on the use of existing or new bus lines or segments in the future fixed-route network, on areas of the city where an integrated MoD service should be offered, on how MoD interacts with the fixed route network via transfer points, and on passenger routes. We develop a path-based formulation of the problem and, for solving it, a branch-and-price algorithm as well as an enhanced enumeration-based approach. Our computational analysis shows that, with both types of algorithms, real-world instances can be solved to proven optimality.

The primary application of this model will be the review of an existing public transport network, since most regions that are ready for this additional layer of intermodality already have an established public transport network. We expect changes to happen incrementally. However, there is certainly the possibility of some cities leaving out stages of the "public transport evolution" and introducing intermodal trips including MoD right from the start. An adjusted version of our model could also be applied in conjunction with further network planning models when a network is designed from scratch, this is discussed in more detail in Section 3.4. Since bus operations have a lower level of fixed costs and a higher degree of flexibility than train operations, we envisage the model to be applied for different times of the day or for different weekdays. It seems to be a reasonable assumption that future fixed-route components of public transport networks will look different on a weekend than during the week, or in the morning and afternoon peak compared to off-peak times. Our model may also be applied to analyze where it could make sense to subsidize on-demand mobility and thus provide strategic direction for the regulator and transport authorities. It may help to assess the interplay between the demand for public transport and the role of MoD. The above aspects are discussed further in Section 3.4.

1.2. Key decisions and interrelationships

For designing an integrated bus transport network that includes MoD operations, the fundamental question is what decision variables are required and what dependencies between them have to be considered. For our model, they are summarized in the following five paragraphs:

Decisions on the existing public transport network. Since we focus on the review of an existing network, the key decision is which segments of the existing network should still be included in the future network.

Zone-based MoD decisions. While classical network design models decide on connections to include and on line plans, we incorporate MoD decisions based on zones, i.e., on areas where the MoD services will be provided. On the one hand, efficiency of MoD depends on vehicle utilization which is higher when entire areas are serviced and not just selected connections. On the other hand, the definition of zones can increase the usability and popularity of the overall system because zone-based MoD consistent with districts can be visualized and communicated easier than scattered connections.

Decisions influencing MoD costs. In addition, costs for MoD legs behave differently than for classical services with fixed routes and frequencies. An exact routing of MoD vehicles is neither feasible (Posada *et al.* (2016) can only solve instances with up to five requests for the *Integrated Dial-a-Ride Problem* (IDARP) with timetabled fixed route service) nor required for a strategic network optimization. Instead, we differentiate between cost aspects that result directly from the network structure and further aspects impacting the desired service level (e.g., the lead time between booking and start of the trip). The key network-dependent aspect impacting costs is the utilization of

MoD vehicles. Modeling such a dependency on utilization requires at least a rough knowledge of where passengers start or end their trips and where they transfer to or from MoD. Hence, we also include decisions on transfer points into the new model, i.e., at which stations passengers interchange between MoD and the fixed-route network. Our modeling of costs further allows for the incorporation of fixed costs, e.g., for infrastructure, as well as incremental costs associated with servicing the induced demand. This is discussed in detail in Section 3.2.

Decisions influencing travel demand. A significant change in service offering like the provision of MoD will certainly impact travel demand. Improved availability with respect to location and time, together with shorter travel times, motivate more people to travel by public transport. However, possible direct MoD trips from origin to destination at market price will also have a reciprocal effect for public transport companies. Ideally, this step would be based on an assessment per OD-pair, both for pairs already served by the public transport network and for those that are newly covered. Due to the modeling complexity of this endeavor and the challenge of obtaining reliable demand data, in particular for relations that currently do not show any observed public transport trips, we have opted for a simplified modeling approach: In a first step, we determine the *base demand* for public transportation that does not change depending on the offered MoD services. In a second step, we estimate *induced demand* caused by an improved service offering. Both, base demand and induced demand are then included as input parameters for the model. More precisely, we allow a predetermined level of induced demand per MoD zone and consider its effects on revenues and costs. Again, OD-specific characteristics can be included in the demand modeling stages. We believe that we do not lose a decisive level of detail, while the model is kept of moderate complexity and thus still allows for solving real-world instances.

Decisions on passenger itineraries. Furthermore, the routing of passengers is another key modeling aspect as it impacts operational costs as well as the service level. The literature differentiates between *passenger route choice*, where passengers choose their routes in a given network based on their preferences, and the *assignment of passengers to routes*, e.g., by an optimization model according to some given objective. For the design of the combined MoD and public network, we are in a hybrid situation: On the one hand, the MoD service provider can almost completely influence (within a reasonable range) the routing on the MoD legs. As long as the overall travel chain is seamless and convenient, passengers will accept being assigned a specific first and/or last feeding leg with MoD, even if a transfer point passed is not on a shortest path through the overall network. On the other hand, once passengers access the fixed-route network, they choose their routes according to individual preferences, meaning that a central assignment will no longer be accurate. We deal with this characteristic by allowing the model to assign passengers to the *MoD routes* as long as a threshold on overall travel time is not exceeded. Within the fixed network, our model routes passengers via the shortest path in order to estimate and ensure an acceptable overall travel time.

The remainder of this paper is structured as follows: In Section 2, we review the related literature. We discuss the modeling of MoD and present our new model and possible extensions in Section 3. The solution approaches based on a branch-and-price and an enhanced enumeration-based algorithm are presented in Section 4. Subsequently, we discuss selected model outputs and their applications to real-world planning questions in Section 5. We conclude by summarizing our findings and discussing possible next steps for research on integrating MoD into strategic public transport network planning in Section 6.

2. Literature review

The following literature review is divided into four parts covering publications demonstrating the potential impact of MoD systems in Section 2.1, multi- and intermodal network planning in Section 2.2, operational modeling of MoD services in Section 2.3, and a discussion of the positioning and contribution of the paper at hand compared to other works in Section 2.4.

2.1. The potential impact of MoD systems

Over the past few years, multiple prominent studies documenting the current and expected future rise of MoD solutions have been published. A few examples of the publishers are intergovernmental organizations and national associations like the International Transport Forum (ITF, Martínez, 2015) and VDV (2015) as well as leading universities, e.g., the MIT and Stanford (Alonso-Mora *et al.*, 2017; Mitchell, 2008; Pavone, 2016). The ITF report, which focuses on a theoretical case study in Lisbon, has been widely discussed. The authors conduct a simulation based on real trip data and analyze the effects of introducing a fleet of self-driving vehicles on a city-wide level, once as a ride sharing system and once as a car sharing system where passengers use the same cars sequentially. These services are assumed to replace the trips currently undertaken by private car and by bus, and in one extreme scenario even today's metro trips. While this is a multimodal setup, no intermodality is considered and each trip is realized with a single mode. The outcome shows that a ride-sharing system could satisfy current demand with only 10.4% of today's car fleet. This number would only rise to 12.8% when also replacing the metro (for a car sharing system the numbers are 16.8% without replacing and 22.8% with replacing the metro, respectively). Furthermore, this could be achieved while reducing average waiting and travel times significantly.

The Association of German Transport Companies (VDV) also sees the pressure on classical public transport rising due to the additional competition by these new offers (VDV, 2015). They recommend that the public transport industry acts quickly while leveraging public transport's core competency to aggregate high numbers of passengers for affordable prices.

Alonso-Mora *et al.* (2017) look at a similar experiment in New York and provide a mathematical model that designs the vehicle routes while still solving real-world instances. The findings, both in terms of the reduced number of vehicles required and the resulting waiting time, support those of the ITF report. Mitchell (2008) focuses more on the conceptual side and depicts a differentiated MoD concept based on a variety of vehicle types and focuses on MoD as a first or last-mile solution, i.e., on an intermodal context. Finally, Pavone (2016) analyzes a full and autonomous MoD solution, however mentions MoD as a last mile solution as a direction for future research.

In short, these studies show that MoD—in particular when combined with autonomous vehicles—has the potential to turn the urban mobility landscape upside down. This underlines the necessity for public transport operators to actively shape the process of integrating MoD and public transport systems to maintain relevance.

2.2. Multi- and intermodal network planning

A comprehensive survey on the general planning and design of public transport networks is provided by Farahani *et al.* (2013). In particular, their Section 3.3 on the *Multimodal Network Design Problem* (MMNDP) provides a collection of relevant references. Note that none of the MMNDP papers considers intermodal passenger trips, and thus the pertinent literature becomes much more sparse as also identified by Farahani *et al.* (2013, Section 7.3.3). Another survey with a focus on line planning is presented by Schöbel (2011).

The line of research on the IDARP does include intermodality by combining fixed and flexible public transport services. A recent overview can be found in (Posada *et al.*, 2016), where the authors schedule flexible and fixed services simultaneously. This approach creates significant complexity so that only instances with five requests are solved.

A more strategic assessment on designing an intermodal network is provided by Aldaihani *et al.* (2004), who model flexible feeding services via a fixed-grid network. This analytical model optimizes the number of MoD zones as well as the frequencies of the fixed services. The overall setup in (Uchimura *et al.*, 2002) is of a strategic nature as well: A hierarchical network with three layers is sketched with MoD as the third layer for distributing passengers within communities. The actual optimization only focuses on the third level once the network structure of levels one and two is given.

The possible interaction between autonomous on-demand services and public transport is studied in (Salazar *et al.*, 2018). In the first step, passenger and vehicle flows are optimized with a multi-commodity flow model for an integrated autonomous MoD and public transport system. Here, the public transport network is fixed and the MoD vehicles have no specific rules on their service areas or on transfer points to the public transport network. In the second step, market dynamics from the customer and MoD operator perspective are modeled in order to determine an optimal pricing and tolling scheme with respect to social welfare. Realistic instances can be solved with this approach also for the intermodal flow problem due to a mesoscopic approach that assumes one person per MoD vehicle and simplifies the modeling of transfers compared to (Posada *et al.*, 2016). Similar to an autonomous fleet, we also assume in this paper that MoD operations are steered centrally, i.e., that rules on operating areas, allowed transfer points etc. can be imposed.

Shen *et al.* (2018) also assess the integration of an autonomous MoD system with public transport. Agent-based simulation is applied to study the consequences of replacing certain bus routes with low ridership figures by MoD services. While the bus routes of the public transport system are fixed a-priori, a wide range of scenarios is analyzed with respect to service quality, total distance traveled as well as financial aspects. The authors find that there are integrated scenarios that simultaneously improve the service quality in all three dimensions. However, there are also scenarios that show an aggravation of the system performance, which underlines the necessity for decision support on determining which areas to serve by which mode.

Li and Quadrifoglio (2010) provide decision support on whether to set up a fixed-route or a demand responsive service with the option of changing between these two service types during the day. Single-vehicle operations are assumed for the demand responsive service and a simulation is performed to determine total passenger travel times for both service types. The computations show that demand responsive services provide shorter travel times for lower demand densities, hence confirming our approach of considering MoD as a valid alternative to fixed-route services in particular in areas of lower demand. Navidi *et al.* (2017) study a similar setup with agent-based simulation and also allowing higher numbers of MoD vehicles in order to realize a zero rejection rate. The results of the simulations show that the demand responsive service provides a superior service level to fixed line services while also being more cost effective in areas of lower demand.

Häll *et al.* (2008) and Edwards *et al.* (2012) present simulation approaches on networks allowing for intermodal trips with MoD as a first or last leg. The paper of Häll *et al.* includes a detailed discussion of the interface between fixed and flexible services and provides a study conducted with the simulation software LITRES-2 for the Swedish town Gävle. One interesting finding is the importance of transfer point placement, which supports our choice of including decisions on transfer point locations explicitly in the optimization model. Edwards *et al.* (2012) work with a 'network-inspired framework for integrated transport systems' that combines traditional public transport with MoD systems and applies this approach to the city of Atlanta. The study concludes that particularly areas of lower density can benefit from an improved service level, which is enabled by the offering of MoD.

Another research area relevant in the context of MoD is *dynamic ride-sharing*, where travelers with similar itineraries are grouped together in the same vehicle, but vehicles are still driven by private drivers. On the one hand, the models developed for dynamic ride-sharing are very similar to MoD system with a dedicated fleet. On the other hand, it might make more sense to set up an MoD system based on ride sharing, in particular if car penetration is high, e.g., in rural areas. Stiglic *et al.* (2018) study an integrated system that offers point-to-point ride-sharing as well as ride-sharing to a transit station or even park-and-ride, where the driver parks his car and uses public transport for the remainder of his journey. In a first step, possible matches are identified based on the required service level. In a second step, an optimization model is set up to maximize the number of matched riders. The computational results show how additional flexibility of drivers, an increased number of system participants, as well as increased train frequencies positively impact the number of successful matches. More general background on dynamic ride-sharing is presented in the survey of Agatz *et al.* (2012).

Rothenbächer *et al.* (2016) design a freight network for combined transport, i.e., each request can be shipped by a combination of road and rail transportation. The presented optimization model decides on the intermodal routing of a given set of requests as well as on hub locations for the transfers and frequencies for the rail connections. Like in our case, a path-based formulation is chosen and a branch-and-price algorithm is implemented that allows for solving realistic instances in reasonable computation time.

Readers interested in more general literature on multi- and intermodal network planning are referred to (Chowdhury and Ceder, 2016) for an assessment of the user's perspective, and to (Willing *et al.*, 2017) for a view on the existing inter- and multimodal mobility landscape and ideas on how it can be enhanced in the future. Further, Jittrapirom *et al.* (2017) provide an overview on Mobility as a Service (MaaS) systems, that always feature MoD elements, including an extensive research agenda that considers aspects like demand modeling, business models etc. The journey planning aspects of including intermodal trips with potential MoD legs are discussed in (Friedrich and Noekel, 2015) and (Horn, 2004). Finally, a more holistic view in context of mobility-oriented development of cities can be found in (Smolnicki and Sołtys, 2016) and referenced papers.

2.3. Operational modeling of mobility on demand

Wang and Odoni (2016) focus on the last-mile problem where passengers travel from a public transport stop or station to the final destination. However, the public transport network itself is not in scope of their study. Instead, they derive analytical expressions for the expected waiting time until boarding and expected riding times and evaluate them against a simulation approach, in the first step for the unit-capacity case, i.e., one passenger per vehicle, and in the second step for vehicle capacities up to 20. The main independent variables considered are the headway of the public transport service, the batch size of customers alighting from the train or bus, the average distance passengers travel to their final destination, the fleet size of the last-mile service, as well as additional statistical parameters. The computational studies show how expected waiting times rise when the utilization of the service increases. These insights can be used to determine an appropriate MoD fleet size to achieve a desired service level.

A similar setup for a stand-alone MoD service is analyzed by Diana *et al.* (2006) where the fleet size is analytically estimated based on service level requirements. The key inputs are the distribution of demand as well as time windows for the passengers. The output is then compared to a simulation model in order to validate the analytical expressions. The computational study shows how the number of required vehicles increases with a growing number of requests and for smaller time windows. Additionally, it can be observed from (Diana *et al.*, 2006, Table 1 in Section 5.3) that the number of vehicles per request actually decreases with a growing number of requests due to better utilization, which supports our approach of carefully reflecting on utilization aspects in a strategic model.

Another modeling approach for a stand-alone MoD service can be found in (Martínez *et al.*, 2015). Passengers with close origins and destinations are clustered in a first optimization step and assigned to vehicle routes in a second step. The last step is the maximization of the operational profit under the condition that passenger requirements on travel duration and arrival time are respected. This process also permits the determination of a suitable fleet size.

Archetti *et al.* (2017) present a simulation study to assess the performance of an MoD system in terms of costs, service quality, and travel time depending on demand density. Standard buses and private cars are available as alternatives and the authors simulate the mode choice of users in their approach.

The classical approach to optimize the routing of demand responsive services is focused on solving the Dial-a-Ride Problem (DARP). Surveys on the DARP are presented in (Cordeau and Laporte, 2007) for literature up to 2007 and in (Ho *et al.*, 2018) for more recent works.

2.4. Positioning and contribution of this work

The main contribution of this paper is the development of a strategic network planning optimization model that allows for intermodal passenger trips with MoD as a possible mode for access and egress legs. Decisions on both the fixed public transport network and the MoD operations are taken simultaneously. To our knowledge, no model with these combined characteristics currently exists.

We briefly reference literature on the two additional important modeling aspects that were discussed in Section 1: the inclusion of dynamic demand and the balancing between passenger routing and assignment. Klier and Haase (2014) design a network and explicitly model passenger paths between OD-pairs, which allows assigning specific demand levels to each path based on a demand model that is calculated before the optimization step. Since the inclusion of MoD significantly changes the public transport offer, dynamic demand approaches are required in our case as well. Regarding passenger routes, Goerigk and Schmidt (2017) and Schmidt and Schöbel (2015) provide discussions on the differentiation between route assignment and route choice. In the first case, the model routes passengers to achieve an optimal objective value. In the second case, passenger route selection is reflected in the model, typically in a bi-level modeling approach. The authors develop two integrated integer programming formulations that allow for passenger route choice. As discussed, a combination of these two approaches fits best in our context: On the one hand, system operators will be able to dictate the MoD legs of the passengers within a reasonable range. On the other hand, passengers will choose routes according to their own preferences once they have entered the public transport network.

Paper	Context			Key decisions	Modeling			Approach†
	Fixed network	MoD	Inter-modality		MoD costs	Induced demand	Passenger routes	
This work	yes	yes	yes and zones	Segments	cost/time aspects considered via transfer points	Considered per zone (not per OD-pair)	Mix of routing and assignment	opt
(Aldaihani *et al.*, 2004)	yes	yes	yes	Number of zones n (all zones offer MoD with one transfer point), frequency of fixed services	Fixed costs per vehicle, variable costs based on demand level and zone size	Not considered	Unique, once n is fixed	ana
(Edwards *et al.*, 2012)	yes	yes	yes	No decisions	Variable costs based on simulated distance traveled per vehicle	Not considered	Unique based on closest transfer point	sim
(Häll *et al.*, 2008)	yes	yes	yes	No decisions	Fixed costs per vehicle, variable costs based on simulated distance traveled per vehicle	Not considered	Mix of routing and assignment	sim
(Li and Quadrifoglio, 2010)	yes	yes	no	Provision of MoD or fixed-route services as feeder to/from a terminal station	Only travel time costs for the customer considered	Not considered	Unique via pre-determined transfer points	ana, sim
(Navidi *et al.*, 2017)	yes	yes	no	Provision of MoD or fixed-route services, potentially as feeder to/from a terminal station	Fixed costs per vehicle, variable costs based on simulated distance traveled per vehicle	Not considered	Unique via pre-determined transfer points	sim
(Posada *et al.*, 2016)	yes	yes	yes	Scheduling of MoD services and intermodal routing of each transport request	Variable costs based on distance traveled per vehicle	Not considered	Routes are fully assigned by the model considering service level requirements on the transfers in detail	opt
(Salazar *et al.*, 2018)	yes	yes	yes	Intermodal passenger routes, pricing and tolling scheme	Variable costs based on distance traveled per vehicle	Not considered	Routes are fully assigned by the model	opt
(Shen *et al.*, 2018)	yes	yes	yes	Inclusion of segments in existing network	Variable costs based on simulated distance traveled per vehicle	Not considered	Unique based on closest transfer point	sim
(Uchimura *et al.*, 2002)	yes	yes	no	MoD vehicle routes	Variable costs based on simulated distance traveled per vehicle	Not considered	Only MoD leg is considered	opt

Table 1: Positioning and contribution of this work compared to other articles; † : opt = optimization, sim = simulation, ana = analytical

Table 1 compares our model to a selection of the existing literature discussed above in Section 2.2.

3. Optimization model for intermodal networks with MoD

This section on the optimization model for intermodal networks with MoD first introduces two important foundations for the model: first, the passenger movements through the network, as they correspond to the central path variables of our model, in Section 3.1, and second a discussion of MoD decisions and the modeling of costs and travel times in Section 3.2. Subsequently, we provide the *integer linear programming* (ILP) formulation in Section 3.3 and comment on further model applications and extensions in Section 3.4.

The following notation is used in the remainder: The current public transport network is given by an undirected graph $(\mathcal{V} \cup \mathcal{W}, \mathcal{E})$ with vertices $\mathcal{V} \cup \mathcal{W}$ and edges $\mathcal{E}$ defined as follows:

$k, l \in \mathcal{V}$	Vertices that represent *origins* and *destinations* of passenger trips.
$(k,l) \in OD \subset \mathcal{V} \times \mathcal{V}$	The set of *OD-pairs*.
$i, j \in \mathcal{W}$	Vertices that represent the current *stops* of the public transport network. We assume that $\mathcal{V}$ and $\mathcal{W}$ are disjoint even though some vertices may refer to the same physical location.
$e \in \mathcal{E}(\mathcal{W})$	Edges (with both endpoints in $\mathcal{W}$) that represent the direct *connections* (two consecutive stops) of the current public transport network.
$s \in \mathcal{S} \subset 2^{\mathcal{E}(\mathcal{W})}$	Pairwise disjoint *segments* (=subsets) of edges that represent the parts of the network that could be removed (in one piece). We refer to these as *omission segments*. A significant set of edges, the core network, is usually not under discussion.

The new possibilities to offer zone-based MoD services are described by:

$z \in \mathcal{Z}$	Potential disjoint *MoD zones* that cover the area in scope. Hence, for every $k \in \mathcal{V}$ there is a unique zone $z_k \in \mathcal{Z}$ indicating the inclusion of the corresponding physical location in the zone.
$\mathcal{T}_z \subset \mathcal{W}$	Subsets of vertices that represent potential *transfer points* $\mathcal{T}_z$ for every zone z. Transfer points are stops of the current network and can be used to connect the public network with MoD. We explicitly allow $i \in \mathcal{T}_{z_1} \cap \mathcal{T}_{z_2}$ for a specific transfer point i and distinct zones $z_1, z_2 \in \mathcal{Z}$. Particularly, the physical location of a transfer point $i \in \mathcal{T}_z$ can be located outside the physical zone z.
$(k,i), (j,l) \in \mathcal{E}(\mathcal{W}, \mathcal{V})$	Edges (with one endpoint in $\mathcal{V}$ and one in $\mathcal{W}$) that represent *access legs* (k,i) and *egress legs* (j,l). $\mathcal{E}(\mathcal{W}, \mathcal{V})$ is the disjoint union of $\mathcal{E}(\mathcal{W}, \mathcal{V})^{walk}$ and $\mathcal{E}(\mathcal{W}, \mathcal{V})^{MoD}$, i.e., we assume that one can a priori decide for each such leg whether passengers walk to the stop (short distance) or must use the MoD service (longer distance). For the latter case, it is required that $i \in \mathcal{T}_{z_k}$ or $j \in \mathcal{T}_{z_l}$ holds, respectively.

3.1. Passenger movements

For each OD-pair $(k,l) \in OD$, our model must ensure that passengers can travel between k and l with a similar service level as in the existing status quo in the future network. For that purpose, we assume that in the current network passengers have chosen a path that is optimal for them, i.e., a quickest connection. If passengers do not use their current route in the future network, we require the new total travel time to not exceed today's travel time by more than a given threshold θ (e.g., $\theta = 10\%$). Let $\mathcal{P}^{kl}$ denote the set of paths from k to l consistent with the passengers service-level requirements.

We describe the paths $P \in \mathcal{P}^{kl}$ using public transport in more detail now. For such a path P, the part realized by walking or MoD is completely determined by the access and egress transfer points $i \in \mathcal{W}$ and $j \in \mathcal{W}$. Thus, the passenger path P can have a maximum of three legs: First, an *access leg* $(k, i) \in \mathcal{E}(\mathcal{W}, \mathcal{V})$, which could either be walking to a close stop or an MoD leg. Second, a *public transport leg* (i, j) in the fixed-route network given by a path in $(\mathcal{W}, \mathcal{E}(\mathcal{W}))$ (here the case $i = j$ is possible for combining two MOD legs or MOD and walking). The third leg is an *egress leg* $(j, l) \in \mathcal{E}(\mathcal{W}, \mathcal{V})$, which can again be realized by walking or by MoD. Here, we assume all passengers traveling from k to l decide for the same mode (walking or MoD) from k to i and from l to j, respectively.

For some OD-pairs, it is also possible that in the future network the passengers use only one direct MoD trip from k to l. For such a pair $(k, l) \in OD$, the trip is denoted by $P^{kl}_{\text{dir}} \in \mathcal{P}^{kl}$. For the sake of convenience, we also define the subset $\mathcal{P}^{kl}_s$ of paths $\mathcal{P}^{kl}$ that passes through at least one edge of an omission segment $s \in \mathcal{S}$. Finally, for $P \in \mathcal{P}^{kl} \setminus \{P^{kl}_{\text{dir}}\}$, the two access and egress transfer points are denoted by i_P and j_P.

3.2. Modeling of MoD decisions and costs

As motivated in the introduction, MoD decisions are not based on edges and lines like in classical network planning approaches, but on zones. The overall area in scope is divided into disjoint zones $\mathcal{Z}$. Defining zones is certainly part of the problem modeling, and we further analyze the impact of different zone sizes in the computational study in Section 5.3. For the moment, we assume potential zones to be given and fixed. For each zone $z \in \mathcal{Z}$, the task is to decide whether and how MoD operations are offered.

Since the accurate inclusion of MoD costs in a strategic model is one of the key challenges when planning an intermodal network, we discuss this topic in some more detail now. Our fundamental decision, per chosen zone $z \in \mathcal{Z}$, is the choice of a subset of transfer points I among those that are available ($I \subseteq \mathcal{T}_z$). We assume that only some subsets I are feasible choices. For example, I must be non-empty and one may want to limit the number of transfer points to a maximum cardinality I^{max} so that only sets $I \in \mathcal{T}_z$ with $|I| \leq I^{max}$ are allowed. This is captured in the definition of the set $T^{\mathcal{I}}_z \subset 2^{\mathcal{T}_z}$.

Clearly, the desired service level impacts MoD costs. Specifically, this includes the (average) lead time between booking and start of the trip, the discrepancy between ideal and actual departure or arrival time, the temporal integration with the timetable of the fixed-route network, the level of permitted detour, the required walking distance if pick-up or drop-off points are used, and the aspired share of satisfied requests. These aspects do not depend on the decisions our model is taking and can therefore be treated beforehand. Indeed, we envisage our model to be applied in conjunction with more operational models that are tailored to the characteristics of the MoD system being designed. The outputs on MoD costs and times of the operational models serve as input parameters to the strategic model we present here. Hence, our optimization model is flexible to include a large variety of aspects without further increasing in size.

We now discuss in more detail that all remaining important cost and timing aspects can be included, in particular:

(i) MoD infrastructure costs per zone

(ii) MoD costs associated to zone-based induced demand
(iii) MoD leg costs for the base demand
(iv) Resulting passenger travel times

MoD infrastructure costs per zone. For a zone $z \in \mathcal{Z}$, any infrastructure that needs to be set up and other general investments that generate fixed costs clearly depend on the chosen transfer points. The model considers such fixed costs explicitly as costs c_{zI}^{fix} based on the decision which transfer points $I \in T_z^{\mathcal{I}}$ need to be established.

MoD costs associated to zone-based induced demand. Recall from Section 1.2 that our model captures the already existing demand of the public transport network, which we refer to as *base demand*, and additional demand that is generated by the extended service offering, the *induced demand*. We treat these two types of demand separately in the MoD cost modeling starting with the induced demand.

More transfer points per zone lead to more options for the passengers, and thus positively impact induced demand. Hence, we assume that it is possible to pre-calculate induced demand d_{ziI} per zone z and transfer point i depending on the set of offered transfer points I. Once the demand level d_{ziI} has been modeled, additional cost and revenue parameters c_{ziI} and r_{ziI} can be calculated and included in our model.

MoD leg costs for the base demand. As the cost (or revenue) of induced demand is already included (see above), the remaining modeling aspect is the incorporation of MoD leg costs for the base demand. Recall that base demand is given per OD-pair $(k, l) \in OD$ as a number of passengers d^{kl}. Clearly, one can consider the zones z_k and z_l separately. We therefore discuss MoD leg costs for the zone $z_k \in \mathcal{Z}$ for an origin $k \in \mathcal{V}$ only.

First, we approximate the *effective distance* traveled by MoD. By using the word 'effective' we account for the fact that actual distances traveled may differ due to detours when additional passengers are picked up. We include this effective distance by capturing start point k and the transfer point i to the fixed-route network explicitly in the model using a cost per passenger parameter c^{ki}. It makes sense that these cost parameters are on a passenger level because costs for on-demand services depend heavily on the number of serviced passengers.

Second, density and structure of demand strongly impact efficiency, as higher densities result in more pooling potential and therefore higher vehicle utilization and fewer detours. This aspect cannot be explicitly and exactly included without significantly increasing the complexity of modeling routes. However, we can still capture the MoD system's efficiency partly in the input parameters of our model. On the one hand, the set of transfer points $I \in T_z^{\mathcal{I}}$ directly impacts the efficiency of the MoD operations within the zone, since feeding all passengers starting from z into one transfer point allows significantly more consolidation than feeding into two or even three transfer points. We have thus opted for including the chosen set of transfer points I in an *inefficiency contribution* c_{zI}^{ineff} for the base demand. The cost c_{zI}^{ineff} is in general higher the more transfer points are selected and the more spread out they are. On the other hand, *local system efficiency* can be partly taken into account when computing c^{ki} depending on the likely MoD demand levels per zone z.

Resulting passenger travel times. Travel times are relevant for identifying possible paths through the network (see Section 3.1) and they follow a similar logic as explained for MoD costs. As we model passenger paths explicitly only for the base demand d^{kl}, they are only required in this case. The travel time to get from k to a transfer point i (or from transfer point j to l) is based on the distance between the stops and adjusted based on preprocessed expected demand levels to account for efficiency. Furthermore, expected initial waiting times have been included in the travel times.

Overall, we can conclude that our model is able to cover the majority of factors that influence MoD costs. Since this is based on the actual operational costs of providing MoD, it represents the most complex setup. In practice, there might be simpler commercial agreements with an operator, e.g., fixed remuneration per passenger trip, that simplify the modeling from the perspective of the network planner.

3.3. Path-based formulation

Our path-based formulation implicitly captures the complicated rules for allowed passenger paths through the network. Three types of binary decision variables are required:

$x_s \in \{0,1\}$	indicates whether an omission segment $s \in \mathcal{S}$ is included in the future network (=1), or not (=0).
$y_{zI} \in \{0,1\}$	indicates whether $I \in T_z^{\mathcal{I}}$ is the set of transfer points of an MoD zone $z \in \mathcal{Z}$.
$\lambda_P^{kl} \in \{0,1\}$	indicates whether path $P \in \mathcal{P}^{kl}$ is selected for an OD-pair $(k,l) \in OD$.

The following coefficients need to be defined; the majority of them have already been introduced and motivated in the above paragraphs:

d^{kl}	Number of passengers traveling in the future intermodal public transport network on the OD-relation $(k,l) \in OD$ (base demand).
d_{ziI}	Additional induced demand passing through transfer point $i \in \mathcal{T}_z$ that can be generated in zone $z \in \mathcal{Z}$ if MoD is offered with transfer points $I \in T_z^{\mathcal{I}}$.
c_s	Costs for operating public transport with predetermined frequency on the omission segment $s \in \mathcal{S}$.
c_{zI}^{fix}	Infrastructure and other fixed costs for operating MoD within zone $z \in \mathcal{Z}$ and for using the transfer points $I \in T_z^{\mathcal{I}}$.
c_{zI}^{ineff}	Inefficiency costs for operating MoD within zone $z \in \mathcal{Z}$ with transfer points $I \in T_z^{\mathcal{I}}$ due to the fact that MoD operations become less efficient if there are more or scattered transfer points.
c^{ki}	Variable costs for feeding a passenger via MoD from $k \in \mathcal{V}$ into station $i \in \mathcal{T}_{z_k}$ (or vice versa).
c^{kl}	Variable costs for transporting a passenger via a direct MoD trip from $k \in \mathcal{V}$ to $l \in \mathcal{V}$. In this case, the costs per km are usually significantly higher as there is decreased consolidation potential.
r_{ziI}	Revenue per passenger generated by the additional demand d_{ziI}.

All input parameters are assumed to be non-negative.

For the sake of brevity, we define the cost parameter (per zone $z \in \mathcal{Z}$ and set $I \in T_z^{\mathcal{I}}$ of its transfer points)

$$c_{zI} = (c_{zI}^{\text{fix}} + c_{zI}^{\text{ineff}}) + \sum_{i \in I} (c_{ziI} - r_{ziI}) d_{ziI}$$

covering infrastructure and other fixed costs, inefficiencies that result from the choice of the transfer points $I \in T_z^{\mathcal{I}}$, and the financial impact of induced demand, which also considers potential additional revenues. With this definition, the path-based formulation reads as follows:

$$\min \sum_{s \in \mathcal{S}} c_s x_s + \sum_{z \in \mathcal{Z}} \sum_{I \in T_z^{\mathcal{I}}} c_{zI} y_{zI} + \sum_{(k,l) \in OD} \left(c^{kl} d^{kl} \lambda_{P_{\text{dir}}^{kl}}^{kl} + \sum_{P \in \mathcal{P}^{kl} : P \neq P_{\text{dir}}^{kl}} (c^{ki_P} + c^{lj_P}) d^{kl} \lambda_P^{kl} \right) \tag{1a}$$

$$
\begin{aligned}
\text{s.t.} \sum_{I \in T_z^{\mathcal{I}}} y_{zI} &\leq 1, && \forall z \in \mathcal{Z} && \text{(1b)}\\
\sum_{P \in \mathcal{P}^{kl}} \lambda_P^{kl} &= 1, && \forall (k,l) \in OD && \text{(1c)}\\
\sum_{P \in \mathcal{P}_s^{kl}} \lambda_P^{kl} &\leq x_s, && \forall s \in \mathcal{S}, \forall (k,l) \in OD && \text{(1d)}\\
\sum_{\substack{P \in \mathcal{P}^{kl}: i_P = i,\\ (k,i) \in \mathcal{E}(\mathcal{V},\mathcal{W})^{MOD}}} \lambda_P^{kl} &\leq \sum_{I \in T_{z_k}^{\mathcal{I}}: i \in I} y_{z_k I}, && \forall (k,l) \in OD, \forall i \in \mathcal{T}_{z_k} && \text{(1e)}\\
\sum_{\substack{P \in \mathcal{P}^{kl}: j_P = j,\\ (j,l) \in \mathcal{E}(\mathcal{V},\mathcal{W})^{MOD}}} \lambda_P^{kl} &\leq \sum_{I \in T_{z_l}^{\mathcal{I}}: j \in I} y_{z_l I}, && \forall (k,l) \in OD, \forall j \in \mathcal{T}_{z_l} && \text{(1f)}\\
\lambda_{P_{\text{dir}}^{kl}}^{kl} &\leq \sum_{I \in T_z^{\mathcal{I}}} y_{zI}, && \forall (k,l) \in OD, \forall z \in \mathcal{Z} : z = z_k \text{ or } z = z_l && \text{(1g)}\\
x_s &\in \{0,1\} && \forall s \in \mathcal{S} && \text{(1h)}\\
y_{zI} &\in \{0,1\} && \forall z \in \mathcal{Z}, \forall I \in T_z^{\mathcal{I}} && \text{(1i)}\\
\lambda_P^{kl} &\in \{0,1\} && \forall (k,l) \in OD, \forall P \in \mathcal{P}^{kl} && \text{(1j)}
\end{aligned}
$$

The objective (1a) comprises the costs of operating the segments of the fixed-route network that are under review and MoD variable and fixed costs based on the decisions on zones, transfer point setup, and passenger paths. Constraints (1b) ensure the selection of at most one subset $I \in T_z^{\mathcal{I}}$. Furthermore, constraints (1c) guarantee that a path through the network is chosen for each OD-pair. The chosen paths need to be consistent with the choice of included segments, i.e., constraints (1d), and with the MoD zones and transfer points, i.e., constraints (1e) and (1f). Constraints (1g) ensure that direct MoD trips can only be realized if MoD is offered in both, the origin and destination zone. Finally, the domains of the variables are stated in (1h)–(1j).

In case there are requirements on the topology of the resulting fixed-route network, e.g., that it should possess no gaps, the model must be extended with additional constraints in the form of $x_{s_1} - x_{s_2} \leq 0$ to ensure the choice of an omission segment s_1 enforces omission segment s_2 to be chosen as well.

3.4. Model application and extensions

In some applications, the primary focus could be on service level, in particular travel times, rather than financial aspects. While travel times are only included as a constraint defining the allowed paths, they can also be included in the objective function. Indeed, every path $P \in \mathcal{P}^{kl}$ immediately yields the change in travel time compared to the status quo for the passengers of the corresponding OD-pair (k,l), and the delta would then be the objective coefficient of λ_P^{kl}. Such an approach could even be extended to also cover travel time changes for the induced demand, when status quo travel times for the current transport modes of the new passengers are assumed.

A similar model could also be used as a component for planning a new public transport network after having determined a rough network structure and a line pool with other planning tools. In this situation, we see two key hurdles to overcome: First, we expect a significantly higher number of potential line segments and less “fixed” edges, i.e., less edges not contained in any omission segment. Depending on the instance size, this could create a computational challenge. Second, demand data as a crucial model input would have a higher degree of uncertainty. MoD costs are modeled to be (in essence) proportional to the ridership, hence it is essential to perform a detailed sensitivity analysis to validate the model recommendations.

Another relevant application is given by a public-private collaboration. Here, the actual MoD operations would be provided by a private company, e.g., from the taxi or ride-sharing industry. This approach could be motivated by significantly lower costs of the private companies, e.g., due to higher driver utilization due to synergies with the core business. First examples of such collaborations can already be observed in practice, mainly in the US (e.g., Jaffe, 2015). From a modeling perspective, the main parameters that change are the MoD costs. As discussed in Section 3.2, our model is flexible enough to reproduce different cost and remuneration structures and could for instance be employed to study whether a certain level of subsidy makes sense from a financial perspective. In this case, the model does not need to consider the actual costs of providing the service, but only the share that is covered by the public authority. The underlying rationale would be to ensure a setup that is still attractive for the private operator, while offering attractive price and service level to the customers also in areas of lower demand, and still being cheaper for the public transport provider than realizing the services himself.

4. Solution algorithms

We solve the real-world instances with two different solution approaches based on the model (1). To reduce the model size, both approaches start with a preprocessing procedure described in Section 4.1. The first approach is a *branch-and-price algorithm* that generates OD-paths dynamically and is discussed in Section 4.2, the second one enumerates all relevant paths and then uses an *integer program* (IP) solver (CPLEX) directly on model (1). We refer to the latter approach as *enhanced enumeration* and present it in Section 4.3.

4.1. Preprocessing

The main drivers of the model size are the potentially exponentially-sized sets of path variables λ_P^{kl} and of constraints (1c)–(1f). Hence, the preprocessing focuses on the reduction of these variables and constraints.

First, one can exploit the fixed part of the public transport network, e.g., the central area of the city where no or only a few edges are considered to be omitted. For all OD-pairs $(k, l) \in OD$ for which the current connecting path is fully contained in this fixed part of the network, the model does not need to take a routing decision as passenger routes remain unchanged. The same argument can be extended for OD-pairs $(k, l) \in OD$, for which at least one path $P \in \mathcal{P}^{kl}$ that respects the constraint on total travel time lies entirely in the fixed network. All these OD-pairs and associated constraints (1c) can therefore be removed from the model. Section 5 shows that a significant reduction of the model results for the real-world instances of our computational study.

The second idea is to reduce the number of OD-pairs by requesting identical MoD routing decisions for "close" OD-pairs with similar characteristics. This is in fact a heuristic approach to reduce the model's size, which makes sense from a consistency and ease-of-use perspective, as it is intuitive to route passengers the same way, e.g., when they share the same origin and have a close destination. Recall that we assume that we can assign passengers to MoD routes as long as a constraint on the total travel time is fulfilled. In order to aggregate such "close" OD-pairs (k_1, l_1) and (k_2, l_2) into a single OD-pair, we require the following: First, both origins k_1, k_2 (both destinations l_1, l_2) must lie in the same zone $z_{k_1} = z_{k_2}$ (and $z_{l_1} = z_{l_2}$). Second, the two OD-pairs must have identical profiles in terms of path preferences. In case both origins k_1, k_2 (and destinations l_1, l_2) share the same set of closest public transport nodes $i_0, i_1, \ldots$ (and $j_0, j_1, \ldots$) within walking distance we can compare their profiles in terms of travel times. For this purpose, we take all possible combinations of transfer points for the three cases (a) one initial MoD leg, (b) one final MoD leg, and (c) one initial *and* one final MoD leg, and consider the duration of the resulting travel chains. If both OD-pairs show the same preference ranking when the combinations are ordered by the travel times, we can treat them as one OD-cluster.

In the case of demand data being available only at the level of the i and j like in the example we discuss in Section 5, this approach would not reduce of the number of OD-pairs. Hence, we suggest an additional aggregation step: If k_1 and i_1 as well as k_2 and i_2 (and l_1 and j_1 as well as l_2 and j_2) represent the same physical location, respectively, these pairs can be aggregated if they show identical path preferences as discussed above and additionally the stops i_1, i_2 (and j_1, j_2) are identically located with respect to the omission segments $s \in \mathcal{S}$. Specifically, they should either be both connected to the fixed part of the network or all neighboring edges should be part of the same omission segment $s \in \mathcal{S}$. To obtain the objective coefficient of λ_P^{kl} for the OD-cluster of (k, l), we sum over the $(c^{k' i_P} + c^{l' j_P}) d^{k' l'}$ for all pairs (k', l') within the cluster of (k, l). For our computational study in Section 5, we only aggregate OD-pairs with identical path preferences. Depending on the specific application, one could even further aggregate pairs with path preferences that are reasonably close.

4.2. Branch-and-price algorithm

The branch-and-price algorithm relies on the following principles (Desaulniers *et al.*, 2005): The *restricted master problem* (RMP) is given by the linear relaxation of model (1) defined over subsets $\mathcal{P}^{kl}_{\mathrm{RMP}} \subset \mathcal{P}^{kl}$ of all possible OD-paths. The linear relaxation of (1) is then solved via column generation. For each RMP solution, pricing subproblems, one for each OD-pair $(k, l) \in OD$, identify negative reduced cost paths $P \in \mathcal{P}^{kl}$ to be added to $\mathcal{P}^{kl}_{\mathrm{RMP}}$. We assume that the direct path P^{kl}_{dir} is always in $\mathcal{P}^{kl}_{\mathrm{RMP}}$ so that only non-direct paths need to be generated. As long as at least one negative reduced cost path exists, the process is iterated. Integer solutions to (1) are then produced by branching.

We start by discussing the subproblems and subsequently present our branching strategy.

4.2.1. Subproblem

Since the number of feasible paths can be huge, we generate them dynamically in a subproblem. For each OD-pair $(k, l) \in OD$, we have to identify paths $P \in \mathcal{P}^{kl}$ with negative reduced costs $\tilde{c}^P$. We denote the dual prices of the constraints (1c), (1d), (1e), and (1f) by $\gamma^{kl}, \delta^{kl}_s, \epsilon^{kl}_i$, and ζ^{kl}_j, respectively. The pricing subproblem for (k, l) is then:

$$\min_{P \in \mathcal{P}^{kl}} \tilde{c}^P = \min_{P \in \mathcal{P}^{kl}} \left(c^{k i_P} + c^{l j_P} \right) d^{kl} - \gamma^{kl} - \sum_{s \in \mathcal{S}: P \in \mathcal{P}^{kl}_s} \delta^{kl}_s - \epsilon^{kl}_{i_P} - \zeta^{kl}_{j_P}. \tag{2}$$

We solve the pricing problem of $(k, l) \in OD$ as a *shortest-path problem with resource constraints* (SPPRC, Irnich and Desaulniers, 2005), where we minimize reduced costs while respecting a constraint on the total travel time. We first describe the underlying SPPRC digraph. This digraph is a network with four layers and it is specific for each OD-pair (k, l). In the associated SPPRC, feasible paths start at a copy of the origin k in layer one and end at a copy of l in layer four.

The vertices of the four layers can be described as follows: We use the notation $v^{(n)}$ for the vertices, indicating a representation of node $v \in \mathcal{V} \cup \mathcal{W}$ in the n-th layer of the pricing network. The first layer consists of a single vertex only, which is a copy $k^{(1)}$ of the origin vertex $k \in \mathcal{V}$, to be used as the network source. The second layer comprises copies $i^{(2)}$ of vertices $i \in \mathcal{W}$ where a passenger could potentially access the fixed-route public transport network. Thus, this layer is restricted to stop points $i \in \mathcal{T}_{z_k} \subset \mathcal{W}$, i.e., where an MoD leg exists or station i is within walking distance from k. Furthermore, at the third layer reside copies $i^{(3)}$ of all stops $i \in \mathcal{W}$ of the existing network. Finally, the copy $l^{(4)}$ of the unique destination vertex $l \in \mathcal{V}$ is used at the fourth layer as the network sink.

Table 2 displays the arcs of the pricing digraph for $(k, l) \in OD$. Each elementary $k^{(1)}$-$l^{(4)}$-path P in this digraph corresponds with a trip from k to l. The computation of a path's travel time and reduced cost relies on the arc travel times and reduced costs shown in the last two columns of Table 2. Both attributes of a path are however not necessarily given by accumulation of the presented values. We discuss details on the computations as well as the labeling algorithm for the actual solution of the (k, l)-specific subproblem in the following paragraphs.

Arc/between Layers	Description	For	Reduced Cost	Travel Time
$(k^{(1)}, i^{(2)})$ 1 → 2	Access leg	$(k, i) \in \mathcal{E}(\mathcal{W}, \mathcal{V})^{walk}$ $i \in \mathcal{T}_{z_k}$ and $(k, i) \in \mathcal{E}(\mathcal{W}, \mathcal{V})^{MoD}$	$-\gamma^{kl}$ $-\gamma^{kl} + c^{ki} d^{kl} - \epsilon_i^{kl}$	t_{ki}
$(i^{(2)}, i^{(3)})$ 2 → 3	First boarding in fixed network	$(k, i) \in \mathcal{E}(\mathcal{W}, \mathcal{V})$	0	t_w
$(i^{(2)}, l^{(4)})$ 2 → 4	Egress leg for trips w/o fixed public transport legs	$(i, l) \in \mathcal{E}(\mathcal{W}, \mathcal{V})^{walk}$ $i \in \mathcal{T}_{z_l}$ and $(i, l) \in \mathcal{E}(\mathcal{W}, \mathcal{V})^{MoD}$	0 $c^{li} d^{kl} - \zeta_i^{kl}$	t_{il}
$(i^{(3)}, j^{(3)})$ 3 → 3	Travel in fixed network	$(i, j) \in \mathcal{E}(\mathcal{W}) \setminus \bigcup_{s \in \mathcal{S}} s$ $(i, j) \in s$ for some $s \in \mathcal{S}$	0 $-\delta_s^{kl}$	t_{ij}
$(j^{(3)}, l^{(4)})$ 3 → 4	Egress leg	$(j, l) \in \mathcal{E}(\mathcal{W}, \mathcal{V})^{walk}$ $j \in \mathcal{T}_{z_l}$ and $(j, l) \in \mathcal{E}(\mathcal{W}, \mathcal{V})^{MoD}$	0 $c^{lj} d^{kl} - \zeta_j^{kl}$	t_{jl}

Table 2: Arcs of the pricing digraph for OD-pair $(k, l) \in OD$ with their reduced costs and travel times

Travel Time Computation. The travel time of a path can either be computed straightforwardly as the sum of the path's arcs travel times (as given in Table 2) or with more detailed network if one wants to integrate expected transfer times (from MoD to public transport [or vice versa] or between different buses in the public transport network).

For the straightforward approach, the travel time information comprises the following elements: First, the time t_{ki} for the access leg $(k, i) \in \mathcal{E}(\mathcal{W}, \mathcal{V})$, which is either given by the MoD travel time including the initial waiting or by the walking time to the station. Second, the time t_w for waiting before initial boarding of a bus, which depends on the frequency of the bus service as well as on the harmonization of the MoD operations with the bus timetable. Third, bus travel time t_{ij} on the edge $(i, j) \in \mathcal{E}(\mathcal{W})$, which also includes the planned stopping time of the bus. Fourth, the time t_{jl} for the egress leg $(j, l) \in \mathcal{E}(\mathcal{W}, \mathcal{V})$, which again either represents MoD travel time or walking time. Note that the above parameters can also be *weighted* times in case the levels of inconvenience perceived by the passengers should be considered.

For a more detailed travel time model, passenger transfer times should be included. In this case, the pricing network needs to contain copies of each stop vertex, one for modeling the access of a certain bus line and one for passing through a stop with a bus of that line. Transfer times can then be added to arcs ending at an access copy of a stop. An approximated value for the transfer time can be estimated on the basis of the frequencies of the corresponding lines. Again, the usual weighting approaches could be applied to reflect the fact that transferring time is perceived to be more inconvenient than in-vehicle time. We have not implemented pricing with the more detailed travel time network and we expect that the difficulty of this problem increases significantly as also discussed in Borndörfer *et al.* (2007, p. 126).

Reduced cost computation. The reduced cost $\tilde{c}^P$ of a path P as defined in (2) contains δ_s^{kl} for an omission segment s only once, independent of how many times (but at least once) arcs of that segment are traversed. The consequence is that $\tilde{c}^P$ is in general not given by the sum of the arc reduced costs (as one could interpret the reduced cost information in Table 2). There are two options now to cope with this complication:

SPPRC-1: Introduction of additional resources, one for each omission segment $s \in \mathcal{S}$

SPPRC-2: Modification of the constraints (1d)

In the first case, we introduce additional binary resources for each omission segment $s \in \mathcal{S}$. The resource for s indicates that the segment s has not been visited by the partial path (=1). The resource is initialized with 1. Propagation sets them to 0 once an edge of the corresponding segment is traversed for the first time. (Implications for cost propagation and dominance in labeling are discussed below.)

In the second case, we replace constraints (1d) by the weaker version

$$\sum_{P \in \mathcal{P}^{kl}} n_s^P \cdot \lambda_P^{kl} \leq |s| \cdot x_s, \qquad \forall s \in \mathcal{S}, \forall (k,l) \in OD, \tag{1d'}$$

where n_s^P is the number of edges in s that are used by P, and $|s|$ is the total number of edges in s. The consequence of this replacement is that now

$$\tilde{c}^P = \left(c^{k i_P} + c^{l j_P}\right) d^{kl} - \gamma^{kl} - \sum_{s \in \mathcal{S}} n_s^P \delta_s^{kl} - \epsilon_{i_P}^{kl} - \zeta_{j_P}^{kl} \tag{2'}$$

allowing to add up the reduced costs of all arcs to obtain the correct reduced cost of P.

While the first option provides a tighter linear programming (LP) relaxation of (1) compared to the one where (1d) is replaced by (1d'), it however requires additional binary resources that slow down the solution of the pricing problem. We analyze these effects in our computational study in Section 5.

Labeling algorithm. The pricing problems, one for each $(k,l) \in OD$, are solved with an SPPRC labeling algorithm. We use the following attributes:

- T^{cost}: the accumulated reduced cost;
- T^{time}: the accumulated travel time;
- $T^{\text{om-seg},s}$: for each $s \in \mathcal{S}$, a binary attribute indicating that omission segment s has not been visited.

While SPPRC-1 uses all attributes, SPPRC-2 uses only the first two. In both cases, the initial label at the start vertex $k^{(1)}$ is $(T^{cost}, T^{time}, (T^{\text{om-seg},s})) = (0, 0, (1))$.

For the label extension, we assume that for each arc (v, w) the reduced cost is $\tilde{c}_{vw}$ as given in Table 2 and the travel time is t_{vw}. In SPPRC-1, the resource extension functions for propagating the attributes along an arc (v, w) are:

$$T_w^{cost} = T_v^{cost} + \begin{cases} 0, & \text{if } (v,w) = (i^{(3)}, j^{(3)}) \in s \text{ for some } s \in \mathcal{S} \\ & \text{and } T_v^{\text{om-seg},s} = 0 \\ \tilde{c}_{vw}, & \text{otherwise} \end{cases} \tag{3a}$$

$$T_w^{time} = T_v^{time} + t_{vw} \tag{3b}$$

$$T_w^{\text{om-seg},s} = \begin{cases} 0, & \text{if } (v,w) = (i^{(3)}, j^{(3)}) \in s \\ T_v^{\text{om-seg},s}, & \text{otherwise} \end{cases} \qquad \text{for all } s \in \mathcal{S} \tag{3c}$$

The new label T_w is feasible if

$$T_w^{time} \leq T^{max} \tag{4}$$

where T^{max} is the upper bound on the journey time. One other helpful observation is that only the arcs $(k^{(1)}, i^{(2)})$ for access legs (from the first layer to the second) can have a negative reduced cost, since the dual prices fulfill $\delta_s^{kl}, \epsilon_i^{kl}$ and $\zeta_j^{kl} \leq 0$. Therefore, labels with positive aggregated reduced costs can be discarded, i.e., the condition $T_w^{cost} < 0$ can be added to (4).

In comparison, the propagation within SPPRC-2 is purely additive, i.e., $T_w^{cost} = T_v^{cost} + \tilde{c}_{vw}$ and $T_w^{time} = T_v^{time} + t_{vw}$. The only feasibility condition is again (4).

Dominance is a key component of SPPRC labeling algorithms. For the weaker model formulation, i.e., SPPRC-2, a label T_1 dominates another label T_2 (both must be resident at the same vertex), if it is not greater in its reduced cost and accumulated travel time attributes. A dominated label can be discarded if the dominating label is kept. For the tight formulation with additional resources, i.e., SPPRC-1, the dominance condition is

$$T_1^{cost} - \sum_{\substack{s \in \mathcal{S}: T_1^{\text{om-seg},s}=1, \\ T_2^{\text{om-seg},s}=0}} \delta_s^{kl} \leq T_2^{cost} \qquad \text{and} \qquad T_1^{time} \leq T_2^{time}. \tag{5}$$

The first condition allows label T_1 to dominate label T_2 even if it is inferior regarding the visits of some omission segments $s \in \mathcal{S}$. However, the reduced cost T_1^{cost} must then be even smaller compared to T_2^{cost} (note that $\delta_s^{kl} \leq 0$ holds). Such an improved dominance relation was coined by Jepsen *et al.* (2008) in the context of vehicle routing.

In addition, it proved beneficial to apply a partial pricing approach (see, e.g., Gamache *et al.*, 1999) that does not solve the pricing problem for each OD-pair $(k, l) \in OD$ in every iteration, but to only consider a pair (k, l) if a path from or to either zone z_k or z_l has been found in the previous iteration. This does not change the algorithm in the initial iterations, as usually paths are found for every zone, but it often significantly accelerates the final iterations. Clearly, all pricing problems need to be solved again once partial pricing failed to identify any negative reduced cost column.

Finally, for the solution of the SPPRC subproblem, we also implemented a bidirectional labeling algorithm (Righini and Salani, 2008). Compared to the discussed monodirectional labeling, pretests did not reveal a significant improvement, neither for SPPRC-1 nor SPPRC-2. This may result from the hard time bounds and the rather restricted options to route from k to l. Therefore, the computational results section presents only results for monodirectional labeling.

4.2.2. *Branching strategy*

We use the following two-level branching scheme that introduces branching decisions in decreasing order of expected significance for the model. At the first level, the binary segment variables x_s are fixed, as these variables predetermine which paths through the network are possible and where MoD is necessary. Priority is given to those segment variables x_s with the largest flow passing through s in today's network.

At the second level, omission segments are fixed, and we branch on the binary MoD variables y_{zI}. The variable selection rule prioritizes MoD variables y_{zI} with value closest to 0.5.

The RMP with fixed x_s and y_{zI} decomposes into independent network flow problems, one for each OD-pair. Therefore, no branching on the λ_P^{kl} variables is required. Since we use the simplex algorithm for solving the RMP, no fractional solutions occur, even when two paths have identical costs.

The branching decisions impact the pricing subproblems in a straightforward way. All decisions can be enforced by removing certain edges from the pricing network. On top of this, we can even abstain from pricing for the OD-pairs that can use the shortest path in today's network, where this network is imposed by the fixed network and the omission segments fixed to one.

Finally, our branch-and-bound node selection strategy is the best-node-first strategy.

4.3. *Enhanced enumeration algorithm*

The number of feasible OD-paths $\mathcal{P}^{kl}$ very much depends on the instance structure. For example, in a perfect star-shaped network, the number of connecting paths between k and l is given by $|\mathcal{E}_k(\mathcal{W},\mathcal{V})| \cdot |\mathcal{E}_l(\mathcal{W},\mathcal{V})|$, where $\mathcal{E}_k(\mathcal{W},\mathcal{V}) \subset \mathcal{E}(\mathcal{W},\mathcal{V})$ and $\mathcal{E}_l(\mathcal{W},\mathcal{V}) \subset \mathcal{E}(\mathcal{W},\mathcal{V})$ denote the set of access and egress legs (k,i) and (j,l) for fixed k and l, respectively. This number is typically small. Hence, we implemented an enumeration-based solution approach that solves model (1) directly with an IP solver.

For each $(k,l) \in OD$, we explicitly construct all possible connecting paths with the help of the pricing network developed in Section 4.2. The actual enumeration results from omitting any dominance between labels.

As one may expect, this basic enumeration approach can lead to memory issues already for small-sized instances and thus needs further enhancement. Indeed, the straightforward labeling does not take into account that passengers choose shortest paths for the fixed-route segment of their journey between potential MoD access and egress legs. This property can be translated into the following dominance relation: One label can dominate a second one if it has a lower accumulated travel time and fulfills two more conditions. First, both labels have visited the same transfer points. Second, the first label must have visited a subset of the omission segments compared to the second label. The latter condition is important in the context of branching on omission segments, because we do not know in advance which segments are finally available. Indeed, with these dominance conditions, any extension of the partial path corresponding to the first label is feasible whenever the same extension is also feasible for the partial path of the second label.

5. Computational results

In this section, we present the computational study based on model (1) and the solution algorithms of Sections 4.2 and 4.3. Section 5.1 gives details on the instances used for the calculations. Subsequently, we systematically determine the most favorable parameter settings for the optimization and provide a comparison of outputs for the branch-and-price and enhanced enumeration algorithms in Section 5.2. Furthermore, we assess the required granularity level for the decision variables in Section 5.3 and interpret selected model outputs in Section 5.4. A sensitivity analysis with respect to the key input factors is presented in Section 5.5.

5.1. Instances

The instances for the computational study are based on real-world data of a mid-sized German city provided by the `LinTim` software (see, e.g., Goerigk *et al.*, 2013), which includes vertices and edges of the existing public transport network with around 20 bus lines and demand data per OD-pair for an average hour. We consider mid-sized cities with bus networks to be a realistic application for our model, because these networks often face financial pressures due to utilization issues, in particular in low density areas and off-peak hours.

For our calculations, we have slightly adjusted the data by replacing (1) the few directed arcs by undirected edges and (2) the group of very close central stops by a single vertex representing the (historic) city center. The latter replacement helps to reduce the network and also makes sense from a practical perspective, as passengers typically choose to walk to their final destination from the most convenient stop of the bus route they are currently using instead of transferring again to another bus line to ride a very short distance.

The given demand is almost symmetric. Therefore, we opted for unidirectional OD-pairs $(k, l) \in OD$ by aggregating the demand for both directions (k, l) and (l, k) in the parameter d^{kl}.

In the `LinTim` datasets, we also found a one-to-one correspondence between the physical locations of OD-vertices $\mathcal{V}$ and public transport stops $\mathcal{W}$: Demand is not available on a more detailed level than the current bus stops. Walking legs between the origin (destination) of the trip and the first (last) public transport stop have still been considered in the travel times as displayed in Table 2.

After these adjustments, the network comprises 238 vertices, 270 edges, 49,449 OD-pairs with non-zero demand, and 405,767 passengers. On this basis, we have generated three sets of instances:

- A district of the city with 30 vertices, 33 edges, and 5,260 passengers to simulate a small city with low demand levels, where it can even be questioned whether a fixed-route network is required at all.
- The city center and selected lines to the outer districts with 145 vertices, 158 edges, and 164,950 passengers to simulate a network with a shape close to a star, where among others the effects of improved linking of the 'rays' can be analyzed.
- The whole network to assess the holistic impact on a city level.

Our model further requires omission segments as well as potential MoD zones and transfer points as additional input parameters. Since these elements do not exist in today's network, we have determined them manually by splitting the city into its main districts and selecting appropriate key stops for the transfer points. Three different setups have been established per instance set for the omission segments and two setups for the zones and transfer points, respectively. Altogether, this gives twelve instances for each of the three instance sets. We have also included additional constraints on the x_s variables to ensure the resulting network is a connected graph.

As the network still resembles a star-shaped network with a limited number of reasonable paths between most of the OD-pairs, we have generated another set of instances to observe how the model and the different solution approaches behave in a different context. We have chosen a quadratic grid, as it abstracts the street-network of a city and offers a multitude of reasonable connections between the majority of OD-pairs. The grid dimension varies between 6×6 and 10×10, where the edge of a grid cell has 1 kilometer (km) length. Potential MoD zones are represented by 3×3

sub-grids and the potential transfer points per zone are given by the central point of the zone itself and its neighboring zones, respectively. Identical demand between every pair of vertices has been assumed. Two setups for the omission segments have been chosen: The first one has two omission segments per horizontal and vertical line, while the second one represents the extreme case where every edge is a separate omission segment. This last case is certainly not a realistic setup. However it is interesting from a performance perspective to see how the different approaches deal with a setup that allows for a maximum number of possible resulting network structures.

		Number of			
Instance set	#	vertices	omission segments	MoD zones	distinct transfer points
District	12	30	4, 9, 13	4, 8	5, 9
Selected lines	12	145	24, 29, 36	22, 28	31, 41
Full network	12	238	27, 37, 48	24, 37	37, 53
Grid $d \times d$, $d = 6, \ldots, 10$	10	d^2	$4(d-1)$, $2(d-1)d$	$\lceil d/3 \rceil^2$	$\lceil d/3 \rceil^2$

Table 3: Omission segment and MoD setups per instance set

An overview of all instances and their characteristics is provided in Table 3. A visualization of the full network for the setup with 27 omission segments, 24 MoD zones, and 37 transfer points is shown in Figure 1.

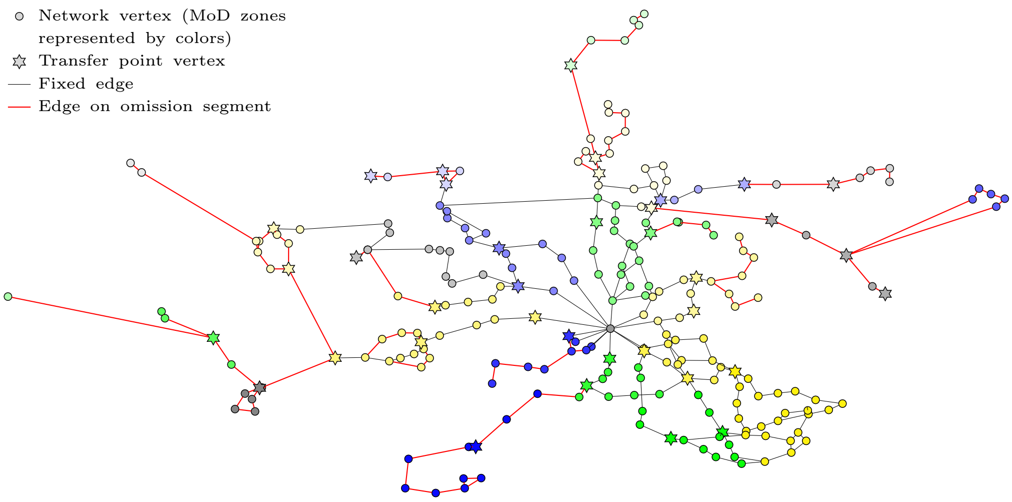


Figure 1: Full network instance with MoD zones and omission segments

We make the following assumptions about the input data (a sensitivity analysis for the key parameters is presented in Section 5.5): Base demand d^{kl} is given in the `LinTim` data, for the induced demand d_{ziI} we assume a start level of 15% of the base demand in the given zone and adjust it based on the transfer point set I. Here, more transfer points are assumed to increase the level of induced demand.

The costs for the fixed-route network are calculated based on costs of 5 € per km that were derived from the annual report of the corresponding public transport operator. MoD costs are certainly the most uncertain input as they depend heavily on how this service is delivered, which service level is aspired etc. As a starting point, we assume 1 € per passenger km and adjust this value based on the demand level in the corresponding zone, hereby decreasing demand levels increase the MoD costs. Costs of around 1 € per passenger km are in the range of current taxi costs when 2 to 3 passengers are on board. An interesting benchmark for MoD costs in a driverless setup is given by current pricing for car sharing, as the costs mainly reflect on providing and managing a fleet of vehicles in both cases. Car sharing costs are in the range of 30-40 € cents per minute in larger German cities, which should have a significant potential to decrease once pooling is introduced and systems grow in scale. Fixed costs for MoD are small in the basic setup as we assume transfer points will be chosen at stops that provide sufficient infrastructure, and also as vehicle costs can be included in the variable costs. A conservative additional revenue of 50 € cent per additional passenger are assumed considering that the MoD pricing mechanism is unknown.

In terms of travel times, we assume average travel speeds of 20 km/h for buses, 25 km/h for MoD vehicles, dwell times of 20 seconds, and access, egress, and waiting times (both for buses and MoD) of 5 minutes, respectively. The parameter θ indicating the maximum allowed increase in total travel time has been set to 20 %.

The preprocessing reduces the number of OD-pairs to consider and thus the number of pricing problems to be solved. We present an example of this reduction for the instance covering the full network with 27 omission segments, 24 MoD zones, and 37 transfer points. Treating paths from k to l jointly with paths from l to k reduces the number of pairs with non-zero demand from 49,449 to 25,359. When factoring out the stable part of the network, where passengers are always routed through the fixed-route edges of the network, we have 15,760 OD-pairs and the number of passengers reduces from 405,767 to 204,009. Subsequently, the aggregation step depicted in Section 4.1 reduces the number of OD-pairs to be considered to 3,031, i.e., on average 5 OD-pairs can be grouped together and their passenger routes determined collectively.

5.2. Technical aspects

All computational tests are performed on a standard PC with an Intel(R) Core(TM) i7-5930 running at 3.5 GHz with 64 GB of main memory using a single thread. Algorithms are coded in C++ using CPLEX 12.7 and compiled in release mode with MS Visual Studio 2015. The time limit for all runs is 1 hour (3600 seconds).

The RMP is initialized with two sets of columns. First, we insert the status-quo paths through the fixed-route network. Second, we insert the MoD-only paths P^{kl}_{dir} for all $(k,l) \in OD$.

Our key objective in this section is understanding the performance of the two branch-and-price algorithms (with the weak and strong formulation) and enhanced enumeration approach. Tables 4 and 5 show computation times and technical characteristics. The columns have the following meaning:

- #Solved tree: Number of instances for which the complete branch-and-bound (B&B) tree was solved;
- #Solved root: number of instances for which the root node (linear relaxation) was solved;
- Comp. time: overall computation time;
- Subp. time: percentage of the overall computation time spent in the subproblem, i.e., pricing time in branch-and price and enumeration time for the enhanced enumeration approach;
- B&B nodes: number of branch-and-bound nodes solved within the time limit;
- RMP iter.: number of (partial) pricing iterations;

Columns gener.: number of generated columns;
Columns: number of columns per OD-pair;
Improv. to status quo: Percentage cost reduction compared to the status quo.

All numbers in Table 4 starting from column *Comp. time* are averages over the six instances considered in the respective row. We omit the instances with 30 vertices, since nearly all instances are already solved to optimality in the root node. Also, the very difficult instances with every edge representing an omission segment are not displayed for the 8×8 and larger grids as not even the root node was solved in these cases.

Nb. of vertices	Setup	Algorithm	# Solved tree	# Solved root	Comp. time	Subp. time	B&B nodes	RMP iter.	Columns gener.	Columns per OD	Improv. to status quo
145	low nb. of transfer points	enh. enum.	6/6	6/6	30	27%	5.7	1	41571	15.2	9.5%
		BaP + strong	6/6	6/6	243	75%	5.7	206	13713	5.0	9.5%
		BaP + weak	4/6	6/6	2762	66%	526.2	1477	20332	7.4	9.5%
	high nb. of transfer points	enh. enum.	4/6	6/6	1817	2%	67.7	1	181640	37.6	10.0%
		BaP + strong	1/6	3/6	3491	45%	28.7	763	49740	10.5	6.7%
		BaP + weak	0/6	6/6	3600	30%	58.8	931	46926	9.9	n.a.
238	low nb. of transfer points	enh. enum.	6/6	6/6	78	50%	2.3	1	101519	19.0	3.4%
		BaP + strong	6/6	6/6	1354	88%	2.3	176	23133	4.4	3.4%
		BaP + weak	0/6	6/6	3600	82%	72.7	694	28911	5.6	n.a.
	high nb. of transfer points	enh. enum.	4/6	6/6	1302	24%	4.2	1	404122	43.4	3.4%
		BaP + strong	2/6	2/6	3320	59%	0.3	167	73563	8.2	3.4%
		BaP + weak	0/6	2/6	3600	42%	1.0	321	52317	5.7	n.a.

Table 4: Results for the real-world instances

As for the algorithmic performance, the enhanced enumeration performs consistently superior to the branch-and-price approaches on the real-world networks. This is due to the network structure being similar to a star and therefore allowing for a limited number of possible paths per OD-pair, see *columns per OD-pair* in Table 4. While the enhanced enumeration approach could not solve the problem to optimality in four cases as well, the root node is solved even for the hardest instances.

For the grid instances, a reverse behaviour can be observed in Table 5. The higher number of possible paths causes significantly long calculation times. In several cases, the total number of columns becomes too large for using the MIP solver directly. Conversely, the column-generation algorithms are faster. For the low number of omission segments, we can compute average speedup factors of 3.4, 7.1, and 4.1 for the 6×6, 7×7, and 8×8 instances, respectively. For the high number of omission segments the 6×6 instance is even solved 60 times faster than by enhanced enumeration. Finally, the 7×7 instance could be solved to optimality in 16.4 hours by column generation, which would still be sufficient for a strategic setup, whereas the enumeration could not tackle this instance any more and did not even finish the enumeration step of all possible paths.

Regarding the two alternative formulations, the tight model is by far superior to the weaker model, mainly because branching trees are much smaller. The only exception is the extreme case of the grid network, where each edge represents its own omission segment, as in this case the two models are equivalent.

5.3. Granularity of decision variables

In this section, we analyze and quantify the impact of different decision variable setups on the resulting network and objective value. This helps address the tradeoff between a more fine-grained

Grid	Setup	Algorithm	# Solved tree	# Solved root	Comp. time	Subp. time	B&B nodes	RMP iter.	Columns gener.	Columns per OD
6 × 6	low nb. of omission segments	enh. enum.	yes	yes	5.4	58%	1.0	1	37530	59.6
		BaP + strong	yes	yes	1.6	78%	1.0	19	5845	9.3
		BaP + weak	yes	yes	3.4	36%	1.0	33	9646	15.3
	high number of omission segments	enh. enum.	yes	yes	420.8	96%	1.0	1	87418	138.8
		BaP + strong	yes	yes	7.0	61%	1.0	53	9966	15.8
		BaP + weak	yes	yes	4.0	34%	1.0	38	10547	16.7
7 × 7	low number of omission segments	enh. enum.	yes	yes	206	11%	3.0	1	118003	100.3
		BaP + strong	yes	yes	29	36%	3.0	73	18913	16.1
		BaP + weak	yes	yes	1964	3%	71.0	723	50275	42.8
	high number of omission segments	enh. enum.	no	no	3600	100%	0.0	1	0	0.0
		BaP + strong	no	no	3600	2%	0.0	172	42495	36.1
		BaP + weak	no	no	3600	0%	0.0	121	40772	34.7
8 × 8	low number of omission segments	enh. enum.	yes	yes	2224	7%	5.0	1	324276	160.9
		BaP + strong	yes	yes	547	14%	5.0	252	45718	22.7
		BaP + weak	no	yes	3600	3%	35.0	517	75607	37.5
9 × 9	low number of omission segments	enh. enum.	no	no	3600	31%	0.0	1	1074224	331.6
		BaP + strong	no	no	3600	2%	0.0	100	76947	23.7
		BaP + weak	no	yes	3600	3%	10.0	241	89476	27.6
10 × 10	low number of omission segments	enh. enum.	no	no	3600	100%	0.0	1	0	0.0
		BaP + strong	no	no	3600	2%	0.0	51	88601	17.9
		BaP + weak	no	yes	3600	4%	3.0	165	102622	20.7

Table 5: Results for the grid instances

modeling, which should lead to better objective values, solution effort, and practical considerations. Table 6 shows the average improvement of the objective value once we have decision variables that are more detailed with respect to segments, zones, and transfer points. The improvement is calculated by a pairwise comparison of the less granular instances (with respect to the aspect in focus) and their more granular counterparts. We finally take the average over all instances. Furthermore, the impact on the computation time for the enhanced enumeration algorithm is shown, e.g., smaller zones yield 0.8 % better results with 19 % longer computation times for the instances with 30 vertices.

Instance set	Change in objective and comp. time through					
	more segments		more zones		more transfer points	
Nb. of vertices	cost	comp. time	cost	comp. time	cost	comp. time
30	−65.4%	+217%	−0.8%	+19%	−3.1%	+860%
145	±0.0%	+194%	−2.0%	+99%	−1.3%	+7 679%
238	±0.0%	+201%	−2.9%	+18%	±0.0%	+428%

Table 6: Impact of decision variable setup on objective value and calculation time

We observe that the transfer point setup has the highest impact on calculation times, while only leading to smaller improvements in the objective value, in particular for the larger-scale instances. Accordingly, we recommend to reduce the number of pre-selected transfer points if computation times are too high. Moreover, a sequential approach could be applied, where first the model is rerun with an increased number of possible transfer points for the zones for which MoD has been offered. In a second step, these zones could be analyzed separately.

The prominent 65.4 % of improvement for the instances with 30 vertices is due to low demand of these instances, so that the optimal solution comprises no fixed-route network at all. Therefore, the first segment setup that still includes certain fixed-route segments is noticeably more expensive. Thus, when modeling smaller networks with lower demand levels, e.g., when analyzing the local transport network within a district, it is advisable to allow all edges to be removed. This solution is purely based on MoD and could be the most beneficial one.

When concentrating on the larger-scale instances, the number of omission segments shows similar behavior as the transfer points, albeit the impact on computation times is much smaller. Finally, modeling smaller zones yields good results in terms of improvement and a reasonable increase in the computation times. We can thus recommend to use very fine-grained zones as long as this is still reasonable from an implementation perspective.

5.4. Interpretation of results

For the following series of experiments we choose the full network depicted in Figure 1 with the same setup with respect to zones, omission segments, and transfer points. The base demand has been reduced slightly by a factor of 0.7, because this setup provides a couple of interesting insights on the variety of passenger routes. This setup is our *base scenario* also used for the sensitivity analysis in Section 5.5.

The optimal solution for the base scenario is shown in Figure 2. Overall, 19 of the 27 omission segments are still included in the solution, ten zones are chosen for MoD operations, and nine distinct transfer points are selected, two of which serve two zones simultaneously.

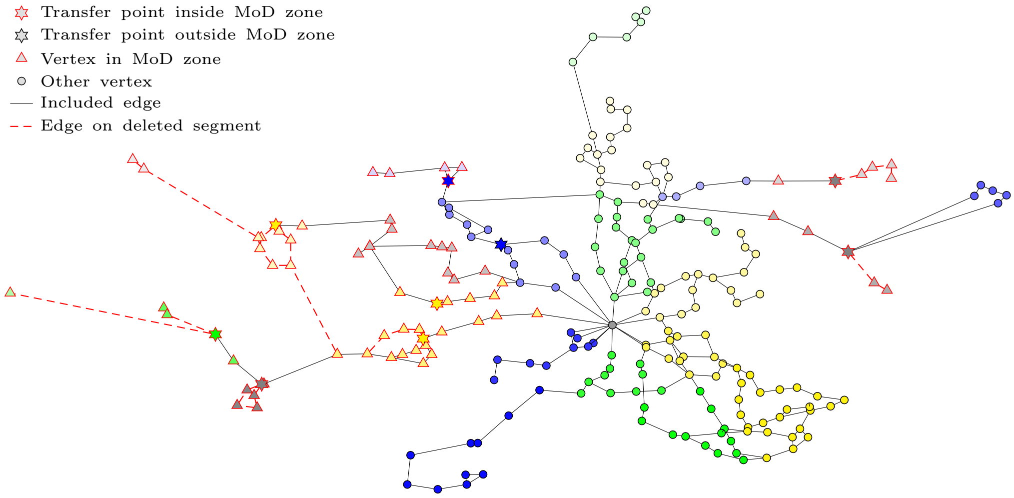


Figure 2: Solution of full network instance with reduced base demand

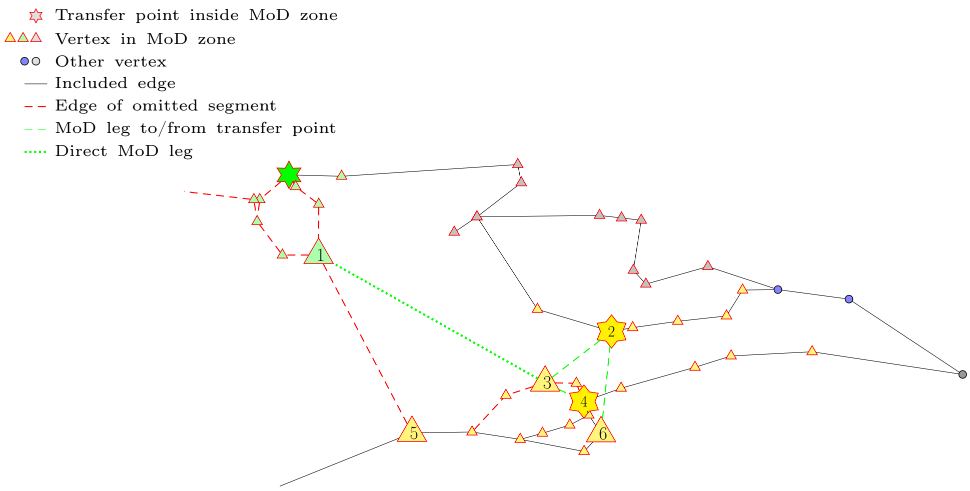


Figure 3: Selected passenger flows for full network instance with reduced base demand

Figure 3 shows a magnification of the western part of the city, visualizing some passenger routes of the optimal solution. As it is not possible to show all passenger routes at the same time, we present four selected MoD legs for a zone (the yellow vertices in the figure) with two distinct transfer points 2 and 4. It can be seen that passengers from the same origin or destination may use different transfer points depending on the other end of their journey, e.g., some passengers starting at vertex 3 use transfer point 2, others use transfer point 4. Also, vertex 6 uses MoD even though it is still connected to the remaining network. These MoD legs are used for connections to the vertices in the upper left of the graph, where the original connection edge $(1, 5)$ in the public transport is omitted. Furthermore, a direct MoD connection has become necessary to travel between vertex 1 and vertex 3. These vertices were originally connected via a (convenient) route with just four stops. This cannot be adequately replaced by routes using the two MoD transfer points with the allowed increase of travel time (here $\theta = 20\%$). Therefore, direct MoD is chosen by the model even though the costs are higher than the MoD costs for access and egress legs due to the decreased consolidation potential.

5.5. Sensitivities

The final set of experiments investigates the sensitivity with respect to the main parameters, i.e., changes in base demand (by scaling the standard demand by a factor), induced demand levels, MoD costs, and the parameter θ.

The impact on objective value, i.e., costs, the number of selected segments, MoD zones, and the total number of transfer points can be seen in Figures 4a–4d. Figure 4a shows the potential of MoD to reduce system costs when demand decreases. In contrast, the main cost reduction possibility for a classical fixed network is increasing headways and thereby making the system less attractive. Additionally, it is worth noting that costs improve again slightly for higher base demand factors due to the increased induced demand.

The sensitivity with respect to induced demand in Figure 4b shows alternating behavior regarding segments, zones, and transfer points. The reason is that there exist multiple solutions of very similar quality, because additional induced demand creates both additional costs and revenues.

In Figure 4c, we observe the expected strong impact of MoD costs on total costs. This result underlines the necessity to obtain a good understanding of the planned MoD system before employing a strategic model like the one we present.

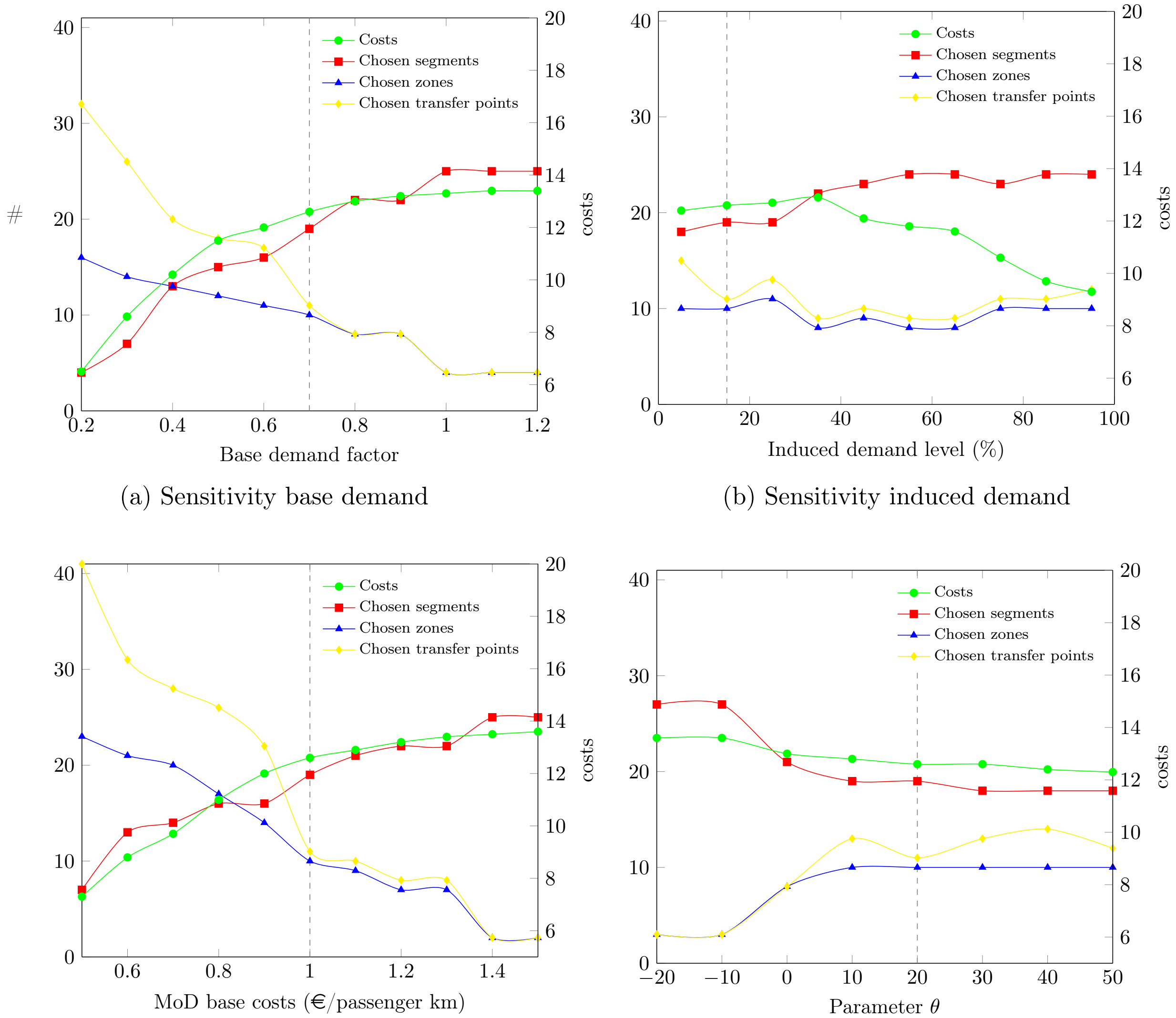
Finally, increasing the parameter θ has the expected impact of reducing costs and allowing for more MoD zones and transfer points as displayed in Figure 4d.

6. Conclusion and outlook

We have presented a strategic network planning optimization model for buses that allows for intermodal trips with MoD as a first or last leg. The model captures important aspects such as dynamic demand and passenger routing, and is the first model in the literature covering all these aspects. Branch-and-price algorithms and an enhanced enumeration approach have been developed able to solve realistic instances to optimality in reasonable computation times.

Considering this is an early step for the integrated planning of intermodal networks including MoD, further research is necessary with respect to multiple modeling and algorithmic aspects. The model objective can be extended to explicitly cover the impact on travel times and consider the cost-service tradeoff. Furthermore, MoD-related aspects can be modeled with a higher level of detail, e.g., a linking with more sophisticated demand models that determine the induced demand on an OD-level. This could improve applicability in rural areas, where current long and inflexible travel times deter the majority of people from using public transportation and where there is high potential for induced demand as a consequence. Also, the costs for the MoD trips could be based more explicitly on demand density instead of preprocessed demand levels as used in this paper. As for the fixed-route network, frequencies could be included to better model the total system costs, in particular in the case of binding capacity restrictions.

Obviously, including these aspects brings challenges for the solvability of the model: Path-specific demand d_P^{kl} would require a different realization of the pricing networks, where the MoD cost contribution is independent of the path. In addition, a more detailed modeling of demand density creates additional indices for the variable feeding costs c^{ki} reflecting the demand level of the zone etc.



(a) Sensitivity base demand

(b) Sensitivity induced demand

(c) Sensitivity MoD costs

(d) Sensitivity parameter θ

Figure 4: Sensitivity analysis

Regarding refined solution approaches, hybrid exact algorithms could be based on the combination of the presented branch-and-price and enumeration algorithms. One may first check whether an enhanced enumeration is feasible and otherwise apply the branch-and-price approach. Alternatively, one may treat some selected OD-pairs with many potential paths via pricing and others with the enumeration approach. On the heuristic side, it could be beneficial to first prioritize, e.g., the largest 20 to 30% of OD-pairs that typically represent some 60 to 80% of the total demand, and to determine the network setup based on these. In a second step, the remaining OD-pairs could then be added and further MoD connections supplemented in case the travel time constraint is violated.

We believe that the integration of MoD and public transport will stay a focal topic taken up by public authorities, transport as well as technology companies, and users. From the research perspective, our paper provides a solid foundation for further optimization-based decision support in this field.

Acknowledgement

We would like to thank Prof. Anita Schöbel (University of Göttingen) for the inspiring discussions and providing us access to the `LinTim` data set before it became publicly available.

References

Agatz, N., Erera, A., Savelsbergh, M., and Wang, X. (2012). Optimization for dynamic ride-sharing: A review. *European Journal of Operational Research*, **223**(2), 295–303.

Aldaihani, M. M., Quadrifoglio, L., Dessouky, M. M., and Hall, R. (2004). Network design for a grid hybrid transit service. *Transportation Research Part A: Policy and Practice*, **38**(7), 511–530.

Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., and Rus, D. (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences*, **114**(3), 462–467.

Archetti, C., Speranza, M. G., and Weyland, D. (2017). A simulation study of an on-demand transportation system. *International Transactions in Operational Research*, **25**(4), 1137–1161.

Borndörfer, R., Grötschel, M., and Pfetsch, M. E. (2007). A column-generation approach to line planning in public transport. *Transportation Science*, **41**(1), 123–132.

Chowdhury, S. and Ceder, A. A. (2016). Users' willingness to ride an integrated public-transport service: A literature review. *Transport Policy*, **48**, 183–195.

Cordeau, J.-F. and Laporte, G. (2007). The dial-a-ride problem: models and algorithms. *Annals of Operations Research*, **153**(1), 29–46.

Desaulniers, G., Desrosiers, J., and Solomon, M., editors (2005). *Column Generation*. Springer, New York, NY.

Diana, M., Dessouky, M. M., and Xia, N. (2006). A model for the fleet sizing of demand responsive transportation services with time windows. *Transportation Research Part B: Methodological*, **40**(8), 651–666.

Economist (2015). Substitutes or complements? *The Economist*. `https://www.economist.com/blogs/graphicdetail/2015/08/taxis-v-uber`.

Economist (2017). Big tech is under increasing pressure as London bans Uber. *The Economist*. `https://www.economist.com/news/business/21729595-exceptionalism-online-giants-eroding`.

Edwards, D., Elangovan, A. K., and Watkins, K. (2012). Reaching low-density urban areas with the network-inspired transportation system. In *2012 15th International IEEE Conference on Intelligent Transportation Systems*. IEEE.

Farahani, R. Z., Miandoabchi, E., Szeto, W. Y., and Rashidi-Bajgan, H. (2013). A review of urban transportation network design problems. *European Journal of Operational Research*, **229**(2), 281–302.

Friedrich, M. and Noekel, K. (2015). Modeling intermodal networks with public transport and vehicle sharing systems. *EURO Journal on Transportation and Logistics*. DOI: 10.1007/s13676-015-0091-7.

Gamache, M., Soumis, F., Marquis, G., and Desrosiers, J. (1999). A column generation approach for large-scale aircrew rostering problems. *Operations Research*, **47**(2), 247–263.

Goerigk, M. and Schmidt, M. (2017). Line planning with user-optimal route choice. *European Journal of Operational Research*, **259**(2), 424–436.

Goerigk, M., Schachtebeck, M., and Schöbel, A. (2013). Evaluating line concepts using travel times and robustness. *Public Transport*, **5**(3), 267–284.

Häll, C. H., Lundgren, J. T., and Värbrand, P. (2008). Evaluation of an integrated public transport system: a simulation approach. *Archives of Transport*, **20**(1-2), 29–46.

Ho, S. C., Szeto, W., Kuo, Y.-H., Leung, J. M., Petering, M., and Tou, T. W. (2018). A survey of dial-a-ride problems: Literature review and recent developments. *Transportation Research Part B: Methodological*, **111**, 395–421.

Horn, M. E. (2004). Procedures for planning multi-leg journeys with fixed-route and demand-responsive passenger transport services. *Transportation Research Part C: Emerging Technologies*, **12**(1), 33–55.

Irnich, S. and Desaulniers, G. (2005). Shortest path problems with resource constraints. In G. Desaulniers, J. Desrosiers, and M. M. Solomon, editors, *Column Generation*, pages 33–65. Springer-Verlag.

Jaffe, E. (2015). Uber and public transit are trying to get along. *CityLab*.

Jepsen, M., Petersen, B., Spoorendonk, S., and Pisinger, D. (2008). Subset-row inequalities applied to the vehicle-routing problem with time windows. *Operations Research*, **56**(2), 497–511.

Jittrapirom, P., Caiati, V., Feneri, A.-M., Ebrahimigharehbaghi, S., Alonso González, M., and Narayan, J. (2017). Mobility as a service: A critical review of definitions, assessments of schemes, and key challenges. *Urban Planning*, **2**(2), 13–25.

Klier, M. J. and Haase, K. (2014). Urban public transit network optimization with flexible demand. *OR Spectrum*, **37**(1), 195–215.

Li, X. and Quadrifoglio, L. (2010). Feeder transit services: Choosing between fixed and demand responsive policy. *Transportation Research Part C: Emerging Technologies*, **18**(5), 770–780.

Martínez, L. (2015). Urban mobility system upgrade: How shared self-driving cars could change city traffic. Corporate partnership board report, International Transport Forum.

Martínez, L. M., Viegas, J. M., and Eiró, T. (2015). Formulating a new express minibus service design problem as a clustering problem. *Transportation Science*, **49**(1), 85–98.

masabi (2018). Key factors influencing ridership in North America. Mass transit rider research report, Masabi Limited. `http://info.masabi.com/mass-transit-rider-research-report-2018-key-factors-influencing-ridership-in-north-america`.

Mitchell, W. J. (2008). Mobility on demand: Future of transportation in cities. Smart cities, MIT Media Laboratory. `http://smartcities.media.mit.edu/pdf/Mobility_on_Demand_Introduction.pdf`.

Navidi, Z., Ronald, N., and Winter, S. (2017). Comparison between ad-hoc demand responsive and conventional transit: a simulation study. *Public Transport*, **10**(1), 147–167.

Pavone, M. (2016). Autonomous mobility-on-demand systems for future urban mobility. In *Autonomous Driving*, pages 387–404. Springer Berlin Heidelberg.

Posada, M., Andersson, H., and Häll, C. H. (2016). The integrated dial-a-ride problem with timetabled fixed route service. *Public Transport*, **9**(1-2), 217–241.

Righini, G. and Salani, M. (2008). New dynamic programming algorithms for the resource constrained elementary shortest path problem. *Networks*, **51**(3), 155–170.

Rothenbächer, A.-K., Drexl, M., and Irnich, S. (2016). Branch-and-price-and-cut for a service network design and hub location problem. *European Journal of Operational Research*, **255**(3), 935–947.

Salazar, M., Rossi, F., Schiffer, M., Onder, C. H., and Pavone, M. (2018). On the interaction between autonomous mobility-on-demand and public transportation systems. *CoRR*, **abs/1804.11278**.

Schmidt, M. and Schöbel, A. (2015). The complexity of integrating passenger routing decisions in public transportation models. *Networks*, **65**(3), 228–243.

Schöbel, A. (2011). Line planning in public transportation: models and methods. *OR Spectrum*, **34**(3), 491–510.

Shen, Y., Zhang, H., and Zhao, J. (2018). Integrating shared autonomous vehicle in public transportation system: A supply-side simulation of the first-mile service in singapore. *Transportation Research Part A: Policy and Practice*, **113**, 125–136.

Smolnicki, P. M. and Sołtys, J. (2016). Driverless mobility: The impact on metropolitan spatial structures. *Procedia Engineering*, **161**, 2184–2190.

Stiglic, M., Agatz, N., Savelsbergh, M., and Gradisar, M. (2018). Enhancing urban mobility: Integrating ride-sharing and public transit. *Computers & Operations Research*, **90**, 12–21.

Uchimura, K., Takahashi, H., and Saitoh, T. (2002). Demand responsive services in hierarchical public transportation system. *IEEE Transactions on Vehicular Technology*, **51**(4), 760–766.

UITP (2017). Public transport embraces the new mobility scene. Press release, International Association of Public Transport.

VDV (2015). Zukunftsszenarien autonomer Fahrzeuge: Chancen und Risiken für Verkehrsunternehmen. White paper VDV, Verband Deutscher Verkehrsunternehmen e. V. (VDV). (In German) https://www.vdv.de/position-autonome-fahrzeuge.pdfx.

Wang, H. and Odoni, A. (2016). Approximating the performance of a "last mile" transportation system. *Transportation Science*, **50**(2), 659–675.

Willing, C., Brandt, T., and Neumann, D. (2017). Intermodal mobility. *Business & Information Systems Engineering*.